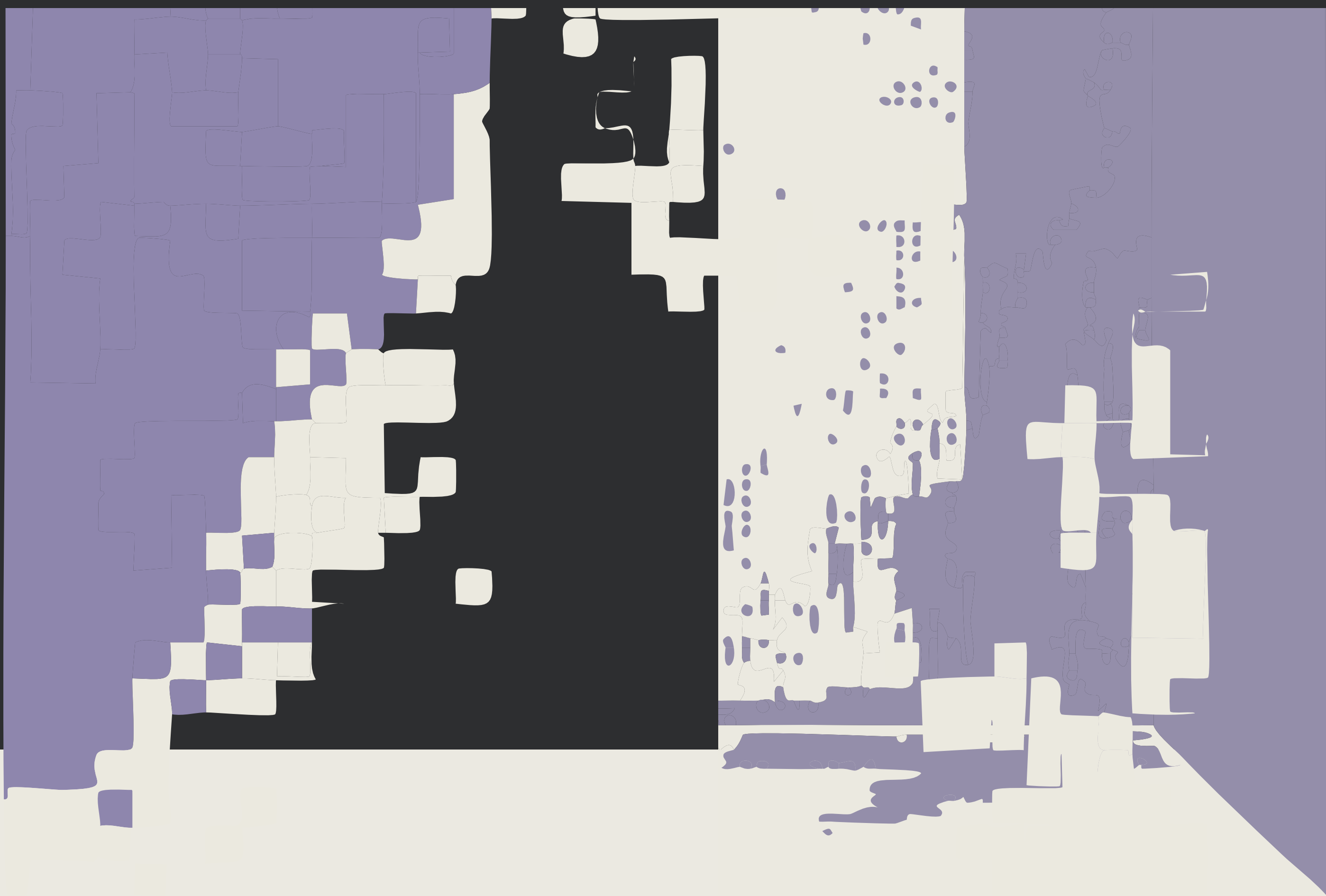
Optimising Crowdfunding Success:

A BOHB-Driven Reward-Tier Strategy for Technology Campaigns



Nathalie Tuijp
University of Twente
Industrial Engineering & Management
08/07/2025

Optimising Crowdfunding Success:

A BOHB-Driven Reward-Tier Strategy for Technology Campaigns

Supervisors:
Dr. Jörg Osterrieder
Dr. Berend Roorda
Ms. Stefana Belbe, PhD Student

Graduate:
Nathalie Tuijp

Programme:
Financial Engineering & Management
Industrial Engineering & Management

University of Twente
08/07/2025

Acknowledgements

This thesis marks the completion of an incredible journey, one that would not have been possible without the support and encouragement of many people in my surroundings. I would like to take this opportunity to express my gratitude to everyone who contributed to helping me finish this thesis.

First and foremost, I am grateful to my supervisors, Jörg Osterrieder, Berend Roorda, and Stefana Belbe, for their support and understanding throughout this process. I truly appreciate the flexibility they afforded me to pursue this research alongside the development of my startup—an unconventional path that I am deeply grateful they allowed me to take. Their guidance and trust in this approach helped me to stay focused and make progress, even when the journey felt challenging.

I am also particularly thankful to Jörg for setting a clear thesis deadline—one that pushed me to stop perfecting and start finishing. His decision to put a date in the calendar was exactly the nudge I needed to move forward and finally deliver a version which contains every Chapter I envisioned. My study advisor Ellen also checked in to make sure I stayed on track and helped me navigate the intricacies of the programme.

I would also like to thank my business partner, Miguel, for understanding that I had to dedicate a bit more time to finishing this thesis. His patience and encouragement during this period made balancing both my startup and my studies possible.

To my dispuut friends, who saw me leave the borrel early more times than I can count but always showed unwavering support—I really appreciate your understanding that kept my spirits high. And of course, a huge thanks to my housemates who, with their endless jokes about being an eternal student, somehow managed to motivate me in their own unique way (I mean they also went to study with me a couple of times, thank you :)

A special mention also goes out to ChatGPT, whose assistance with coding was very helpful, and asking me critical questions when I asked for it. I mean, who is to expect me not leveraging the capabilities of an LLM, while writing about them? I do declare that aside from the coding and a little help for generating elements for my front page, no other direct content produced by AI is presented as my own work.

Finally, I would like to thank my friends and family for their unconditional support and belief in my work. Your encouragement kept me going during the late nights and long weekends, and I am truly grateful to have had you by my side.

Abstract

Crowdfunding platforms like Kickstarter offer opportunities for entrepreneurs and creators to fund their projects. Most of these projects have a reward-tier structure to fund their campaigns. The reward-tier-structure largely influences the outcome of the campaign. Therefore, this thesis aims to optimise this structure by developing a BOHB (Bayesian Optimization and Hyperband) framework with the integration of LLaMA-based embeddings, identifying the most effective reward-tier strategies to enhance campaign success rates. The BOHB framework is specifically chosen, as it is particularly effective for high-dimensional, non-convex search spaces like those found in crowdfunding campaigns. Its adaptive resource allocation and multi-fidelity optimisation allow it to efficiently explore vast parameter spaces, identifying optimal reward strategies with reduced computational cost. To complement traditional numerical features such as funding goals, number of backers, and reward levels, the research integrates LLaMA embeddings which are incorporated in the model to predict the campaign success, giving a higher accuracy to the model. These embeddings capture the semantic richness and emotional tone of campaign descriptions and reward titles. By combining advanced hyperparameter optimisation with the use of LLaMA embeddings, the model identifies optimal reward configurations that enhance the probability of campaign success.
The study uses the publicly available Kickstarter database WebRobots.io, alongside additional data collected by a custom webscraper script. This thesis uses the data in the category Technology from the Kickstarter platform, followed by a data analysis in Python.
The basic RF model scored 0.7428 on accuracy, 0.7169 on precision, 0.7326 on recall, 0.7247 on the F1-Score and 0.8150 on AUC-ROC. The best performing model, the XGBoost BOHB model with all-MiniLM-L6-v2 Embeddings, scored 0.8180 on accuracy, 0.8177 on precision, 0.8182 on recall, 0.8180 on the F1-Score, and 0.9011 on the AUC-ROC. This work contributes to the academic understanding of crowdfunding dynamics and provides actionable insights for researchers interested in NLP in crowdfunding.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning	First occurrence
AdaBoost	Adaptive Boosting	p.27
AI	Artificial Intelligence	Title page
AUC-ROC	Area Under the Receiver Operating Characteristic Curve	p.7
BART	Bidirectional and Auto-Regressive Transformer	p.44
BERT	Bidirectional Encoder Representations from Transformers	p.28
BO	Bayesian Optimisation	p.41
BOHB	Bayesian Optimisation and Hyperband	Title page
EDA	Exploratory Data Analysis	p.51
EVS	Explained Variance Score	p.43
GA	Genetic Algorithm	p.43
HB	Hyperband	p.43
HPO	Hyperparameter Optimisation	p.41
LLaMA	Large Language Model Meta AI	Title page
LLM	Large Language Model	p.15
LoRA	Low-Rank Adaptation	p.46
MAE	Mean Absolute Error	p.43
MAPE	Mean Absolute Percentage Error	p.43
ML	Machine Learning	p.15
MLP	Multi-Layer Perceptron	p.15
NLG	Natural Language Generation	p.44
NLP	Natural Language Processing	p.18
NLU	Natural Language Understanding	p.44
PSO	Particle Swarm Optimization	p.43
RF	Random Forest	p.7
RLHF	Reinforcement Learning from Human Feedback	p.44
SD	ScienceDirect	p.38
SH	Successive Halving	p.41
SHAP	SHapley Additive exPlanations	p.67
SMAC	Sequential Model-Based Algorithm Configuration	p.43
SR	Systematic Review	p.21
SVM	Support Vector Machine	p.27
TPE	Tree-structured Parzen Estimator	p.42
XGBoost	Extreme Gradient Boosting	p.7

1. Introduction

This thesis explores the optimisation of reward-tier strategies in crowdfunding campaigns, focusing on the data available from the biggest crowdfunding platform in the Technology category. Crowdfunding made it possible for entrepreneurs and creatives to secure funding by offering reward tiers to incentivise backers at different pricing levels and tactics. One commonly known strategy is the 'Early Bird' option, where a limited number of early supporters receive a discount for their early commitment, which creates urgency and momentum. However, the optimisation of reward structures can be challenging cause of complex pricing strategies, exclusivity, and backer behaviour.

This research aims to optimise crowdfunding success by developing a Bayesian Optimization and Hyperband (BOHB) framework to enhance predictive modelling and reward-tier strategies for technology-based campaigns. Crowdfunding success is defined as a binary variable dependent on if the funding goal is obtained. In this thesis, this binary variable is used to measure the success and accuracy score of the found configurations of the variables.

BOHB is a Machine Learning (ML) technique specifically suitable in complex optimisation problems, which is applicable here with the high-dimensional optimisation problem in crowdfunding. It makes use of guided configuration sampling of Bayesian Optimization and Hyperband for efficient resource allocation, which balances the exploration and exploitation effectively (Falkner et al., 2018). Since textual relations play a big part in crowdfunding success, we use LLaMA (Large Language Model Meta AI) embeddings to extract nuanced textual features from campaign descriptions and reward names. LLaMA embeddings offer richer semantic representations, which improves the predictive capability of machine learning models. Large Language Models (LLMs) now play a major role in not only research, but also daily life (Chang et al., 2023) and new models keep being released, offering better performance every time. These LLaMA embeddings are integrated into an optimisation process of two phases. In the first phase, BOHB is employed to optimise the hyperparameters of both Multi-Layer Perceptron (MLP) and Extreme Gradient Boosting (XGBoost) models for the success prediction of campaigns. After evaluation, XGBoost is chosen as the best performing model, demonstrating a higher accuracy than MLP. Building upon this foundation, the second phase introduces reward-tier information, including the number of rewards, average price, early bird and exclusive flags, sentiment analysis, reward price distribution, and clustering of reward types. BOHB is applied here as well, but then to fine-tune the parameters of the XGBoost model, leading to improved predictive performance. In the final model, customisable configurations are introduced on specific campaign parameters such as campaign duration, to simulate various scenarios. This specifically gives more insight in the reward structures and increases the probability of crowdfunding success.

The thesis is supervised by Jörg Osterrieder, Berend Roorda, and Stefana Belbe. They are all specialised in either finance, machine learning, or both. Motivated by personal interest and the academic objectives of the Financial Engineering and Management program, this research aims to bridge the gap between theory and practice. In addition to testing the model, this study will deepen the understanding of how AI can assist in managing crowdfunding campaigns. By doing so, it will provide practical recommendations for entrepreneurs looking to maximise funding potential through strategic campaign design.

1.1 Context

Crowdfunding

"Crowdfunding is the act of collecting monetary contributions together with feedback and suggestions from a crowd of contributors (either in form of donation or in exchange for some forms of reward) through an open call on enabling web platforms" (Butticè et al., 2018).

Crowdfunding has several applications. It is well-known as an instrument to collect donations when for example a natural disaster occurred, with organisation Giro555 as an example (Giro555, n.d.). There are also other non-profit organisations which continuously use crowdfunding to fund their projects. One example is the Ocean Cleanup (The Ocean Cleanup, n.d.). Their aim is to have removed 90% of floating ocean plastic by 2040.

Although crowdfunding can be used in commercial applications, it also serves as a valuable tool for empowering emerging entrepreneurs and creative individuals. It enables these often under-resourced actors to bring innovative ideas to market, thereby promoting diversity and competition in industries typically dominated by large corporations. Crowdfunding platforms such as Kickstarter and Indiegogo have become popular venues for entrepreneurs and creatives to launch projects and seek financial support from the public. Both these platforms make use of reward-based crowdfunding, where the backer of a project receives a reward in return, which is often the product that is offered in exchange for a sum of money. Since its launch in 2009, Kickstarter has successfully funded over 266,000 projects with more than 24 million backers contributing over $8 billion (About – Kickstarter, n.d.). Indiegogo, which was founded in 2008, has realised over 800,000 projects with more than 9 million backers (Learn About Indiegogo | Indiegogo, n.d.). A key difference between these platforms is the funding model: while Kickstarter requires projects to meet their funding goal to receive any money, Indiegogo offers a flexible funding option, allowing project creators to continue even if the goal is not reached.

Figure 1: The Ocean Cleanup, they use an innovative system to clean plastic from oceans (The Ocean Cleanup, n.d.)

Figure 2 (left): Kickstarter Logo (About – Kickstarter, n.d.)
Figure 3 (right): Indiegogo Logo (Learn About Indiegogo | Indiegogo, n.d.)

Crowdfunding success, defined as obtaining the campaign goal, is influenced by several factors. This includes the target funding goal, social proof (such as early backer support), campaign duration, team experience (if they have successfully run past projects), and the multimedia used (Elrashidy et al., 2024).

Artificial Intelligence

Various forms of research have been published, and more often a form of Artificial Intelligence (AI) is used. AI is commonly characterized by the capability of machines to exhibit cognitive functions akin to human intelligence, or ultimately outperform humans with certain specific tasks. Often, Machine Learning (ML) is seen as the same as AI. However, AI is an overarching term that encompasses additional techniques and requirements besides ML. A full AI solution would have automated data identification, testing, and decision-making. ML involves manual data identification, testing by hand, and human decision-making. Because it is difficult to implement a full working system on AI, most focus in research is on ML, to target a specific subprocess (Aziz & Dowling, 2019). This can be very specific, such as using ML to analyse the videos used on Kickstarter (Korzynski et al., 2021).

Reward-tier strategies

One of the most important factors is the reward-tier strategy, which significantly influences campaign success. Reward-tier strategies involve offering different rewards or perks to backers at varying price points, such as the earlier mentioned 'Early Bird' option. However, as projects increase in complexity, optimising reward tiers becomes more challenging. An example of this is the game Exploding Kittens, the most backed project on Kickstarter, which raised $8.7 million. The platform's scalability features helped the creators avoid under-ordering or over-ordering, demonstrating how reward tiers, combined with effective campaign management, can drive success (Kuchera, 2015). An example of a reward tier structure is shown in Figure 4, of the BioLite FirePit, a product which is a smokeless woodburning firepit, used with camping (BioLite FirePit, 2019).
In crowdfunding, there are many different characteristics that contribute to success or failure of a campaign. Few examples are the setting the funding goal, the clarity of the timeline, and the number of rewards. For all those three parameters, there is an optimum value which could vary per category of project, that results in a higher probability of crowdfunding success. More information and studies on these subjects are elucidated in Chapter 2.

Figure 4: Reward Tiers BioLite FirePit: different configurations of products (BioLite FirePit, 2019)

All these parameters are influenced by human behaviour, and this differs in every situation. To illustrate this with an example, we look at the number of reward options. Elitzur et al. (2024) identified the overchoice phenomenon in crowdfunding, demonstrating an inverted U-shaped relationship between the number of reward options and campaign performance. This suggests that while offering diverse reward tiers attracts backers, too many options may overwhelm them, leading to decreased campaign success. However, studies do not seem to agree on the most optimal number of rewards, leading to the highest change of project success, and lightly contradict each other, most probably as behaviour changes over time and differs per demographic group. The Discussion Chapter reflects on this subject.
One other important aspect in crowdfunding is a certain behaviour theory: signalling theory. This theory is essential for understanding how backers interpret the information presented in a campaign (Cumming et al., 2024; Moy et al., 2024; Tajvarpour & Pujari, 2022; Zhai & Shen, 2024). Reward tiers and clear timelines both act as signals that can convey reliability and professionalism, directly influencing backers' decisions. Striking the right balance between clarity and feasibility is key to maximising campaign success.
Signalling theory also highlights the role of language, making Natural Language Processing (NLP) critical for this thesis. As an example, Tajvarpour & Pujari (2022) found that informal language reduces the crowdfunding success, whereas punctuation increases the success. Or that risk rhetoric creates a negative external perception reducing performance, whereas reward rhetoric creates a positive external perception increasing performance. This is another example that shows the value of semantics used in crowdfunding campaigns, illustrating why incorporating a form of text analysis is of high importance. Seemingly small factors can matter—higher levels of communication openness are linked to greater project success (Yasar et al., 2022).
As crowdfunding continues to evolve, the availability of large datasets and the increasing sophistication of AI offer new opportunities for campaign optimisation. With the use of AI and specifically ML, campaign creators can better understand backer preferences, predict campaign outcomes, and as this thesis shows, refine the reward structures.

1.2 Problem Statement

Crowdfunding operates under several models, but this thesis focuses specifically on reward-based crowdfunding, which is the system used by platforms like Kickstarter (Aziz et al., 2023). While these platforms provide vast opportunities for entrepreneurs and creatives to reach potential backers, many campaigns still fail. One of the reasons is a lack of strategy in their reward tiers (Greenberg et al., 2013; Kuppuswamy & Bayus, 2018; M. Zhou et al., 2016). Selecting the wrong reward-tier strategy can confuse or overwhelm potential backers, leading to underfunding or campaign failure. Research points out issues such as the overchoice phenomenon, where too many reward options can negatively impact campaign performance (Elitzur et al., 2024), and the role of signalling in shaping backer perceptions (Cumming et al., 2024; Moy et al., 2024; Tajvarpour & Pujari, 2022; Zhai & Shen, 2024). Yet, despite these insights, most campaign creators rely on intuition or trial-and-error approaches, resulting in inconsistent outcomes.

Despite the availability of over 15 years' worth of data from crowdfunding platforms, which has led to a rich body of literature, there remains a significant opportunity to optimise reward-tier strategies using modern analytical techniques. This not only offers an opportunity to gain insight into the behaviour of backers but also offers a higher change for creative individuals who do not have the financial means themselves, to bring their project alive. AI has gained attention in recent years across finance, being used in areas like trading, portfolio management, risk modelling, and text mining (Y. Li et al., 2023). Although AI has gained a large presence in the crowdfunding domain, applications specifically in optimising reward tiers, where predictive analytics and insights from backer behaviour could be game-changers, still remain relatively underexplored.

The challenge is not that reward-tier optimisation has been totally ignored, but that AI advances are evolving so rapidly that newer, more sophisticated methods are becoming available. Whereas there exist many studies that research parts of the reward tier strategy, e.g. making use of early-bird options (Wessel et al., 2019), there are very few that integrate multiple aspects, since this is difficult to capture with traditional approaches. This thesis uses NLP and Bayesian Optimization and Hyperband (BOHB) which can better account for backer behaviour, language patterns, and dynamic preferences than traditional static or rule-based approaches.
Recent breakthroughs in AI, such as LLaMA, provide unprecedented capabilities in understanding campaign narratives and backer sentiment, but their potential for optimising reward-tier strategies in crowdfunding has not yet been fully realised.
Without a data-driven method in place for reward-tier optimisation, many campaigns fail to reach their funding goals, even when they have innovative underlying ideas or products. This not only undermines the potential of crowdfunding as a powerful financing mechanism but also limits growth opportunities for entrepreneurs. Therefore, there is a strong need for an AI-enhanced solution that can help campaign creators design and optimise reward tiers, balancing diversity with backer preferences and improving overall campaign performance.

Summarised, **the problem statement** of this thesis is *to address the gap found in literature in optimising reward-tier strategies in crowdfunding campaigns by developing a data-driven framework that integrates the latest AI techniques, such as BOHB (Bayesian Optimization and Hyperband) and LLaMA embeddings, to analyse campaign parameters and optimise reward-tier structures.*

The **main research question** is: *How can AI-driven methods, specifically BOHB and NLP, be applied to optimise reward-tier strategies in crowdfunding campaigns, maximising campaign success and improving backer engagement?*

1.3 Research Questions

Following the problem statement, the research delves into the following questions:

1.1) How do different campaign features (e.g., funding goals, reward levels, early backer incentives) interact to influence the financial success and risk profile of a crowdfunding campaign?
The characteristics that influence a crowdfunding campaign are described, the general knowledge about the optimal parameters is discussed. Relevant literature will be reviewed to build on existing knowledge. This review is described in Chapter 2.

1.2) How do different reward-tier structures influence backer behaviour in terms of funding commitment, what role does the use of language have here, and how can AI-driven optimisation enhance this interaction?
This question shows the relevance of this research. It aims to quantify how improving reward-tier structures will increase backer engagement. This is discussed in Chapter 2 as well.

2.1) How can BOHB optimise reward tiers by balancing reward diversity with backer engagement in crowdfunding campaigns, thereby minimising financial risks associated with underfunding?
This question elucidates on the characteristics and mechanisms of BOHB. The necessity of BOHB for optimising reward tiers is demonstrated by comparing it to other common optimization methods. Relevant literature on BOHB is reviewed to demonstrate the novelty and advantages of this ML method in this context, which constitutes the first part of Chapter 3.

2.2) How can LLaMA or other NLP techniques be utilised to identify key textual features in campaign descriptions and updates that enhance the prediction of crowdfunding campaign success?
This question also touches on the current literature and explains how LLaMA and other NLP techniques can be leveraged to analyse textual content to predict campaign success and related risks. This, shown in Chapter 3, shows how NLP improves prediction accuracy in optimising reward tiers.

2.3) How can the integration of BOHB and NLP models contribute to a more accurate prediction of campaign funding outcomes and minimize the financial risks of failing to meet funding goals?
This question explores the synergistic effect of combining BOHB and NLP models, supported by relevant literature, in the context of this thesis. This is included in the last part of Chapter 3.

3) What are the optimal hyperparameters and configurations for BOHB in combination with NLP and other ML models to maximise prediction accuracy, and how do they compare to a baseline model?
After data analysis and preparation in Chapter 4, Chapter 5 and 6 answer this question with the development and testing of the model that is made in this thesis.

4) How can AI-driven optimisation of reward-tier configurations enhance crowdfunding campaign success by balancing campaign outcomes, reducing backer choice complexity, and minimising financial risks?
This question is covered in Chapter 6, where the model is evaluated, and the results are shown.

1.4 Approach

This subsection elucidates the approach, referencing the research questions outlined earlier. The methods for conducting this research are described in four key areas: literature review, data collection, model design, and model testing.

Literature study
In Chapters two and three, a literature study is conducted. A systematic search string is composed using relevant keywords, and a date range will be selected alongside other selection criteria, such as the field of studies. The databases of Scopus (Scopus Search | Elsevier, n.d.) and ScienceDirect (ScienceDirect.Com | Science, Health and Medical Journals, Full Text Articles and Books., n.d.) are used. This study will focus on answering the research questions outlined in clusters 1 and 2.
To discover the existing research in the specific topic of this thesis developing the combination of reward tier optimisation using BOHB and NLP, a Systematic Review (SR) is conducted (Kankanamge et al., 2019). While SR is often used for scarce and disjointed literature, it is also applicable here to synthesise the fragmented knowledge in this area. SR is a methodological approach to synthesise and summarise the state of knowledge on a given topic or research question. It follows strict guidelines in three stages: objectivity, transparency, traceability, and replicability (Cradock-Henry et al., 2019). This is executed in three stages (Butler et al., 2021), for the flow in Chapter 3, where the execution is described, the approach is described here:

- Stage 1: Planning. All search strings are defined as keywords and are used to search the title, abstract, and keywords of relevant publications. The databases of Scopus and ScienceDirect are utilised. A date range is selected based on the novelty of the subject (if the topic is novel, no date range will be applied). Based on the number of papers found, a further selection is made.
- Stage 2: Review. The articles are reviewed based on primary inclusion and exclusion criteria, focusing on title, abstract, and keywords. Duplicates are removed, and only accessible articles are included. The full text of the selected articles will be reviewed to determine their relevance to the research subject.
- Stage 3: Results. The papers are analysed using qualitative techniques. One example of a technique that can be used is called Pattern Matching. This is mainly used for case studies and qualitative research in general (Attard Cortis & Muir, 2022), which is the desired type of literature, as this thesis also uses qualitative data. The found papers can then be categorised and cross-checked against other literature. In this thesis, there is no need for Pattern Matching, since the result is very limited in size.

Data collection
The model requires training with an adequate volume of data. Crowdfunding data is collected using a readily available database of Kickstarter campaign information provided by WebRobots.io (Web Robots IO, 2025) in combination with data collected by web scraping. We choose to use the subset of Technology campaigns created on Kickstarter. After the data collection, the relevant information is cleaned, prepared and analysed. A critical aspect of this process is determining which data is most relevant for training the model. For example, can data from certain periods (e.g., during Covid-19) be considered less relevant due to unusual market conditions? (Elrashidy et al., 2024) show that Covid influenced the reaction to risk disclosure and using URL's. Addressing potential biases, time dependencies, and reflecting on the data's quality and representativeness is crucial. Chapter 4 gives an extended overview of the data and describes the distribution, reflects on the literature, and how it influences the model. We now provide a list of all features that are used for the models:

- Reward tier features: Reward count, Average reward price, Early bird presence, Sentiment of reward titles, Top reward name category, Exclusivity
- Campaign metadata: Goal in dollar, Campaign duration, Subcategory, Country
- Textual Content: Blurb and slug, Reward titles, Sentiment label, Embeddings (LLaMA or sentence-transformers/all-MiniLM-L6-v2)

Designing the model
This research leverages a BOHB framework alongside advanced Natural Language Processing (NLP) techniques to optimise reward-tier strategies in crowdfunding campaigns. The end goal of the model is to use different parameters for campaigns, such as a set campaign duration, to find the best reward tier configurations, such as the average price. It is an iterative process, building on a basic model to extend the model with multiple layers, to ultimately arrive at the final working model. There are four different steps in designing the final model:

- The first version of the model is a basic Random Forest (RF) model that can predict campaign success based on different campaign parameters such as the campaign duration and campaign goal. This model did not make use of LLaMA embeddings, but sentence-transformers/all-MiniLM-L6-v2 embeddings, to compare later on.
- The second version of the model did incorporate the LLaMA embeddings and compared the performance of three different ML models: RF, XGBoost, and MLP. To find the optimal settings of these models, BOHB is leveraged here as well.
- In the third version of the model, XGBoost is used together with BOHB to determine the optimal parameter settings for reward-based attributes, such as the number of rewards, or the use of exclusivity, without giving any specific campaign information such as the country.
- The final model now continues on the structure of the third model, but now accepts different customisable configurations for campaign parameters, so it has a higher relevance to practical applications.

The model architecture is informed by insights from the literature review and employs a range of publicly available Python libraries. Key components include HpBandSter (Falkner et al., 2018b) and Optuna (Akiba et al., 2019) for the implementation of BOHB, Hugging Face Transformers (Wolf et al., 2020) for NLP tasks, Pandas (Mckinney, 2010) for data manipulation, and PyTorch (Paszke et al., 2019) for model training and embedding extraction. Central to this approach is the integration of LLaMA-derived embeddings for capturing textual nuances in campaign descriptions and updates. In addition, a separate sentiment analysis model is employed to extract backer sentiment, enhancing the understanding of narrative-driven influences on campaign success. The design prioritises not only prediction accuracy but also computational efficiency, optimising the resource-intensive processes associated with NLP tasks. Further elaboration on the model's design and development process is presented in Chapter 5, where experimental setups, hyperparameter optimization, and evaluation metrics are discussed in detail.

Model testing and validation
In Chapter 6, the model's reliability, performance, and accuracy are evaluated using metrics specific to classification tasks, reflecting the goal of predicting campaign success. Key metrics include accuracy, precision, recall, and F1-score. These metrics are particularly valuable for assessing model performance in the presence of class imbalance, as they offer a more nuanced view than accuracy alone. Additionally, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) assesses the model's ability to distinguish between successful and unsuccessful campaigns across various thresholds. Moreover, a Sensitivity Analysis is performed to investigate the sensitivity of important features for the performance of the model.
This chapter will also provide a reflection on the used models and how it compares to the literature.

1.5 Concluding words Chapter 1

In summary, this chapter has introduced the key challenges and opportunities in optimising reward-tier strategies for crowdfunding campaigns, emphasizing the importance of AI-driven methods like BOHB and NLP for improving campaign success. The research questions are outlined to explore how these methods can enhance backer engagement, minimize financial risks, and refine campaign strategies. The approach is introduced, and the four different design steps of the model are shown. The subsequent chapters will delve deeper into the literature, data collection, model development, and testing methodologies, setting the foundation for the practical application of these AI techniques in optimising crowdfunding outcomes.

2. Literature Review - Crowdfunding Campaign Features

This chapter summarises the current overview of knowledge in the field, highlighting a gap in the literature and demonstrating the extensive possibilities and applications of Machine Learning in crowdfunding. This chapter gives a list of the ML methods used in the most relevant papers, and no definitions are required to understand the State of the Art. For a more thorough explanation, see Appendix A. Understanding the choice of methods is important; however, the specific details of methods relevant to this research are explored in greater depth in Chapter 3. To achieve this, the chapter addresses the following research questions:

- *How do different campaign features (e.g., funding goals, reward levels, early backer incentives) interact to influence the financial success and risk profile of a crowdfunding campaign?*

This question has a high relevance to this thesis, as during the development of the model, the parameters in the data are chosen. If certain parameters have a negligibly small influence they could be omitted to limit the complexity of the calculations of the model. These different campaign features are reflected upon during data preparation. In the data Chapter, number four, the literature is compared with the findings in the dataset that we use in this thesis.

- *How do different reward-tier structures influence backer behaviour in terms of funding commitment, what role does the use of language have here, and how can AI-driven optimisation enhance this interaction?*

Building on the first question, this explores how campaign features interact with reward-tier structures. These relationships are explored in the subchapter of 'Reward-tier Structures'. This part zooms in on the backer behaviour related to reward tiers. Furthermore, it demonstrates how the data related to crowdfunding campaigns is intricate enough to present an optimisation problem ideally suited for AI.

2.1 State of the Art

This subchapter serves to show the findings of the initial review of the literature. A review of relevant papers reveals that there is a strong focus on the prediction of the success of the campaigns in crowdfunding. This is one of the most evident measures there is, since there is always data available about the status of the campaign (failed or succeeded), which is quite objective and unambiguous. Many studies make use of the same data, which is often retrieved from Kickstarter, since it has one of the most complete and vast databases there is in crowdfunding. It is quite interesting to see that even though the researchers work with the same data (mostly varying in the period of time), that the research is still extensive and many different research questions and methodologies are applied. The volume of research on crowdfunding continues to grow, and the focus areas are evolving. There used to be a focus on social capital, peer-to-peer lending, decision-making, venture capital, and open innovation. However, nowadays, other topics such as COVID-19, fintech, and machine learning receive more attention(Aziz et al., 2023). It can form a challenge to find a new combination between the research question(s) and methodology, therefore, a comprehensive dive into the literature is desired. Some research goals are categorised, and the methods used are described. While not exhaustive, this review illustrates the breadth of research available and which type of categories are of relevance for this thesis.

Team Characteristics and their Impact

A significant portion of literature delves into the team behind the campaign. They discuss the personality, and prior crowdfunding performance of the entrepreneur and arrive to interesting conclusions that for example having a female entrepreneur leads to a higher success rate, and narcissistic business owners launch less successful campaigns (Aziz et al., 2023). The personality is often analysed, and one common finding is the positive relationship between openness and crowdfunding success (Neuhaus et al., 2021). Furthermore, another category looks from a societal perspective, such as the challenges women face in using crowdfunding for financing (Saluja, 2024). While these findings are insightful, they may have limited practical applicability, as the composition of the team is often predetermined.

Investor and Backer Motivations

Other studies shift perspective to examine what motivates backers to invest. For example, they are more likely to repeatedly pick similar reward tiers rather than adhering to specific times for their investments. The longer they have been on the platform, the more this "sticking" behaviour shows up in picking rewards, although it appears less influential in determining the timing of their investments (Xiao & Yue, 2018). Additionally, another study by Hoegen et al. (2018) identifies six main categories of factors that influence backer decisions: financials and campaign characteristics, project and product quality, founder perception and attributes, social interactions, community and third-party involvement, and contextual elements. Among these, the final category, investor characteristics, was found to influence the impact of the other categories but did not directly contribute to the decision-making process itself. While investor characteristics may hold relevance for this thesis, it is expected these nuances are captured with ML algorithms. If unexpected results occur, this section of the literature may be revisited to explore these dynamics further.

Temporal and Event-Based Analysis

As mentioned, one of the key topics the last few years was how Covid impacted the crowdfunding campaigns. One finding is that there were more expressed sad emotions in a campaign's description after the COVID-19 outbreak (J. Wang et al., 2022). Another study based in the UK finds that competition and Covid had a positive effect on crowdfunding success, while the Brexit had a negative influence (Vu & Christian, 2024). Some studies look at the development in crowdfunding throughout the years, such as the legislation that has come into place (Y. Zhao et al., 2019). It is important to keep in mind that the occurrence of these events or change in regulations could give different results in studies executed some years ago, in respect to the current ones. One measure that could possibly be applied is to put a higher weight on more recent campaigns than on older campaigns.

The Effects of Crowdfunding

Some papers discuss the relevance of crowdfunding for either the company or society that crowdfunding can give, as well as the challenges it introduces (Hussain et al., 2023). Crowdfunding through online platforms has broadened access to capital, empowering entrepreneurs to create products and services that cater to diverse public needs. In addition, it can help companies to validate their idea and build a community of supporters (Hoque, 2024). Another finding is that crowdfunding could expand the international reach of the company, and strengthens its commercial networks (Troise et al., 2023). Crowdfunding is particularly valuable for firms facing high uncertainty about consumer preferences, such as developers of innovative consumer products. Furthermore, the structure of reward-based crowdfunding, particularly the constraints imposed by platforms, aligns firms' incentives with the interests of their backers, reducing the likelihood of moral hazard (Chemla & Tinn, 2020).

However, crowdfunding also comes with risks: the public disclosure of a large amount of information can expose entrepreneurs to intellectual property theft, fraud, and misuse. Entrepreneurs may also face operational challenges, such as the strain caused by overfunding. Additionally, crowdfunding disrupts traditional financial institutions, raising questions about its regulatory and legal implications. At the same time, it represents a democratisation of finance, addressing inequities in the current system by providing new funding opportunities for women and minority entrepreneurs (Hoque, 2024). These findings provide important context for this thesis. While the risks and benefits of crowdfunding do not directly influence the optimisation of reward-tier structures, they underline the broader relevance of this research. For example, a well-optimised reward-tier structure could help mitigate risks like financial instability and operational strain, while also making campaigns more appealing to diverse groups of backers. By addressing these inefficiencies, this research contributes to making crowdfunding a more sustainable and equitable financing option, aligning with the societal benefits discussed in the literature.

Prediction of Project Success

The prediction of crowdfunding project success is one of the most extensively researched topics in the field. This focus is particularly relevant to this thesis, as the optimisation of reward tiers is evaluated based on campaign success. Recent studies frequently employ ML methods, making it important to analyse commonly used approaches and their suitability for specific research questions. For more detailed information on ML methods, refer to Appendix A.

Common Characteristics in Success Prediction Studies

Most success prediction studies share several key characteristics. They often incorporate features such as funding goals, campaign duration, categories, images, videos, and reward levels. Social signals, including Facebook connections, Twitter metrics, and the number of backers, are also frequently examined. Early studies primarily predicted whether a campaign reached its funding goal (binary success). Over time, research expanded to predict more nuanced outcomes, such as time-to-success, pledged funding range, and the factors influencing novice versus experienced creators (Ahmad et al., 2017; Y. Li et al., 2016; Tran et al., 2016; Ryoba, 2021; M. Zhou et al., 2016). In the most recent years, visual and textual elements are also incorporated into the research (Blanchard et al., 2022; Gündüz, 2024; Z. Tang et al., 2023).

Early Classification Techniques

In earlier studies, researchers relied on static ML methods, often constrained by the limited data available at the time. For example:

- Support Vector Machine (SVM): (Etter et al., 2013) achieved an accuracy of over 76% by focusing on Twitter data.
- Combination Methods: Greenberg et al. (2013) achieved 68% accuracy using decision tree algorithms, SVMs, and AdaBoost. Similarly, M. Zhou et al. (2016) reported 73% accuracy, claiming it exceeded educated guessing by 14%.
- Logistic Regression: M. J. Zhou et al. (2015) obtained 73% accuracy with logistic regression and recommended SVM as a classification algorithm, a choice supported by Oduro et al. (2022).

However, these early methods often yielded inconsistent results. For instance, Cordova et al. (2015) achieved only 67% accuracy when applying logistic regression to technology projects. Later studies began incorporating more advanced algorithms, which improved prediction accuracy.

Advancements in Machine Learning Algorithms
As ML techniques progressed, accuracy rates increased significantly. Notable advancements include:

- Random Forest and Naive Bayes: Tran et al. (2016) used these models with Twitter features, achieving 81% accuracy for Random Forest and 78–80% for other models.
- Survival Analysis: Yan Li et al., (2016) combined survival analysis with logistic regression and SVMs, reporting accuracies of 76% and 78%, respectively.
- Random Forest with AdaBoost: Ahmad et al. (2017) achieved over 94% accuracy, marking a major improvement over earlier methods.

Emerging Techniques and Deep Learning
Recent years have seen the emergence of deep learning and other innovative approaches:

- Deep Neural Networks: Lee et al. (2018) employed sequence-to-sequence models with sentence-level attention and Hierarchical Attention-based Networks, achieving 89–91% accuracy, including 76% accuracy using only text data.
- Multi-Layer Perceptron (MLP): Yu et al. (2018) showed this model outperformed Random Forest with AdaBoost, reaching 93% accuracy.
- Deep Learning vs. Traditional Methods: W. Wang et al. (2020) demonstrated that deep learning achieved the best results (92.3% accuracy) compared to decision trees, random forests, logistic regression, SVM, and K-nearest neighbours.
- Gradient-Boosted Decision Trees: Y. Guo et al. (2021) found that they provide 70.2% accuracy based on launch-time information alone.
- Metaheuristic Approaches: Ryoba et al.(2021) applied a whale optimization algorithm with K-nearest neighbour, achieving 90.28% accuracy using a subset of nine features.

Visual and Language Models
Visual and textual elements are increasingly explored in crowdfunding prediction research:

- Visual Data: Blanchard et al. (2022) applied flexible ML models, including Lasso, Ridge, Bayesian additive regression trees, and XGBoost, to analyse visual content.
- Language Models: Language models such as bidirectional encoder representations from transformers (BERT) now dominate the field. For instance, Gündüz, (2024) demonstrated that BERT significantly outperformed FastText. Z. Tang et al. (2023) introduced a Deep Cross-Attention Network that outperformed BERT by integrating video and text data.

More detailed information about the language models is provided in Chapter 3.

Key Observation
One consistent finding across studies is the robust performance of Random Forest models. For instance, Zhong, (2022) highlights Random Forest's ability to overcome issues like low stability and overfitting, often outperforming other classifiers. Similarly, Haitham et al. (2024) reaffirm the model's strong predictive power, making it a useful benchmark for this thesis.

2.2 Influence of campaign features

From the literature, it's evident that specific campaign features, such as funding goals, reward levels, and backer incentives, play crucial roles in influencing a campaign's success and its associated financial risks. In this subchapter, all factors are stated and explained how they interact with different outcomes. It is important to note that the factors are identified as non-linear and can interact dynamically for each project (Oduro et al., 2022). This outline also shows the diversity and divergency of the factors. They are grouped in eight critical key themes to achieve the desired objectives.

1. Clear and obtainable Funding Goals
Many papers point out that the funding goal amount is a clear factor of the success of a project (Ahmad et al., 2017; Kuppuswamy & Bayus, 2018; Mollick, 2014; Ryoba et al., 2021; M. Zhou et al., 2016). Elitzur et al. (2023) Found that in their model, the log of the funding goal was the most influential metric, with text variables included second. Clear financial goals are fundamental to building trust and reducing backer uncertainty. Lukkarinen et al. (2016) and Pinkow & Emmerich (2021) demonstrate that achievable funding targets foster confidence among backers, directly improving campaign outcomes. Frydrych et al. (2014) found that lower funding targets signal legitimacy by creating modest, achievable expectations. This is in line with the effect of extra contributions just before reaching the target goal due to the psychological effect of goal proximity., and when it is reached, the contributions tend to decrease (Kuppuswamy & Bayus, 2017). Butticè et al. (2018) shows that higher targets are negatively associated with success. Conversely, excessively ambitious goals increase the risk of unmet expectations in the case of medical campaigns, as noted by Aleksina et al. (2019). The balance between confidence and financial potential is crucial for setting realistic goals. Y. Wang et al. (2021) emphasise that low funding goals increase perceived achievability, while Coakley et al. (2021) argue that overly conservative targets may limit financial success. Transparent milestones, as highlighted by Weber et al. (2023), further enhance backer trust by clearly delineating how funds will be utilised.

2. Overfunding Risks and Clear timelines
If the project duration is too long, it negatively impacts the success of the project (Fernandez-Blanco et al., 2020). Clear timelines improve backer confidence and enhance financial success. Y. Wang et al. (2021) found an inverted U-shaped relationship between timeline clarity and crowdfunding success, in the technology category. Frydrych et al. (2014) noted that shorter campaign durations reflect realistic and focused planning, signalling legitimacy to backers. Moreover, clear timelines mitigate the negative effects of overfunding by ensuring efficient resource allocation (Y. Wang et al., 2021). While overfunding provides resource advantages, it introduces risks such as project delays and unmet backer expectations. Blockbusters increase platform visibility but may monopolize backer resources (Z. Wang et al., 2022). Weber et al. (2023) and Sendra-Pons et al. (2024) emphasise that signalling realistic goals and maintaining transparent communication are essential to managing these risks.

3. Frequent Updates and Communication style
Following up on transparent communication, this alongside frequent updates significantly influence backer trust and engagement. Hui et al. (2014) and N. Wang et al. (2018) demonstrate that regular updates keep backers informed and engaged, reducing risks associated with uncertainty. The use of social media generally aids the project (Jankü et al., 2023). Maintaining engagement with backers over time helps sustain momentum and improves outcomes (Sendra-Pons et al., 2024; Q. Zhang et al., 2017). Another recent study found a significant relationship between quality/trust signals and campaigns' funding success (Elrashidy et al., 2024). Transparent communication, particularly regarding project milestones and progress, fosters credibility (Aleksina et al., 2019; Q. Li & Wang, 2024; M. Zhou et al., 2016). Beyond transparency, narrative legitimacy in rewards-based crowdfunding often stems more from the online community within the platform than from visual pitches alone (Frydrych et al., 2014). Besides having a clear communication style, the style of voice of the description of the project has importance as well.

Clarity and emotional appeal in project descriptions are critical factors driving backer contributions (K.-F. Yang et al., 2023; X. Zhang et al., 2021) as well as positive sentiment (Adamska-Mieruszewska et al., 2021; Jankü et al., 2023; W. Wang et al., 2017). Positive sentiment in descriptions increases campaign success rates, especially in tech-focused campaigns (Yosipof et al., 2024). Furthermore, a product description should contain buzzwords and have a semantic richness – have many words which trigger a larger thought pattern (Babayoff & Shehory, 2022). Furthermore, simple, understandable language must be used (Adamska-Mieruszewska et al., 2021). Even the verbs needed to request the reader to invest in the project matter (Carradını & Nystrom, 2024). Linguistic style, defined by features like concreteness, preciseness, and interactivity, has been shown to significantly impact success rates, particularly for social campaigns (Parhankangas & Renko, 2017). Most successful descriptions make use of general persuasion principles, such as reciprocity (Mitra & Gilbert, 2014). Frydrych et al. (2014) highlight that aligning the narrative to the platform community's expectations enhances backer trust, reinforcing the campaign's legitimacy. Rose et al. (2021) add that campaigns featuring products in early development stages or with long delivery times often feel psychologically distant to backers, reducing contributions and success rates. To counter this, outcome-focused mental simulation—helping backers imagine the benefits of the product—can effectively increase engagement and funding commitment.

4. Visual Elements
Combining textual and visual data enhances success prediction, as it complements textual content to signal professionalism and reduce perceived risks (Blanchard et al., 2022). They can enhance engagement as well (Barnes, 2024; Sendra-Pons et al., 2024; M. Zhou et al., 2016), because most likely the positive first impressions from the images significantly influence backer decisions (Q. Guo et al., 2022). Moreover, visual appeal amplifies network effects – where people share and interact more with the campaign, driving engagement across related campaigns (Y. Wang et al., 2022). Furthermore, the number of videos seems to enlarge the campaign success (Ahmad et al., 2017; Kuppuswamy & Bayus, 2018; Mollick, 2014; Ryoba et al., 2021; M. Zhou et al., 2016). While Frydrych et al. (2014) emphasise narrative over visuals in signalling legitimacy, visuals still serve as a vital supplement for enhancing campaign professionalism. Elrashidy et al. (2024) found the use of photos to have an adverse effect on the success of the campaign, however, mixing multimedia has a positive effect.

5. Social Proof and Endorsements
The network effect is quite important. Network effects amplify campaign visibility and success, especially in technology niches (Y. Wang et al., 2022). The number of comments is an indicator of how successful the campaign will be (Fernandez-Blanco et al., 2020). Strong online communities amplify the effectiveness of campaign features (Marinova, 2019). Plus, social proof and endorsements amplify campaign credibility, leading to increased backer trust and contributions. Active engagement on social media is highly correlated to crowdfunding success (Q. Zhang et al., 2017). Support from the platform for the project and frequent updates both are positively correlated with the success of the project (Martínez-Cháfer et al., 2021). Q. Li & Wang (2024a) and Jankü et al. (2023) highlight the role of third-party endorsements and strong social media engagement in driving backer participation. Social interaction drives engagement, especially in innovative product campaigns (Pati & Garud, 2021). Psychological determinants like social proof and relatability enhance backer motivation (Popescul et al., 2020). It is important to address the public adequately, cultural alignment of strategies improves performance, emphasizing the need for context-specific approaches (Haasbroek & Ungerer, 2020). Cultural and social alignment of narratives boosts campaign performance (Rama et al., 2022). Furthermore, aligning the goals to the demographics of the audience is important (Belleflamme et al., 2014). Even the identity of the entrepreneur seeking funding matters (Sendra-Pons et al., 2024). Important is that the online presence is not disclosed (Elrashidy et al., 2024). Solo-founders have a lower chance succeeding (McCarthy et al., 2023).

6. Early Backer Incentives and Engagement
Early backer incentives create momentum critical for campaign success. Research by Etter et al. (2013) and Kindler et al. (2019) shows that campaigns with strong initial engagement are more likely to succeed. Similarly, Hui et al. (2014) show that visible support from early adopters creates a bandwagon effect, motivating further contributions. C. Zhao et al. (2019) argue that these early contributions set a positive tone, attracting subsequent backers. This phenomenon underscores the importance of well-timed rewards or exclusive offers to secure early support.

7. Category of Project
The category of the project is important in two ways: different categories have different characteristics that perform well, and a certain category membership can offer an advantage because of a 'blockbuster'. This is a project that performs extremely well. Incremental innovations are more likely to succeed than radical innovations. Non-profit projects are generally more successful than for-profit projects (Butticè et al., 2018). Some studies show examples of how different attributes work differently for other categories. For movie crowdfunding campaigns, storytelling and milestone-based rewards significantly improve backer engagement (M.-Y. Chen et al., 2022). Plus, positive sentiment in descriptions is particularly effective in tech-related campaigns (Yosipof et al., 2024). Lastly, clear academic value and detailed explanations attract diverse backers in academic funding (Sauermann et al., 2019). In general, tailoring campaigns to niche markets ensures higher success rates and stronger backer connections (Corsini & Frey, 2023). One other interesting thing to note is that environmentally oriented campaigns attract diverse backers and higher success rates (Hörisch & Tenner, 2020). They also resonate with environmentally conscious backers (Bento et al., 2019). Environmentally focused campaigns attract global backers, especially in developed markets (X. Tang et al., 2024). The category membership also can cause similar projects to fail or succeed. Blockbusters increase visibility but may monopolize resources, impacting smaller campaigns (Z. Wang et al., 2022).

8. Tailored and Personalized Reward Structures
Reward levels are included in predicting the outcome of the campaign and influence the success (Greenberg et al., 2013; Kuppuswamy & Bayus, 2018; M. Zhou et al., 2016). In the study of Elitzur et al. (2023) the reward options were the fourth most important criterium, taking into account text variables. Greater variety and personalization increase campaign attractiveness, and socially interactive rewards (e.g., merchandise) correlate with higher success rates (Butticè et al., 2018). The supply of multiple reward options is important (Kunz et al., 2017). One study showed the most important feature in their outcomes was the number of reward options (Zhong, 2022a). One example that shows the significance is the overchoice phenomenon exists in crowdfunding as well, with the tipping point of 33 options (Elitzur et al., 2024), which seems contradictory to the findings of (Martínez-Cháfer et al., 2021), who found the number of reward options are independent of the project success. Reward levels tailored to specific backer demographics enhance participation and financial success. Mitra & Gilbert, (2014) and Adamska-Mieruszewska et al. (2021) argue that simplicity and clarity in reward descriptions increase accessibility for a broad audience. Persuasive elements in text interact with reward tiers to influence outcomes (Allison et al., 2017). Babayoff & Shehory (2022) further highlight that semantic richness and emotional resonance in reward descriptions improve backer engagement, especially in niche markets. Furthermore, personalised and tiered rewards significantly enhance backer engagement. Fernandez-Blanco et al. (2020) and W. Wang et al. (2024) demonstrate that tailored rewards, such as milestone-based incentives, align with backer preferences and drive participation. For example, niche campaigns in the video game and sustainability sectors benefit greatly from reward customisation (Bento et al., 2019; M. Y. Chen et al., 2021). However, in sustainable projects, the complexity of managing personalised rewards must be carefully addressed to avoid inefficiencies (Corsini & Frey, 2023). Reward tiers and funding structures significantly affect product development timelines and risks (Candogan et al., 2024).

Importance of Research

The interaction of these eight campaign design factors—clear financial goals, achievable funding targets, personalised rewards, early backer incentives, frequent updates, overfunding management, social proof, and tailored reward levels—plays a critical role in crowdfunding success. Language use has also been shown to significantly influence outcomes: in a large-scale study of 45,000 Kickstarter projects, linguistic features alone accounted for approximately 58% of the variance in project success (Mitra & Gilbert, 2014). This figure does not reflect perfect prediction but highlights the unexpected importance of narrative and persuasive phrasing, since in this study only language (and not other features) was analysed. Given that language impact varies by category, this thesis focuses on technology campaigns, where prior research indicates that positive sentiment and phrasing are particularly predictive (Yosipof et al., 2024). A multimodal model can aid the accuracy, as an example: combining text, visuals, and social features significantly improves success prediction (Cheng et al., 2019). Furthermore, as all elements interact differently with each other (Oduro et al., 2022), ML algorithms seem in place, since they can account for many different scenarios. Lastly, the reward structure is a vital element in the crowdfunding campaign (Greenberg et al., 2013; Kuppuswamy & Bayus, 2018; M. Zhou et al., 2016), and therefore, forms a good topic for additional research.

2.3 Reward-tier Structures

There are 4 major crowdfunding models: equity-based, donation-based, loan-based, and reward-based. The lending and the reward-based models raised most funds across the globe (Wangchuk, 2021). Since most public information is available, the reward-based model is researched in this thesis. This model carries a medium amount of risk for both the founders and the pledgers, as the highest risk is that the founder cannot provide the pledger with the reward, and the money is refunded (Hossain & Oparaocha, 2017). Rewards consist of either material compensation or social acknowledgment (Kraus et al., 2016). The amount of risk and the reward tiers are linked to each other (Candogan et al., 2024). As mentioned in the previous paragraph, the reward options heavily influence the success of the campaign (Greenberg et al., 2013; Kuppuswamy & Bayus, 2018; M. Zhou et al., 2016). Understanding these interactions is crucial to optimising campaign outcomes.

Number of Rewards
As mentioned, this is a vital element (Zhong, 2022), although not all studies seem to agree (Martínez-Cháfer et al., 2021). There are several psychological concepts which play a role in this matter. Y. Lin et al. (2016) discuss the choice overload hypothesis, where the potential backer is dazed by the number of options, also referred to as the overchoice phenomenon (Elitzur et al., 2024). Y. Lin et al. (2016) find that the overchoice phenomenon does not exist in crowdfunding for projects with a maximum of 30 reward options. Elitzur et al. (2024) find a very high tipping point of 33, showing that in general, more reward options increase success probability. However, another study by Cai et al. (2021) finds that there is an inverted U-shaped relationship between the campaign success and the number of rewards. They find the most optimal reward number is 10.

Strategic Pricing
The pricing of the rewards interacts with the number of options. One clear example is the dual-process theory, which suggests that people either think with their System-1 processing (fast, automatic and intuitive), or System-2 processing (slow, deliberate and analytical). Crowdfunding backers tend to use System 1, which is associated with several cognitive biases. One of those is the middle-option bias, where backers may choose the middle reward option more than they choose other options (Simons et al., 2017). Other effective tactics are pricing awards with a number ending with a 4 or a 9, and bundling, but for crowdfunding only with a maximum of 4 options (Keisar & Lev, 2023). Another tactic which is frequently studied is applying a lottery. More backers are attracted when using this strategy, but the backers that would pledge a big amount of money hold back when this strategy is implemented, which causes a lower funding amount (Gong et al., 2021). Moreover, a donation option could be added to the reward tiers. This is mainly for prosocial-cause projects beneficial to reach campaign success (J. Chan et al., 2023).

Language and Rewards
Mitra & Gilbert (2014) and Adamska-Mieruszewska et al. (2021) argue that simplicity and clarity in reward descriptions increase accessibility for a broad audience. Furthermore, semantic richness, emotional resonance (Babayoff & Shehory, 2022), and persuasive elements interact with the reward tiers and can improve backer engagement (Allison et al., 2017). Not only the content, but even the quantity of the words used matters. Bi et al. (2017) found that the word count of the introduction in reward-based crowdfunding typically signals the project quality. More readers will invest with a higher word count.
Another paper studies the use of the Machiavellian rhetoric and measure the relationship between the frequency of Machiavellian rhetoric use and crowdfunding performance. This type of language is divided into eight facets: revenge, intimidation, betrayal, manipulation, ingratiation, supplication, self-disclosure, and persuasion. The use of revenge, self-disclosure, and intimidation have negative effects on the campaign. However, signals of ingratiation and persuasion have mixed positive effects, where ingratiation increases the number of backers, but not the funding success. For persuasion it is the other way around, and betrayal rhetoric is positively related to both measures (Calic et al., 2021). Furthermore, it is determined that an effective entrepreneurial narrative in a reward-based crowdfunding campaign consists of the following elements: 1) problem/need; 2) project; 3) product; 4) team; and 5) venture (Crescenzo et al., 2022).

Timing
Lin et al. (2016) also find that projects with late-added rewards receive relatively more funds, independent on if they are successful or unsuccessful. For successful projects, they receive 1.4 times more of raised-goal rations than those without. Another tactic that enhances the campaign performance is to signal information regarding the future retail price (Sewaid et al., 2021).

Personalization and Exclusivity
There are several ways to implement personalisation as a strategy in reward options. Applicable options to enhance the campaign success are to use socially interactive rewards such as merchandise (Butticè et al., 2018), tailored rewards such as milestone-based incentives (Fernandez-Blanco et al., 2020; W. Wang et al., 2024), or the use of exclusivity for early backers with early birds (Wessel et al., 2019). Moreover, rewards that offer an intrinsic motivation, which regard a product that is sold as individual ownership, such as a prototype or limited edition, allow the project to achieve bigger funding (Cappa et al., 2021; Maiolini et al., 2023).
Y. Lin et al. (2016) show that in both successful and unsuccessful projects, projects with limited rewards have a higher raised-goal ratio. They devote this to the fact that scarcity appeals, and that the limited-edition products mostly result in irrational purchases and faster sales. However, with too many exclusive options, there is an adverse effect (Keisar & Lev, 2023). Dynamic updates to limited-edition options can mitigate this, renewing backer interest and driving contributions (L. Yang et al., 2020).
With early-bird options, there is the so-called "phantom effect", where backers choose the equivalent but undiscounted reward option more frequent. In traditional offline retail, a sold-out hurts, but here, it is effective to promote sales. This effect is especially strong with a moderate amount of discount for early bird options. Plus, the effect is reliant on social proof; if many others choose the early bird option, and there is a huge discount, people are less likely to buy the undiscounted option (M. Chen et al., 2021; Wessel et al., 2019). Lastly, reward levels that let pledgers participate in and experience the project are correlated with project success (Regner & Crosetto, 2021).

AI implementation
In the past, traditional methods were used to predict the success of crowdfunding campaigns, such as linear regression or logit. These are largely dependent on parameterization and assumptions (Cavalcanti et al., 2024). There are several more recent studies, as shown in 2.1: State of the Art, which use Machine Learning and arrive to more accurate predictions for the success of the campaigns (Cavalcanti et al., 2024; Elitzur et al., 2023). In the very recent years, many ML methods have been tested, but it seems BOHB in combination with LLaMA embeddings has not yet been studied, although research seems to point out Bayesian semi-parametric approaches would be the next type of model to test (Oduro et al., 2022). The next chapter will delve into how the application of this type of approach can aid the prediction of success of reward-based crowdfunding campaigns.

2.4 Concluding words Chapter 2

This literature review highlights there is an extensive body of research about crowdfunding campaigns, enabled by the availability of large public databases like Kickstarter. The role of campaign features and advanced ML methods in predicting and enhancing campaign success is frequently emphasized (Blanchard et al., 2022; Elitzur et al., 2023). Key findings indicate that reward-tier structures, clear financial goals, and personalised strategies are critical factors influencing backer behaviour and funding outcomes (Greenberg et al., 2013; Kuppuswamy & Bayus, 2018). Additionally, the integration of textual, visual, and social features has been shown to significantly enhance prediction accuracy, particularly in technology-related projects where language plays a central role in engaging backers (Babayoff & Shehory, 2022; Mitra & Gilbert, 2014). This illustrates the reward tier structures can be complex, and of high importance for the success of a campaign. Using ML to control some of those parameters will therefore benefit campaign creators.

Furthermore, the literature underscores the robust performance of Random Forest as a benchmark ML model, particularly when analysing complex and non-linear relationships in crowdfunding data (Haitham et al., 2024; Zhong, 2022). However, gaps remain in applying advanced Bayesian semi-parametric approaches. The novel approach to utilising BOHB combined with LLaMA embeddings holds promise for optimising reward-tier strategies, particularly through its ability to analyse the semantic richness, emotional resonance, and persuasive elements in language that significantly influence campaign success (Allison et al., 2017; Calic et al., 2021). These insights set the stage for the development of AI-enhanced methodologies to improve campaign outcomes and reduce financial risks in reward-based crowdfunding.

3. Literature Review - AI-Driven Optimisation and NLP in Crowdfunding

This chapter continues the literature review initiated in the previous chapter. It starts off with a Systematic Review (SR) to capture the present applications of the combination of optimisation using BOHB and NLP in crowdfunding, to further highlight the gap present in literature and to be able to build further on existing literature. After this, 3 research questions are answered necessary to grasp the specific background knowledge required to create and tune the model to its best capabilities:

- *How can BOHB optimise reward tiers by balancing reward diversity with backer engagement in crowdfunding campaigns, thereby minimizing financial risks associated with underfunding?*

This question discovers why the characteristics of Bayesian semi-parametric approaches make BOHB suitable to apply in the optimisation of crowdfunding campaigns. It explains the specific mechanism of BOHB, and its applications in another context. It critically reviews its limitations and possibilities. This subchapter is therefore also part of the methodology, as it shows which underlying concepts and formulas are used.

- *How can LLaMA or other NLP techniques be utilised to identify key textual features in campaign descriptions and updates that enhance the prediction of crowdfunding campaign success?*

This question continues on the importance of language used in crowdfunding campaigns, as shown in the previous chapter. Moreover, it continues the methodology used in this thesis. It shows how the use of LLaMA could significantly improve the accuracy of the model, and explains the mechanism, its shortcomings, advantages, and applications behind this model.

- *How can the integration of BOHB and NLP models contribute to a more accurate prediction of campaign funding outcomes and minimise the financial risks of failing to meet funding goals?*

The answer to this question shows how BOHB and the use of NLP can be used together, and why it has such advantages. Furthermore, it completes the theoretic part of the models used in this thesis.

3.1 Systematic Review

A Systematic Review (SR) is carried out in three steps, as described earlier in Chapter 1. The goal of the SR is to find all relevant literature on the specific subject: the use of BOHB and NLP in the application of crowdfunding. The first step to conduct this SR is the planning stage, where all the search strings are defined to search the title, abstract and keywords in the databases of Scopus and ScienceDirect. Since the topic is novel, as BOHB was only used starting 2017 (Klein, Falkner, Springenberg, et al., 2017; L. Li et al., 2017), no date range needs to be selected. After this step, there remains a certain selection of papers which is further filtered in the two other stages. Since ScienceDirect only allows a maximum of eight Boolean search strings, the string is limited to use this number of operators. In the search strings there is put a focus on a crowdfunding context, or the reward-tiers, an optimisation algorithm, specifically Bayesian or HyberBand, plus some form of NLP. The following table shows the search strings applied and the number of results it gives, executed on the 14th of January 2025, in Table 1:

Stage 1:

Search Strings	Scopus:	ScienceDirect:
(“crowdfunding” OR “Kickstarter”) AND (“optimization” OR “BOHB” OR “Bayesian”) AND (“text analysis” OR “natural language processing” OR “NLP”)	3	198
(“crowdfunding” OR “reward-tier strategy”) AND (“Bayesian” OR “Hyper-Band”) AND (“Language Model” OR “text analysis”)	0	23
(“crowdfunding” OR “Kickstarter”) AND (“machine learning” OR “AI”) AND (“LLaMA” OR “Language Model” OR “text analysis”)	5	140
(“crowdfunding” OR “Kickstarter”) AND (“Bayesian optimization” OR “BOHB”) AND (“BERT” OR “LLaMA” OR “NLP”)	0	1
(“crowdfunding” OR “reward tiers”) AND (“optimization” OR “Bayesian”) AND (“Language Model” OR “NLP” OR “text mining”)	5	178
(“crowdfunding success” OR “Kickstarter”) AND (“Bayesian optimization” OR “BOHB”) AND (“Language Model” OR “text-based insights”)	0	1
(“crowdfunding campaigns” OR “Kickstarter”) AND (“reward-tier strategy” OR “optimization”) AND (“BERT” OR “LLaMA” OR “text analysis”)	1	36
(“crowdfunding” OR “Kickstarter”) AND (“AI-driven optimization” OR “Bayesian methods”) AND (“Language Model” OR “text analysis” OR “NLP”)	0	5

Table 1: Stage one SR: Search Strings

Stage 2:

In the second stage a review takes place of the papers. Here, only accessible articles are included, and only articles written in English. Only research articles and papers are included, others such as book chapters are excluded.

After this, the abstracts are read to filter the papers which are not relevant, the subjects which do not discuss crowdfunding nor ML. Subsequently, duplicates are removed, with the exclusion of duplicate articles from the results of queries and duplicate articles between the two databases (removed from the SD column). Finally, the papers are proofread, and the final selection is made by eliminating retracted or irrelevant articles. There are 6 elements on which the papers are checked, if it concerned the topic of crowdfunding, if reward-tiers are present, if they use ML methods, if they use textual analysis, if a LLM or something such as BERT is used, and lastly, if an advanced hyperparameter optimization strategy, such as Bayesian Optimization, Hyperband,

Scopus	SD	Access	Scopus	SD	EN	Scopus	SD	Research	Scopus	SD
14	582		10	556		10	556	articles	10	499

Scopus	SD	Subject	Scopus	SD	Duplicates	Scopus	SD	Reading	Scopus	SD
10	499	Abstract	7	106		7	46	papers	6	37

Table 2 (up): Stage two SR: part 1; Table 3 (down): Stage two SR: part 2

BOHB, or a comparable method, is applied. If at least three of those characteristics are present, the paper is considered relevant to the topic. In total, 43 papers remained relevant.

Stage 3:
No paper checks every box. This shows the newness of the subject. In the remaining selection, 22 papers matched three of the characteristics, 19 papers matched four, and only two papers had five of the criteria. Only seven studies used a LLM such as BERT, and solely three of the studies applied Bayesian Optimization or something similar. The full overview of papers can be found in Appendix B.

Most relevant papers
One relevant paper which matched 5 of the criteria, regards how linguistic styles (subjective vs. objective) in different sections of crowdfunding narratives influence funding success (W. Wang et al., 2021). Analysing over 328,000 Kickstarter campaigns, the authors used Bayesian inference and text mining to assess narrative structure across different sections (e.g., title, abstract, rewards, biography). Their results suggest that subtle linguistic positioning (e.g., placing objective content at the start of a description) can modestly but significantly affect campaign outcomes. The model incorporating these linguistic features improved prediction accuracy by 1.76 percentage points, from an already strong baseline of 76.37%. While the gain is modest, it demonstrates the value of incorporating structured language analysis in crowdfunding prediction models and highlights the potential of NLP-enhanced methods to fine-tune campaign effectiveness.
The other paper which checked five of the six criteria, employs the BERT model to assess the linguistic quality of crowdfunding descriptions, revealing how nuanced metrics like average BERT scores (indicative of lower writing quality) predict funding outcomes differently in “story” and “risk” sections. This paper underscores the significance of leveraging NLP models like BERT to uncover the nuanced impact of linguistic quality on funding outcomes, offering valuable insights for tailoring reward-tier descriptions to optimise backer engagement. Its findings suggest that integrating advanced AI techniques can enhance both the strategic design and textual refinement of crowdfunding campaigns, directly aligning with this research focus on optimising reward-tier strategies (C. S. R. Chan et al., 2021).

Furthermore, the abstracts of all 43 papers are analysed on their keywords. Using a Python script, the number of certain keywords is analysed. The result is shown in Figure 5. The keyword frequency analysis reveals that terms such as “crowdfunding” (170), “project” (125), “campaign” (57), “success” (53), “model” (36), and “description” (32) are highly prevalent across the selected abstracts. These frequencies reflect the

Criteria	**Number of papers**
Subject of Crowdfunding Campaigns	43
Reward-Tier Campaigns	16
Uses a form of Machine Learning	40
Uses a Textual Analysis	43
Makes use of an LLM & Similar (e.g., BERT)	7
Makes use of Bayesian Optimization & Similar	3

Table 4: Properties of different papers in SR

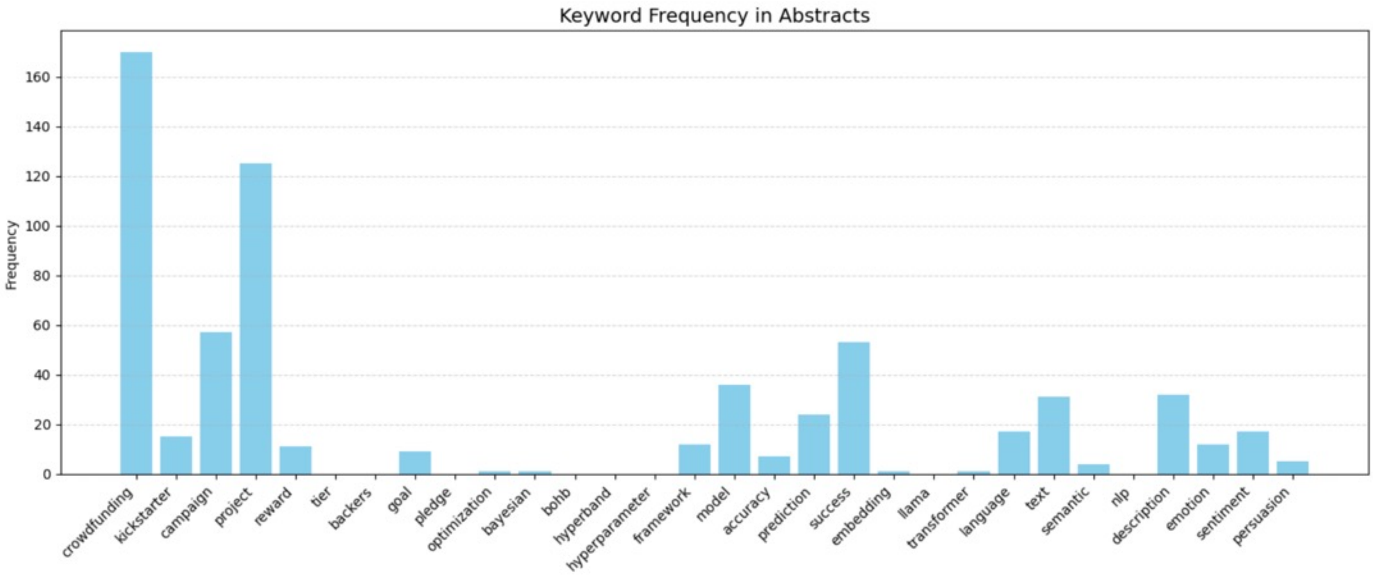


Figure 5: Keyword Frequency in Abstract

dominant focus of academic literature on understanding campaign-level dynamics, success factors, and modelling approaches in crowdfunding. Notably, keywords directly connected to this thesis, such as "reward" (11), "framework" (12), "prediction" (24), and "accuracy" (7), are also present but to a lesser extent, highlighting a gap in detailed optimisation strategies for reward-tier structures. Furthermore, the near absence of terms like "BOHB," "hyperparameter," "LLaMA," and "nlp" suggests that the integration of advanced hyperparameter optimisation and language embeddings remains an underexplored area.

3.2 Bayesian Optimisation and Hyperband

Hyperparameter Optimisation (HPO) aims to find the optimal of an unknown black box function, where either BO or random search is most often used (Cho et al., 2020). BOHB consists of two HPO methods: Bayesian Optimization and Hyperband. Short definitions and explanations are given with simple illustrations to explain the concept, where Appendix C gives more detail and the pseudocode. This section shows why these mechanisms are suitable for use in crowdfunding optimisation problems.
The technique Bayesian Optimization predicts how well a set of hyperparameters will perform, by using a probability model based on past results (L. Li et al., 2018). It relies on a surrogate function, which is a simpler, approximate model that is faster to evaluate than testing the actual hyperparameters. In addition, it uses an acquisition function to decide which set of hyperparameters to evaluate next. The acquisition function balances two goals: exploring new, less-tested options and exploiting promising configurations that are likely to perform well. By iterating this process, Bayesian Optimization identifies the best set of hyperparameters as its output (Bischl et al., 2023).

In this thesis, Bayesian Optimisation (and Hyperband) is used to tune parameters of other ML models, to boost their performance. Furthermore, it is extended to search for the optimal reward-tier structure in crowdfunding campaigns: such as the ideal number of rewards or average pledge amount. For example, when optimising the number of rewards, the algorithm treats this number as a variable whose effect on campaign success is not known in advance. It begins by evaluating a few configurations; say, campaigns with 3, 5, or 7 reward tiers, and measures how successful these are using an XGBoost prediction model. These are observations, displayed in Figure 6. The performance feedback from XGBoost helps Bayesian Optimization fit a surrogate model that estimates how different numbers of reward tiers affect campaign success. This is also shown in Figure 6. Then, the acquisition function selects the next number of rewards to test, balancing the trade-off between trying unexplored values (exploration) and refining promising ones (exploitation).

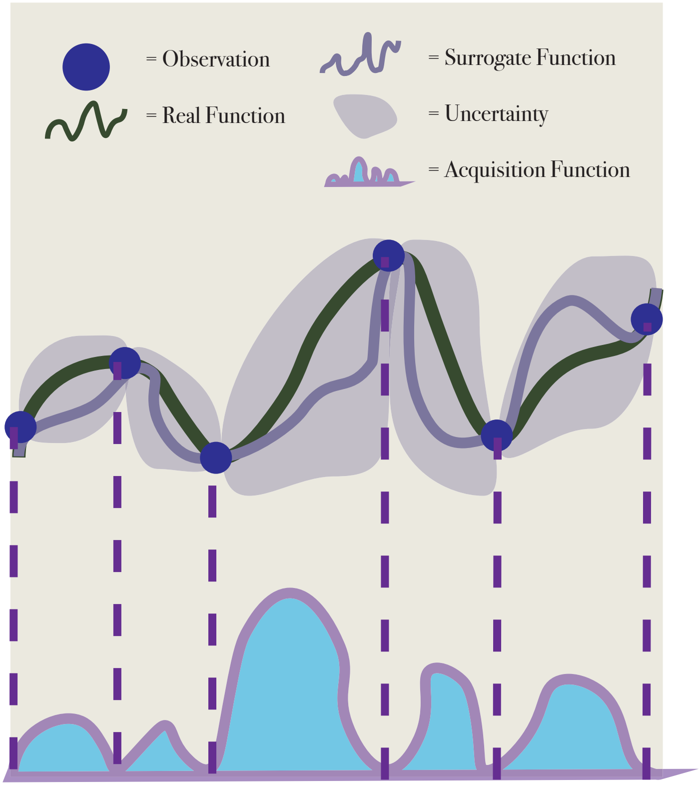


Figure 6: Simplification of Bayesian Optimisation

This process iteratively narrows in on the number of rewards that maximises predicted campaign success. This is more efficient than trying yourself or using a method like grid search. As mentioned, the figure accompanying this explanation illustrates the mechanism: it shows the actual performance curve (real function), the surrogate model, uncertainty across the space, and how the acquisition function chooses where to search next.

Hyperband is most closely related to a repeated process of Successive Halving (SH). This is a process where a fixed budget is assigned to configurations, also referred to as candidates. First, it uses low resources, but at every stage, the worst-performing configurations are removed, and a higher focus is put on the most promising ones. Ultimately, only the best configuration (candidate) remains. The budget can be divided between exploring a lot of candidates or evaluating fewer candidates with more detail (Bischl et al., 2023). The combination of both offers a faster convergence to optimal configurations. It uses the guided configuration sampling of Bayesian Optimization and Hyperband for efficient resource allocation, which balances the exploration and exploitation effectively (Falkner et al., 2018). Since crowdfunding is so high-dimensional, it makes it very suitable to apply in this context. A simplified representation of Hyperband is shown in Figure 7, where in the left of the image, it is shown how the budget of different candidates is re-evaluated from the situation on the left, to the right.

The combination of both offers a faster convergence to optimal configurations. It uses the guided configuration sampling of Bayesian Optimization and Hyperband for efficient resource allocation, which balances the exploration and exploitation effectively (Falkner et al., 2018). Since crowdfunding is so high-dimensional, it makes it very suitable to apply in this context.

Implementation and Libraries Used
Several studies have applied the BOHB algorithm to optimize ML models. One found setup involves splitting data into 80% for training and 20% for testing, with both sets scaled to have a mean of zero and standard deviation of one. This standardisation is particularly effective for gradient-based models, while tree-based models like XGBoost do not require it but can still benefit from normalisation to make their results easier to understand. However, another method which is often used is 10-fold cross-validation (often used in ML to reduce the bias) to evaluate hyperparameter configurations. Cross-validation tends to provide a more robust estimate of model performance by reducing the impact of data variability, making it especially valuable for smaller datasets or when overfitting is a concern. The performance metric to be optimised (e.g., validation loss or accuracy) is specified during coding. The choice of evaluation metric is crucial and should align with the objectives of the crowdfunding prediction models. Since this research focuses on classifying campaign success, metrics such as accuracy, F1-score, and AUC (Area Under the Curve) are prioritised for their ability to measure both overall correctness and the balance between precision and recall. Given the potential for class imbalance—where successful campaigns may be less frequent than unsuccessful ones—metrics like F1-score and balanced accuracy are particularly informative, ensuring that minority classes are accurately represented in the model's performance evaluation. Mostly, studies allocate a search budget of 50 trials/evaluations (Falkner et al., 2018; Sani et al., 2021).
For BO and other ML models for comparison, the following libraries are used:

- Scikit-learn: used for ML algorithms such as RF (Sani et al., 2021).
- Scikit-optimize: used for Bayesian optimisation with Gaussian Process and Random Forest surrogate models (Sani et al., 2021).
- Hyperopt: used for Bayesian optimisation based on a Tree-structured Parzen Estimator (TPE) surrogate (Sani et al., 2021).

Different libraries have been employed to implement the BOHB process:

- Optuna: uses a TPESampler for Bayesian optimisation and a HyperbandPruner to execute BOHB efficiently (Nguyen & Liu, 2025).
- Ray-Tune: enables large-scale, distributed BOHB optimisation and manages computationally intensive tasks effectively (Im et al., 2025).

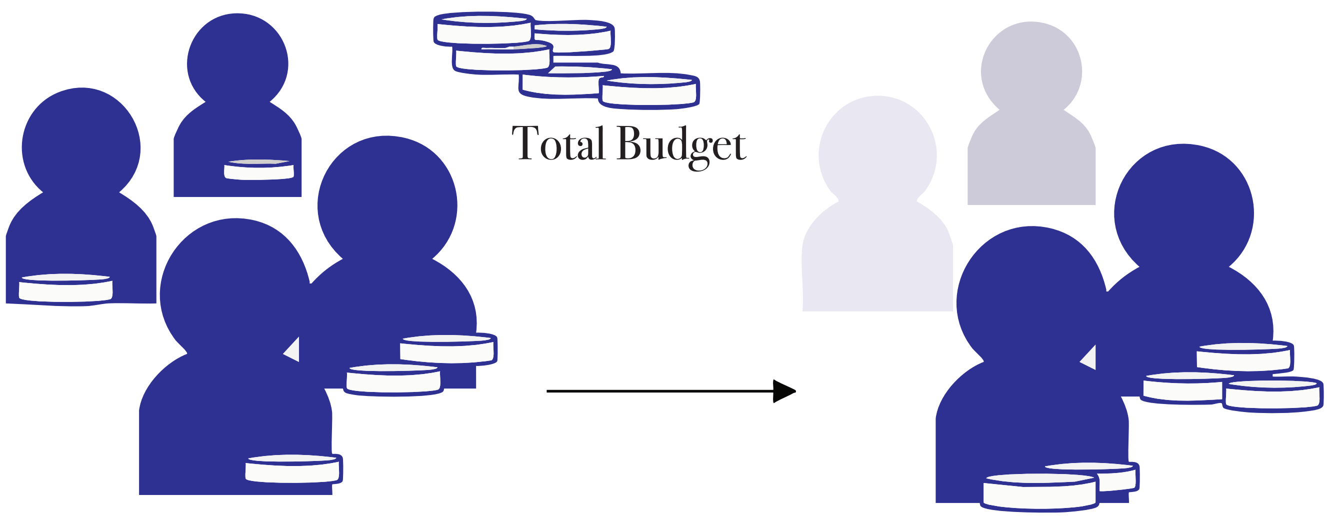


Figure 7: Simplification of Hyperband

For deep learning tasks, models are built and trained using PyTorch, with BOHB optimizing key hyperparameters such as learning rate, hidden layers, batch size, and activation function (Y. Zhang et al., 2024). Distributed frameworks like Ray-Tune ensure efficient exploration of large search spaces while minimising computational overhead.

Moreover, two studies have publicly shared their BOHB implementations:

- Falkner et al. (2018) provide an open-source implementation of BOHB and Hyperband, available at https://github.com/automl/HpBandSter. Their approach leverages the Multivariate KDE for modelling density distributions and uses random forests to estimate classification error and training time.
- Nguyen & Liu (2025) implemented BOHB using Optuna, which combines a TPESampler for Bayesian Optimisation and a HyperbandPruner for resource allocation. A step-by-step tutorial for replicating their case studies is provided in their paper.

Comparison with other HPO Methods

To evaluate the results obtained with BOHB, it can be compared with other optimisation methods like BO, with the three different surrogate models: Sequential Model-Based Algorithm Configuration (SMAC), TPE, or Spearmint (L. Li et al., 2018), as well as Hyperband (HB), random search (Sani et al., 2021), or greedy search (Im et al., 2025). Studies have shown that BOHB generally outperforms these models (Klein, Falkner, Bartels, et al., 2017). While some advanced models like FABOLAS can surpass BOHB, they typically require more computational resources (Klein, Falkner, Bartels, et al., 2017). Besides Bayesian methods, Genetic Algorithms (GAs), Particle Swarm Optimization (PSO), and hybrid models are also applied for HPO, but they are often more computationally expensive (Swaminatha Rao & Jaganathan, 2024).

Advantages:

By combining the strengths of both Bayesian Optimisation (BO) and HB, their drawbacks are eliminated as well (Joseph et al., 2024). While BO is sensitive to assumptions about the problem space, BOHB remains robust and improves sampling, while HB provides early stopping and efficient resource allocation (L. Li et al., 2018). This enables BOHB to find good configurations early in the optimization process (Falkner et al., 2018). One study by J. Yang et al. (2024) successfully applied BOHB to optimise window size, batch size, and layer units for car rental price prediction, handling multi-dimensional search spaces, limited computational budgets, and noisy outcomes. This illustrates BOHB's suitability for optimising reward tiers in crowdfunding, as it efficiently manages high-dimensional parameter spaces and noise (Falkner et al., 2018).

BO methods are already able to outperform random search (L. Li et al., 2017), but when combined with HB, it gives optimal results. BOHB is often selected due to its high performance, comparative computational complexity and the availability of asynchronous execution (Yaloveha et al., 2022). It has been shown that BOHB performs other HPO methods when tuning SVM and DL models (L. Yang & Shami, 2020). One study compared an optimised CatBoost model with BOHB to outperform all other tested optimisation methods (L. Lin et al., 2024).

Challenges and Limitations

While highly effective, BOHB has limitations in extremely noisy settings, where low-fidelity evaluations may be misleading (Falkner et al., 2018). Moreover, the reliance on the KDE could limit the scalability in very high-dimensional spaces. L. Yang & Shami (2020) mention that BOHB's evaluations with smaller budgets must be representative of the entire dataset to prevent slower convergence than standard BO models.

Some models are faster than BOHB, such as HyperSTAR—which evaluates 50% fewer configurations for better performance (Mittal et al., 2020), or MFES-HB, which achieves 3.3-8.9 speedups over BOHB (Y. Li et al., 2021). Lastly, Hyper-Tune is also shown to achieve a speedup of 11.2x over BOHB (Y. Li et al., 2022). However, there is a scarcity in literature on these methods, and they have not been tested as extensively as BOHB. For example, HyperSTAR has only been tested on image classification datasets (Mittal et al., 2020).

Evaluation/Empirical Benchmarking

BOHB can be benchmarked against other HPO methods using standard libraries like HPOlib (L. Li et al., 2018). For comparison, mean MAE, RMSE, and R2 scores are often averaged across multiple samples, as demonstrated by Sani et al. (2021), and they also ran the default hyperparameter settings in the Scikit-learn package or XGBoost. Other evaluations include Explained Variance Score (EVS), Mean Absolute Error (MAE), R2, and Mean Absolute Percentage Error (MAPE) (Ling et al., 2022).

3.3 LLaMA and key textual features

The rapid advancements in Natural Language Processing (NLP) have significantly improved textual analysis and decision-making in finance. NLP represents and analyses human language computationally. It consists of two components: Natural Language Understanding (NLU) and Natural Language Generation (NLG). The former is the ability to analyse language in various ways, and the latter to produce language itself. NLP is used in various ways, such as machine translation, text categorisation, information extraction or summarisation. Nowadays, there are a lot of databases available with labelled data which can for example be used for a sentiment analysis (Khurana et al., 2023). Large Language Models (LLMs), such as LLaMA (Large Language Model Meta AI), released in February 2023 (Minaee et al., 2024), represent a major breakthrough in NLP due to their ability to process and generate human-like text at scale, and excel in both NLU and NLG. LLMs are advanced language models, and they have massive parameter sizes and very high learning abilities (Chang et al., 2023). LLaMA is currently one of the best performing LLMs, comparing with the performance GPT-4 and Claude 3.5 Sonnet (Grattafiori et al., 2024), and it is an open-source model, allowing everyone to use it for free. LLaMA was trained on several datasets, of which two-third was Common Crawl (an open repository of web data), but also C4, Github, Wikipedia, books, ArXiv, and Stack Exchange (Le et al., 2023). Using an LLM has several advantages. First, it is capable to handle sequential data efficiently, which makes it capable of handling parallelisation and capturing long-range dependencies in the text. Another benefit, not utilised in this thesis, but useful to name for future research, is Reinforcement Learning from Human Feedback (RLHF). This allows the model to be finetuned due to user input (Chang et al., 2023). It natively supports multilinguality, coding, reasoning, and tool usage (Grattafiori et al., 2024).
All LLM are based on a transformer architecture, where there is a certain input, the source sequence, received by an encoder, that converts it into a vector that represents the semantics. Subsequently, a decoder processes this vector to the desired output language. The terminology can be quite ambiguous, since some papers like Le et al. (2023) mention a model like GPT-3 has these two components, while their architecture is slightly different, since it does not explicitly contain an encoder. LLaMA and GPT-4 are autoregressive language models. They predict the next word based on the previous words; they only have a Decoder. BERT is a masked language model, where it predicts omitted words in a sentence based on the previous and next words. It makes use of a Bidirectional Encoder. Last, BART (Bidirectional and Auto-Regressive Transformer), is a hybrid model in-between the two categories, and uses an encoder-decoder architecture. It corrupts a sequence and then predicts the original sequence (Min et al., 2023). In Figure 8, the 3 different architectures are visualised. The Transformer architecture (Vaswani et al., 2017) relies on self-attention to process text efficiently and capture long-range dependencies. More details on this architecture are stated in Appendix D. This foundational model has enabled significant improvements in NLP tasks.

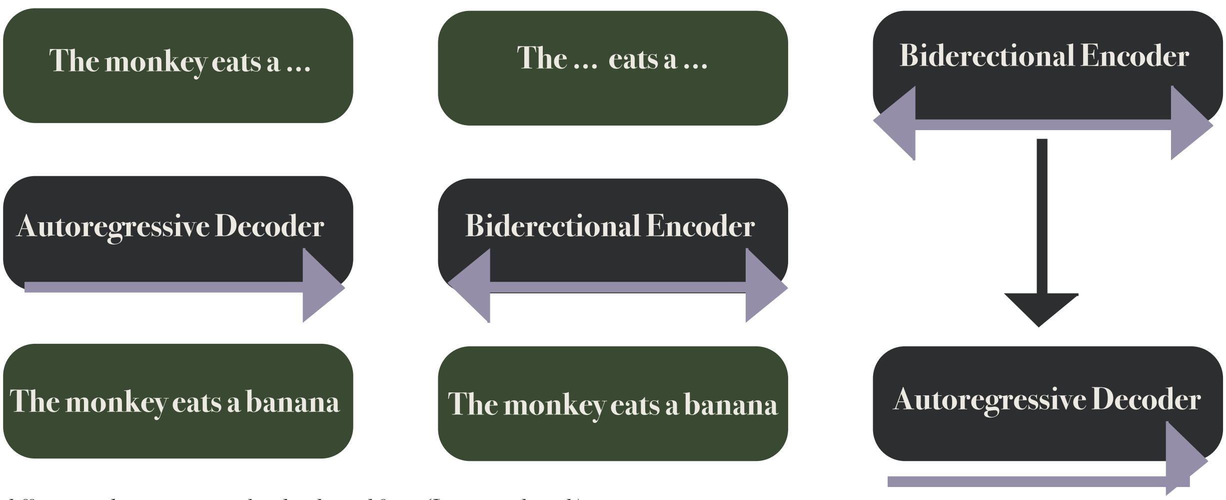


Figure 8: different architectures visualised, adapted from (Lewis et al., n.d.)

Text Analysis for Crowdfunding

Prior research has shown that textual factors, such as project descriptions, peer review sentiment, and linguistic patterns significantly influence crowdfunding success (L. Li et al., 2022; Peng et al., 2022; K. F. Yang et al., 2023). This demonstrates the relevance of deploying NLP on crowdfunding data, aligning with findings that textual signals impact campaign outcomes. In this thesis, we make sure to implement a Sentiment Analysis.

Sentiment Analysis identifies the emotional inclination of a text, either with three classes (positive, neutral or negative) or binary (positive and negative). We do note that LLaMA 3.2 faces challenges in handling nuanced expressions like sarcasm and irony (Buscemi & Proverbio, 2024). Therefore, we aim to use an additional model to capture specifically the sentiment. This dual-model approach makes sure that there is a deep contextual understanding and accurate sentiment detection.

To achieve optimal representation of textual features, we do apply LLaMA 3.2 embeddings to campaign descriptions and reward structures. This enables the identification of influential keywords and linguistic signals that contribute to predictive modelling. Unlike traditional embeddings like Word2Vec and GloVe (Otter et al., 2021), LLaMA 3.2's Transformer-based architecture allows it to understand long-range dependencies and contextual relevance more effectively.

Embedding Approach

Recent developments in NLP have led to rapid evolution of models. Traditional approaches such as n-gram analysis and penalized logistic regression were still used in the context of crowdfunding as recently as 2022: Peng et al. (2022) examined linguistic factors in crowdfunding success with these methods. However, transformer-based models have since become the new standard, as they can capture contextual and semantic nuance beyond simple word frequencies.

In this thesis, embeddings, which are dense vector representations of textual inputs, are used as features for ML, rather than relying on generative capabilities of LLMs. These embeddings are crucial for capturing semantic and emotional tone in crowdfunding campaign descriptions and reward tiers. In this thesis we chose to compare 2 different embedding models; sentence-transformers/all-MiniLM-L6-v2 (All-MiniLM-L6-V2, n.d.) and LLaMA 3.2 – 3B Instruct (Llama-3.2-3B-Instruct, n.d.). The former is a lightweight transformer model optimised for sentence similarity and semantic search. It is widely used for embedding short texts due to its fast runtime, small memory footprint, and excellent performance on classification and clustering tasks. According to the Massive Text Embedding Benchmark (MTEB) (MTEB, n.d.), MiniLM ranks among the top-performing open-source models in multiple sentence-level tasks—making it ideal for structured Kickstarter texts.

LLaMA 3.2 – 3B Instruct is a compact yet powerful LLM from Meta, as explained earlier. In this thesis, it is solely used as a feature extractor. LLaMA is chosen for its parameter efficiency, open-source availability, and strong performance on semantic benchmarks. While typically used for generation, the model is repurposed in this work to convert campaign texts into meaningful embeddings, making them suitable for downstream prediction tasks with models like XGBoost.

These models were selected not only for their technical strengths but also for their practical advantages. Unlike OpenAI's text-embedding-ada-002, which is a paid, closed-source model requiring API access, both LLaMA and MiniLM are free and reproducible, aligning with the academic standards of transparency and local experimentation.

Other Embedding Models

While models like BERT, RoBERTa, and GPT-4 are also well-regarded, each comes with trade-offs. BERT remains a strong benchmark, especially when computational resources are limited. GPT-4 is far more powerful, but primarily designed for generation, with high inference costs and limited access. Additionally, while OpenAI's text-embedding-ada-002 performs well in general-purpose applications, its usage incurs costs and lacks full local control.

In contrast, LLaMA 3.2 – 3B outperforms GPT-3 (175B) on many NLP benchmarks (Minaee et al., 2024) and smaller models in the LLaMA family have demonstrated superior alignment in terms of both helpfulness and harmlessness (Grattafiori et al., 2024). This makes it a strong candidate for embedding extraction. We do acknowledge that LLaMA has some shortcomings. Since LLaMA is not specifically optimized for fine-grained sentiment detection (Buscemi & Proverbio, 2024), this work supplements its embeddings with a dedicated sentiment analysis model: cardiffnlp/twitter-roberta-base-sentiment (Barbieri et al., 2020). This transformer-

based model from Hugging Face is fine-tuned on large-scale social media sentiment data, which closely resembles the tone and brevity of Kickstarter campaigns. It effectively classifies text into positive, neutral, or negative categories, offering a clearer signal of emotional appeal in campaigns.
While the approach to compare LLaMA and MiniLM for this thesis, as these are two well-known performing models, in the future other embedding models could be researched, as there are 276 models shown in the MTEB. The Discussion chapter will discuss this further.

Fine-tuning and Adaption
While LLaMA supports advanced fine-tuning techniques, the focus of this thesis is on its embedding capabilities for predictive modelling rather than full fine-tuning. We still discuss the possibility, since this could be used for later research. Fine-tuning large language models like LLaMA is computationally intensive, requiring significant resources to update millions of parameters. Recent methods like delta-tuning (Ding et al., 2023) optimize a small subset of parameters while freezing the rest, reducing computational costs. Techniques such as Low-Rank Adaptation (LoRA) further minimize resource consumption by breaking down attention weight updates into smaller, more manageable components (Ding et al., 2023). While these methods are state-of-the-art for efficient fine-tuning, they are not applied in this thesis due to the focus on LLaMA pre-trained embeddings instead of model retraining. LLaMA has many possibilities to be optimised or finetuned. Another way of implementing LLaMA is prompt engineering. This is a practical alternative for task-specific guidance without altering model parameters. P. Liu et al. (2023) highlight how structured prompts can guide model responses, either through manual template engineering or automated learning strategies. Prompts can be discrete, using actual words, or continuous, operating directly within the model's embedding space. This method is efficient, as it enhances adaptability without the computational burden of fine-tuning.
In this thesis, hyperparameter optimisation of the success-prediction model (e.g. with XGBoost), in combination with the embeddings, is prioritised over fine-tuning and prompt engineering. To enhance prediction accuracy and generalisation without the computational cost of full model adaptation, only LLaMA embeddings are used. However, there is reflected on the possibility to use these methods in the Discussion of this thesis.

Implementation and Libraries used
Meta provides all LLaMA3 models freely on https://llama.meta.com. They openly release the flagship model, with pre-trained and post-trained versions to accelerate a responsible path towards the development of artificial general intelligence (Grattafiori et al., 2024).
In this study, we use PyTorch 2, an advanced version of PyTorch that integrates TorchDynamo, a Python bytecode transformation tool that optimises computation at runtime, and TorchInductor, a compilation backend that generates high-performance code for GPUs (CUDA) and CPUs. These enhancements improve model efficiency, reducing computational overhead while maintaining PyTorch's flexibility. PyTorch 2's optimisations help improve runtime efficiency and scalability, which supports smoother integration of large models like LLaMA during embedding generation. (Ansel et al., 2024). Additionally, Hugging Face plays a crucial role in modern NLP research, offering a unified ecosystem for transformer-based models. The Transformers library provides seamless access to pre-trained models like BERT, GPT, and LLaMA, enabling fine-tuning, evaluation, and dataset management (Wolf et al., 2020). The library integrates seamlessly with PyTorch, simplifying the use of pre-trained models like LLaMA and MiniLM for embedding extraction in this thesis.. The exact methods and more specific libraries are described in Chapter 5.

Advantages and Challenges
One advantage of LLaMA is that it is completely open source. There are a few other models (OPT, GPT-NeoX, BLOOM, and GLM), but they are not competitive with the state-of-the-art. Second, they offer both smaller and bigger models. Their smallest model, LLaMA-13B, outperforms GPT-3 on many benchmarks despite being ten times smaller. The model can be run on a single GPU, and their 65B-parameter model is competitive with advanced models such as Chinchilla or PaLM-540B (Touvron, Martin, et al., 2023). LLaMA has been evaluated against state-of-the-art language models across multiple NLP benchmarks, demonstrating strengths in efficiency and knowledge-intensive tasks while revealing areas for improvement. Although originally

evaluated on reasoning benchmarks, LLaMA's strong performance on a variety of NLP tasks highlights its ability to capture complex semantic patterns: useful when converting crowdfunding narratives into embeddings. While LLaMA lags behind in long-context QA tasks, its strong factual representation ability remains valuable for summarising shorter texts like Kickstarter campaigns. Its efficient language modelling, shown by low perplexity scores, indicates a strong internal representation of language, making LLaMA embeddings a good fit for text-based machine learning models. (Touvron, Lavril, et al., 2023).

Empirical Evaluation and Benchmarking
Evaluating NLP models is challenging because natural language is inherently flexible, meaning the same information can be expressed in multiple valid ways. Normally, models require specialised benchmarks such as GLUE and SuperGLUE to assess their performance (Chang et al., 2023). Since we are only dealing with classification tasks, we solely need to include the following:

- Accuracy, Precision, Recall, and F1-score – Frequently used to assess model correctness in text classification, sentiment analysis, and named entity recognition (Alomari, 2024).
- Confusion Matrices & Agreement Rates – Used to compare model outputs and understand misclassification trends (Repede & Brad, 2024).

3.4 The Integration of BOHB and NLP

While LLaMA on its own is a powerful model for generating high-dimensional text embeddings, it does have some shortcomings, particularly for nuanced sentiment analysis (Chang et al., 2023) and logical inference (Rawte et al., 2023), since it has difficulties with nuanced expressions such as sarcasm and can have an optimistic bias (Buscemi & Proverbio, 2024). For analysing crowdfunding campaigns, this can be an issue, since it highly relies on textual features (L. Li et al., 2022; Peng et al., 2022; K. F. Yang et al., 2023). One measure to overcome these constraints, is to integrate BOHB to optimize the hyperparameters of RF, MLP and XGBoost, which are driven by LLaMA 3.2 embeddings. HPO improves the accuracy of predictive models by fine-tuning parameters such as learning rate, number of layers, and batch size (Inman et al., 2023). Automating this selection process through BOHB not only enhances predictive accuracy but also significantly reduces computational overhead.
This subchapter will ultimately answer this question: *How can the integration of BOHB and NLP models contribute to a more accurate prediction of campaign funding outcomes and minimise the financial risks of failing to meet funding goals?* This is the final theoretical research part of this thesis.

Applications of BOHB in Text-based Models
BOHB is often used to find the optimal parameter settings for ML models. Relevant applications of BOHB include:

- Text classification and sentiment analysis: optimising LSTM and Transformer models to improve sentiment detection and text categorization (Genc et al., 2020; Han et al., 2020)
- Document classification: enhanced accuracy in document labelling tasks though efficient hyperparameter searches (J. Guo & Li, 2019).
- Language modelling: finetuning hyperparameters for sequence generation tasks, improving fluency and coherence (Mittal et al., 2020).

The use of BOHB for optimising ML models in this study is designed to maximise the representation power of the LLaMA embeddings, enhancing the prediction accuracy for campaign success.

Compatibility of LLaMA embeddings and BOHB
LLaMA generates high-quality embeddings that contain the semantic meaning of the campaign description and the reward tiers. However, to leverage these embeddings effectively, we need well-optimised ML learning methods tuned by BOHB. Several key benefits make BOHB particularly well-suited for fine-tuning ML models, as well as using it combination with many embeddings:

- Efficient hyperparameter search: BOHB balances exploration and exploitation through its acquisition function and guided sampling (Bischl et al., 2023). Instead of relying on grid search or random search, it intelligently selects hyperparameter configurations that are likely to yield better performance, reducing the number of required evaluations.
- Resource-efficient training: the Hyperband component of BOHB ensures that poorly performing configurations are pruned early, reducing the computational burden. This strategy has been successfully applied in domains such as arrhythmia detection (Han et al., 2020) and reinforcement learning (Baek et al., 2023), where it significantly decreased the training time.
- Scalability for large models or large embeddings: BOHB has been used effectively in high-dimensional optimisation problems, such as material science (Y. Zhang et al., 2024) and energy applications (How et al., 2022). Given the complexity of the large amount of LLaMA embeddings, BOHB's scalability ensures that hyperparameter tuning remains computationally feasible even with large parameter spaces.

The suitability of BOHB for this thesis is further reinforced by its ability to optimise both continuous and discrete hyperparameters, handle noisy performance evaluations, and dynamically allocate resources where they are most needed (Falkner et al., 2018).

Implementation Strategy

The implementation strategy of this thesis is complex but can be reduced to the following steps:

- A transformer-based model from Hugging Face, cardiffnlp/twitter-roberta-base-sentiment, is used to classify each text as positive, neutral, or negative.
- Model implementation: LLaMA 3.2 is integrated with Hugging Face's Transformers, to generate embeddings from textual data, such as campaign descriptions and reward tiers.
- Hyperparameter optimization: the main BOHB will be executed with Optuna, as this has most documentation available and is easy to integrate with HuggingFace and Pytorch (Nguyen & Liu, 2025). BOBH is used for both tuning the ML models (RF, MLP, and XGBoost) as well as the parameters in reward tier strategies.
- Evaluation Metrics: Performance is measured using Accuracy, Precision, Recall, F1-Score, and AUC-ROC to evaluate campaign success predictions effectively.

This pipeline ensures that there is a high computational efficiency, plus accounts for a high predictive accuracy. It forms a balance between computational resources and performance.

Expected Advantages and Challenges

BOHB has proven to be an effective HPO method across various domains, as well with the integration with NLP. It has the advantages of a faster convergence, a reduced computational cost. Studies (Azadi et al., 2025) have shown that integrating BOHB with Monte Carlo sampling techniques enhances accuracy and efficiency, a reduced computational cost, and has a better generalisation on NLP tasks, reinforcement learning studies (Baek et al., 2023) highlight its effectiveness in optimizing NN layers for text-based tasks. However, it can also pose some challenges associated with its application:

- Computational cost of multiple trials: It can make optimization costly in complex domains like stock market predictions (J. Guo & Li, 2019).
- Risk of overfitting: BOHB's iterative tuning may lead to overfitting if not carefully managed. Studies (Nguyen & Liu, 2025) report high accuracy but prolonged computation times.
- Trade-off between efficiency and performance gains: the efficiency gains provided by BOHB may sometimes come at the cost of performance consistency. Studies on battery state-of-charge predictions (How et al., 2022) found that while BOHB optimised model complexity, trade-offs between accuracy and computational efficiency had to be balanced.

By leveraging BOHB, with careful management of computational trade-offs and potential overfitting risks, we can successfully employ this technique to find the optimal reward tier structure.

3.5 Concluding words Chapter 3

This chapter has established the novelty of using BOHB alongside LLaMA embeddings for optimising reward tiers in crowdfunding. Through the Systematic Review (SR), it was demonstrated that no prior studies have explored this specific combination, confirming the unique contribution of this research.
LLaMA is one of the best LLMs available, and due to its open-source availability, it is favoured over the other models. It is used to generate embeddings to catch the semantic meaning of the language used in the reward structures. Integrating BOHB with LLaMA embeddings, alongside hyperparameter tuning of ML models is a strategic approach to enhancing the model's efficiency and performance in a domain where the use of language has high importance. BOHB offers a structured and resource-efficient optimisation method, balancing exploration and exploitation to identify the most effective configurations.
With this foundation, the model set-up can be performed by implementing LLaMA embeddings in HuggingFace, defining the hyperparameters by BOHB in Optuna, and eventually evaluating the model with metrics such as the F1 score. The chapters following this chapter will explore the data used in this thesis and test the hypothesis of the expected performance of LLaMA embeddings incorporated in a tuned ML model (RF, MLP, or XGBoost) by BOHB, alongside the optimisation of the reward parameters performed by BOHB as well.

4. Data Collection and Preparation

This chapter outlines the data collection process, and the key decisions made during data preparation. The dataset combines structured campaign data from (Web Robots IO, 2025) with additional reward-tier information collected via a custom browser-based Tampermonkey userscript. This additional data is essential to reward-specific information required for analysing and optimising reward configurations.

Following collection, the data is cleaned and formatted to ensure it is complete, consistent, and ready for analysis. An Exploratory Data Analysis (EDA) was then conducted to evaluate the dataset's representativeness, identify anomalies, and uncover initial patterns, supported by visualisations to enhance interpretability. We reflect on the literature if the data matches previous studies.

Subsequently, the dataset is tailored for ML using LLaMA embeddings and BOHB. Feature selection is applied to retain only relevant variables for predicting campaign success. Sentiment analysis is performed to extract emotional tone from campaign text, and textual fields are vectorised into numerical formats interpretable by ML models.

Finally, the data is split into training, validation, and test sets to ensure robust evaluation. The resulting feature matrix—composed of numerical, categorical, and semantic representations—forms the foundation for the predictive modelling and optimisation in the following chapters.

4.1 Data Collection & Description

The data used in this research, consisting of information from over 14000 Kickstarter campaigns, is collected through a combination of publicly available data provided by Web Robots IO (Web Robots IO, 2025) and custom webscraping using a browser-based Tampermonkey userscript. This section outlines the data sources, scraping methodology, limitations, and the key variables included in the final dataset.

Data Sources

The primary dataset employed in this research is obtained from Web Robots IO, a widely used source for publicly available Kickstarter data. Since March 2016, they have run monthly crawls using an automated scraper collecting Kickstarter campaign data in both CSV and JSON formats. The version used for this thesis is scraped in March 2025 and includes data on more than 50,000 campaigns, divided into 50 CSV files. Each file has a consistent schema, with 42 columns detailing campaign attributes such as the number of backers, funding goals, currency conversions, campaign duration, creator metadata, project descriptions, and categorical classifications.

The campaigns in the dataset span Kickstarter activity from 2009 to 2025, mainly covering projects launched in English-speaking regions including the United States, United Kingdom, Canada, and Australia. Fortunately, this dataset contains direct URL links to each campaign's reward page. These links proved essential for extending the dataset with tier-specific reward information, which is not available in the original structured files.

To enable analysis of reward-tier configurations at scale, it is necessary to extend the dataset by collecting granular reward-level data for each campaign. This includes the tier descriptions, pricing, estimated delivery dates, and number of backers per tier. These are all features one needs to optimise a reward tier strategy.

Scraping Method

There are multiple methods explored to retrieve the needed reward-specific data. These include sending authenticated requests to Kickstarter's GraphQL API using session cookies and CSRF tokens, manually extracted from browser sessions. However, all of the attempts to collect the data resulted consistently in 403 Forbidden errors, largely due to Kickstarter's use of Cloudflare protection, which uses powerful bot-detection mechanisms such as TLS fingerprinting and dynamic header validation. Subsequent efforts using Selenium with ChromeDriver, both in standard and "undetected" configurations, were partially successful in loading Kickstarter pages, but still failed to bypass Cloudflare's JavaScript challenge and CAPTCHA system. Manual attempts to extract updated session cookies and use them in scripted HTTP requests also failed to yield reliable access, indicating that Kickstarter's security architecture required more than just static cookie headers to authorise requests.

Ultimately, the reward data is successfully collected using a custom-built Tampermonkey userscript that runs directly in the browser. This script operates within a legitimate browser session, thereby inheriting all the necessary authentication and bypass mechanisms already completed by the user. The script processes batches of 200 Kickstarter campaign reward pages at a time, scraping the visible content from each page. The most important information such as the URL, reward title, price, number of backers, expected delivery date, detailed description, and included items, are extracted and compiled into structured CSV files. In total, 75 batches were processed manually, resulting in an extensive dataset of reward-tier information across approximately 14,000 campaigns.

Limitations

Although Web Robots IO also offers datasets for Indiegogo, the data quality and structure are significantly more limited. Indiegogo datasets do not include reward page URLs, making it infeasible to extract detailed tier-level data in a scalable manner. For this reason, Indiegogo campaigns are excluded from the scope of this thesis.

In addition, the study focuses exclusively on Kickstarter campaigns within the Technology category. This choice is made based on literature and the big share this category offers. Technology projects tend to have more complex and varied reward structures, and prior research has shown that the integration of textual, visual, and social features significantly enhances predictive performance, particularly in technology-driven campaigns where narrative plays a central role in attracting and engaging backers (Babayoff & Shehory, 2022; Mitra & Gilbert, 2014).

Key Variables

In summary, the dataset used in this thesis comprises roughly 14,000 Kickstarter technology campaigns, enriched with additional reward-specific information. For each campaign, the following key variables are available:

From the reward-tier scraping process:
- The number of backers per reward tier,
- The reward description,
- The monetary price of each tier,
- The expected delivery date of each reward,
- The title and label of each reward,
- The individual items or benefits included in each tier,
- The original URL to the reward page (for verification or manual inspection).

From the original Kickstarter dataset (Web Robots IO):
- The campaign title and short description (blurb),
- The full project description (textual content),
- The funding goal and amount pledged,
- The number of backers and funding success status,
- The campaign launch and deadline dates,
- The category and subcategory of the campaign (e.g., Technology > Gadgets),
- The creator information and project location (e.g., country, city),
- The currency and converted pledged amount (USD),
- The campaign duration (in days),
- Various social and visual signals, such as presence of a video, images, or FAQs.

This combined dataset, which combines both structured and unstructured features, forms the basis for the subsequent analysis, with the goal of understanding and optimising reward tier configurations in crowdfunding campaigns.

4.2 Data Preprocessing & Cleaning

This section shows the data cleaning and preprocessing procedures after collecting and scraping campaign and reward information from Kickstarter. These steps ensure the data can be used without any errors and is representable, needed for the exploratory analysis and ML development.

Original Dataset
The original dataset is compiled using data provided by Web Robots IO, which systematically scrapes Kickstarter project data from their public website. Initially, the dataset included information on 16,909 campaigns. However, due to Web Robots IO's strategy of scraping multiple subcategories-rather than top-level categories exclusively, duplicate entries were introduced. Since Kickstarter enforces rate limits on API and page requests, scraping in a single run is impossible. As a workaround, smaller subcategories are scraped independently, which increases coverage, causing some duplicates. The duplicate campaigns are identified and removed, resulting in 14,906 unique campaigns. Of these, 21 campaigns have reward pages that were inaccessible at the time of scraping, likely removed by the project creators, yielding a final usable base of 14,885 campaigns. Of these campaigns, a total of 100,599 individual reward tiers were collected using a custom Tampermonkey browser script, preserving the one-to-many structure between campaigns and their associated rewards. This scraping process creates two distinct but linked datasets: the campaign information, and the reward information. This separation is intentional, as one campaign could contain even more than 30 rewards, making it difficult to integrate the datasets into one.

Data Cleaning Procedures
Both datasets are cleaned fully to ensure quality, uniformity, and suitability for ML tasks. Key cleaning operations include:

- Handling Missing and Erroneous Values
  - o Rows missing critical information (e.g., reward URL, goal, pledged amount, or campaign dates) are removed.
  - o Infinite and implausible numeric values are replaced with NaN and, where appropriate, imputed with neutral values such as zero.
- Format Standardization
  - o All date fields (e.g., created_at, deadline, Delivery) are parsed into datetime objects.
  - o Campaign duration is derived and supplemented with the launch year.
  - o Monetary values (goal, pledged, reward prices) are converted into USD using the provided fx_rate column, producing consistent *_usd fields.
  - o A binary indicator funding_success is added to label campaign outcomes.
- Categorical and Textual Field Cleanup
  - o Categorical fields (e.g., country, category) are cleaned and standardised in casing.
  - o The price fields are stripped of currency symbols and cleaned of noise.
  - o All reward prices are converted to floating-point values to ensure compatibility
- Structural Integrity and Dataset Separation
  - o A reward_count column is added in the campaign-level dataset to quantify the number of tiers per project.
  - o The campaign statistics are kept separate of the reward statistics, to ensure they are only counted as one entry and not X * number of rewards

The result is a pair of fully cleaned and harmonised datasets-one structured for reward-level analysis and the other for campaign-level modelling.

4.3 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is an important first step in the data analysis process. It summarises the main characteristics of the dataset using statistical and visual methods in order to better understand the structure, detect patterns, identify anomalies, and generate hypotheses (Exploratory Data Analysis, 1977). To quote John Tukey: “The greatest value of a picture is when it forces us to notice what we never expected to see”. In this section, key variables from both the campaign and reward tiers are explored through descriptive statistics and graphical representations, made with Tableau and Python. Particular attention is given to how these features relate to funding success, reward pricing behaviour, and textual elements like campaign descriptions.

Data in a nutshell

First, we present some general descriptive statistics to describe the reach and the characteristics of the dataset.

Number of Campaigns	14,885
Average number of rewards	6.76
Average pledge amount per campaign	$55998.20
Average backers per campaign	282.24
Average backers per successful campaign	577.45
Success Rate	47.06%
Average Campaign Duration	35.92 days

Table 5: General Statistics Dataset

Aside from these statistics, we examine the geographical reach, which is shown in Figure 9. There might be regional differences for the campaign success characteristics. In this case, there is a large dominance of campaigns in the United States. This may imply cultural standardisations for rewards. Plus, the success rates differ quite substantially per country, as can be seen in the Appendix E. Furthermore, in Figure 10, the different subcategories of the campaigns are presented. The bigger the circles, the higher the average success rate per subcategory. Some of the numbers are quite remarkable, and therefore, the number of entries is checked per subcategory. Makerspaces contains only 22 campaigns, and therefore, the sample size seems too small to draw conclusions. However, subcategories such as Gadgets and Hardware contain 1760 and 428 campaigns respectively. These subcategories are most successful but can be largely influenced by the period of time. Literature shows the category of the campaign influences the campaign success: different categories have different characteristics that perform well, and a certain category membership can offer an advantage because of a ‘blockbuster’ (Butticè et al., 2018).

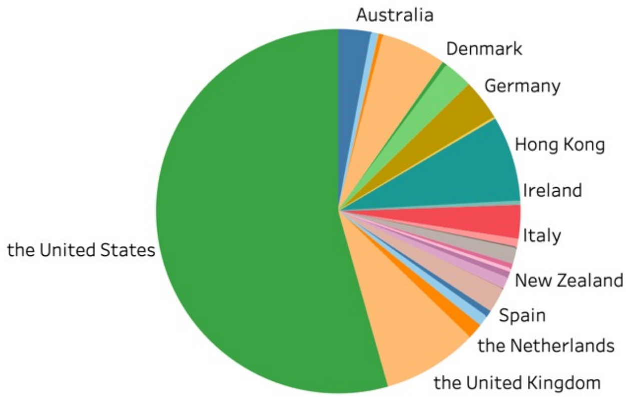


Figure 9: Country distribution of dataset

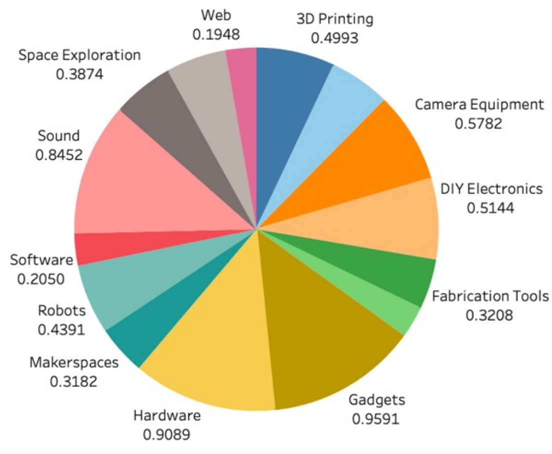


Figure 10: Sub-category distribution of dataset

Reward Tier Behaviour

The central focus of this thesis is the reward tier structure; thus, it is quite important to examine the general characteristics of the campaigns regarding this specific subject. Are they quite uniformly distributed, or are there many different set-ups? First, there is the number of rewards per campaign. According to literature, having multiple reward options is important (Kunz et al., 2017) and in one study, the most important

feature in their outcomes was the number of reward options (Zhong, 2022). In Appendix E, the number of rewards per campaign. It shows most campaigns have between 1 and 15 rewards. As shown, the most optimal reward number is determined to be 10 (Cai et al., 2021), many campaigns are still below this number. Furthermore, the pricing of the rewards is shown in Appendix E. A closer inspection of the reward price distribution reveals sharp spikes at psychologically appealing price points such as $99, $199, $299, and so on. This pattern reflects the widespread use of the 9-ending pricing strategy, a well-established psychological pricing technique used by marketers to influence consumer perception. According to Kumar & Pandey (2015), such prices are cognitively perceived as significantly lower than the next rounded threshold, for example, $199 is associated more with $100 than with $200. This strategy exploits the "level effect", where consumers unconsciously round down prices, and the "image effect", where 9-ending prices suggest promotions or value deals. The spikes observed in the data suggest that many creators adopt this strategy, possibly to maximise backer appeal at perceived value points, especially among price-conscious or low-involvement consumers. However, for crowdfunding this is only beneficial for a maximum of 4 options (Keisar & Lev, 2023). In Figure 11, the price distribution of the rewards is shown, for successful campaigns and for unsuccessful campaigns. For successful campaigns, the average reward price is slightly lower, but not enough to state that lower reward prices have a higher chance of success.
The average number of backers per reward is 41.75, however, once examined more closely, this is not closely located to the median. In Appendix E it is shown most rewards have 0 up to 10 backers, and there are a few outliers that are referred to as the so-called 'blockbusters'. One noticeable fact is the small spikes at 50, 100 and 200 backers. This is most likely because creators of campaign can create a maximum number of backers which is used most often for early birds.

Campaign Dynamics
Beyond reward-specific features, the broader structure and configuration of a campaign plays a critical role in determining its likelihood of success. These dimensions cannot be ignored as they interact with the reward

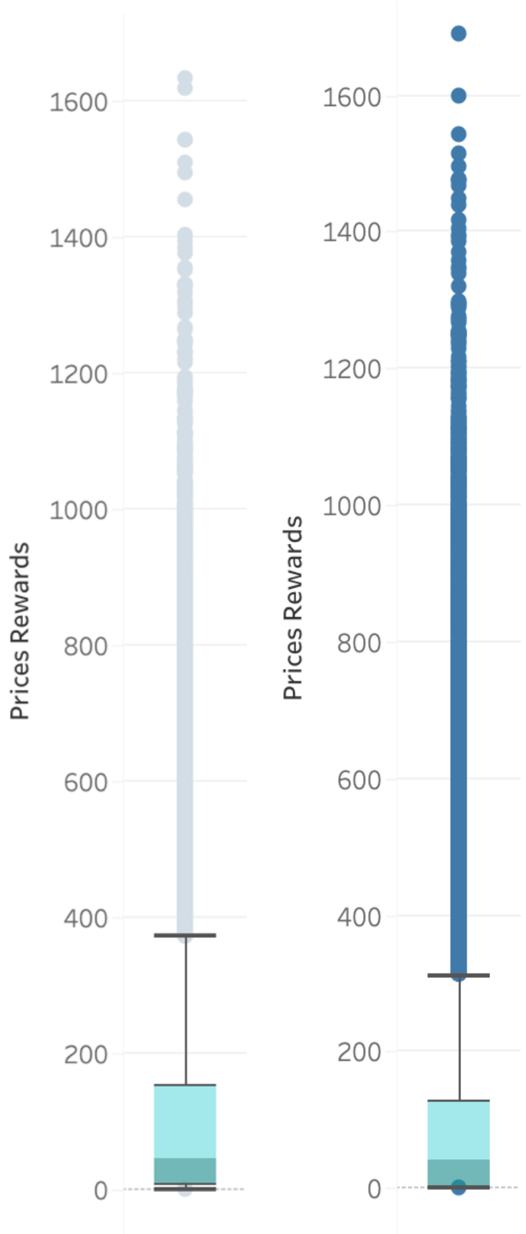


Figure 11: Box-plots Reward Prices (left unsuccessful campaigns, right successful)

tiers presented. This section explores how high-level campaign attributes such as funding goals, actual pledge amounts, campaign duration, and backer engagement interact with campaign performance. These elements reflect strategic choices made by creators regarding ambition, reach, and momentum and can offer insight into underlying behavioural or market-driven patterns.

In literature described in Chapter 2, papers presented that campaigns that have an obtainable funding goal, thus slightly lower funding goal, have a better possibility of reaching their goal (Ahmad et al., 2017; Kuppuswamy & Bayus, 2018; Mollick, 2014; Ryoba et al., 2021; M.

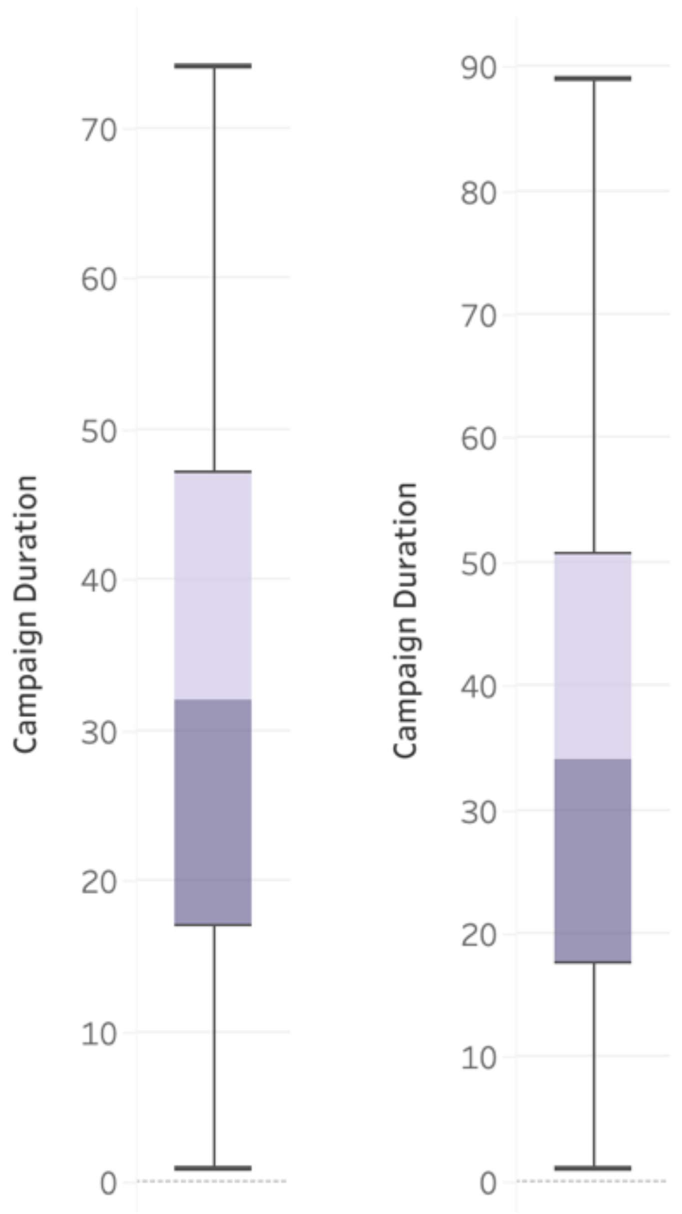


Figure 12: Box-plots campaign duration (left unsuccessful, right successful)

Zhou et al., 2016). Moreover, a long project duration can negatively impact the success of the project Fernandez-Blanco et al. (2020) and Y. Wang et al. (2021) found an inverted U-shaped relationship between timeline clarity and crowdfunding success specifically in the technology category. Last, combining textual and visual data enhances success prediction (Blanchard et al., 2022).

Figure 14 visualises the relationship between campaign goals and actual pledged amounts, with the colour indicating funding success. As expected, successful campaigns (shown in dark blue) cluster above the diagonal where pledges exceed goals. Interestingly, many of these campaigns significantly overshoot their target, suggesting strong momentum effects or social proof dynamics. On the other hand, a dense layer of failed campaigns (light blue) lies along the x-axis, indicating a substantial number of projects that fail to gain traction at all. Notably, campaigns with moderate goals – particularly between $1,000 and $50,000 – appear to have the highest density of success, reinforcing the idea that realistic goal setting may serve as a tactical advantage in driving both performance and backer engagement. One last noticeable characteristic is that almost no campaigns very close to their goal have failed. This can be seen by the small gap between the light blue and dark blue dots closer to the origin of the plot. This is in line with the psychological effect of goal proximity presented by Kuppuswamy & Bayus (2017).
Figure 12 presents the box plots of the campaign durations for successful and unsuccessful campaigns. This is contradicting with the findings of literature, as here it seems that slightly shorter campaigns seem less successful on average, but there is no large difference in campaign duration.

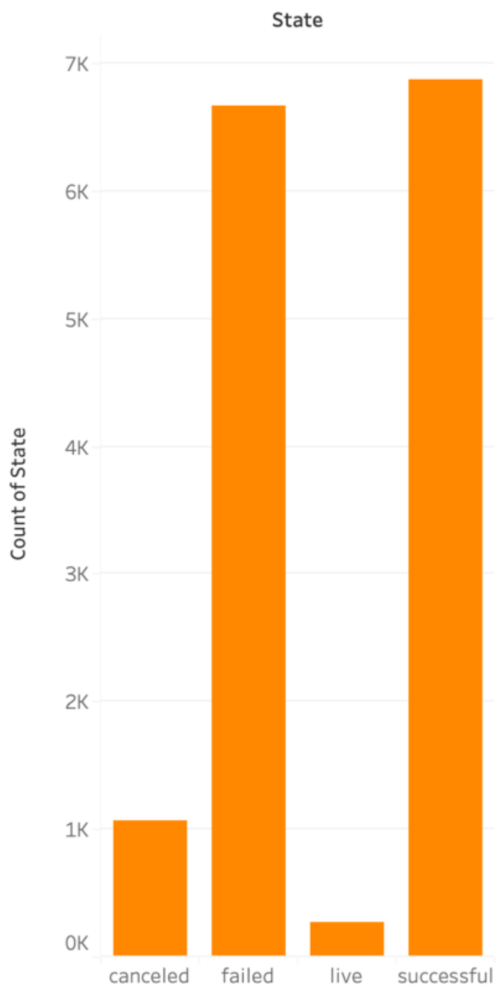


Figure 13: Status campaigns

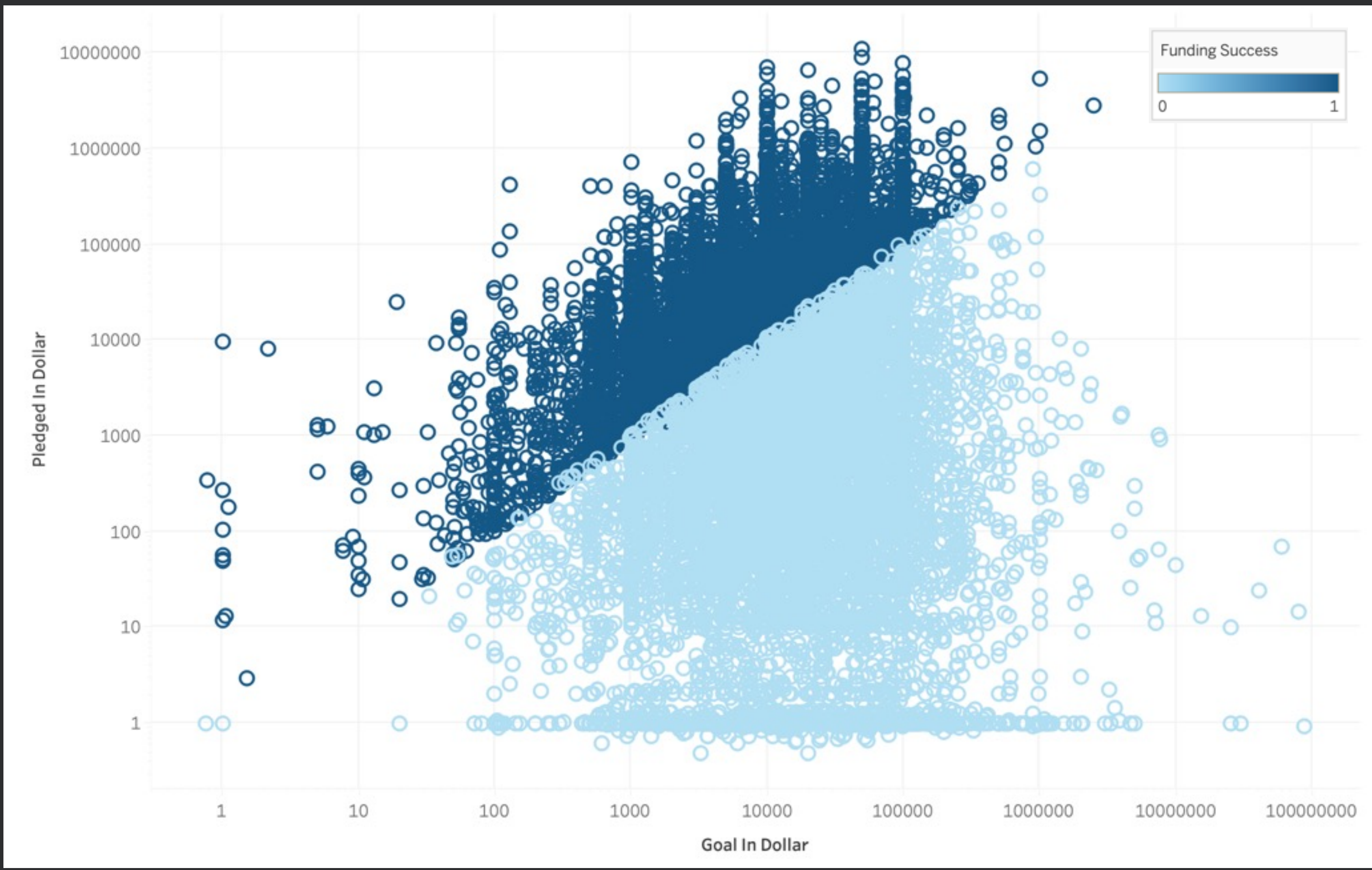


Figure 14: Pledged amount vs. Goal

Figure 15: Percent funded vs. Goal for cancelled campaigns

Furthermore, the presence of a picture and a video was checked. All campaigns in the sample contained a picture, 12026 had a video added. The average success rate of campaigns with a video was 53.71%, whereas of the 2859 campaigns without a video the average success rate was 19.10%. This confirms the findings of literature that more visual elements enhance success.

Figure 13 shows the state of all campaigns. Once zoomed in on the cancelled campaigns, one notices that the majority did not reach its goal. For few campaigns however, it would be a possibility that the demand could not be met by the supply chain, as this is a common issue in production. In Figure 15 one can observe that few campaigns, looking at the logarithmic scale, had reached a much higher pledged amount than their goal.

Correlation Analysis

With all these presented variables, we can create a correlation matrix, visualised in Figure 16. The most interesting relationships are observed in the collum or row 'funding_success'. This shows the observation that has been made earlier, a higher reward count correlates slightly with a higher success rate. Two other measures which slightly correlate with a higher campaign success, are the 'pledged in dollar' and 'backers count'. One would expect a higher correlation; however, success is defined as a binary variable and thus dampens the correlation.

Furthermore, the 'campaign duration' has a slight negative correlation with the funding success. This confirms the findings by literature. While the correlation matrix indicates a slight negative relationship between

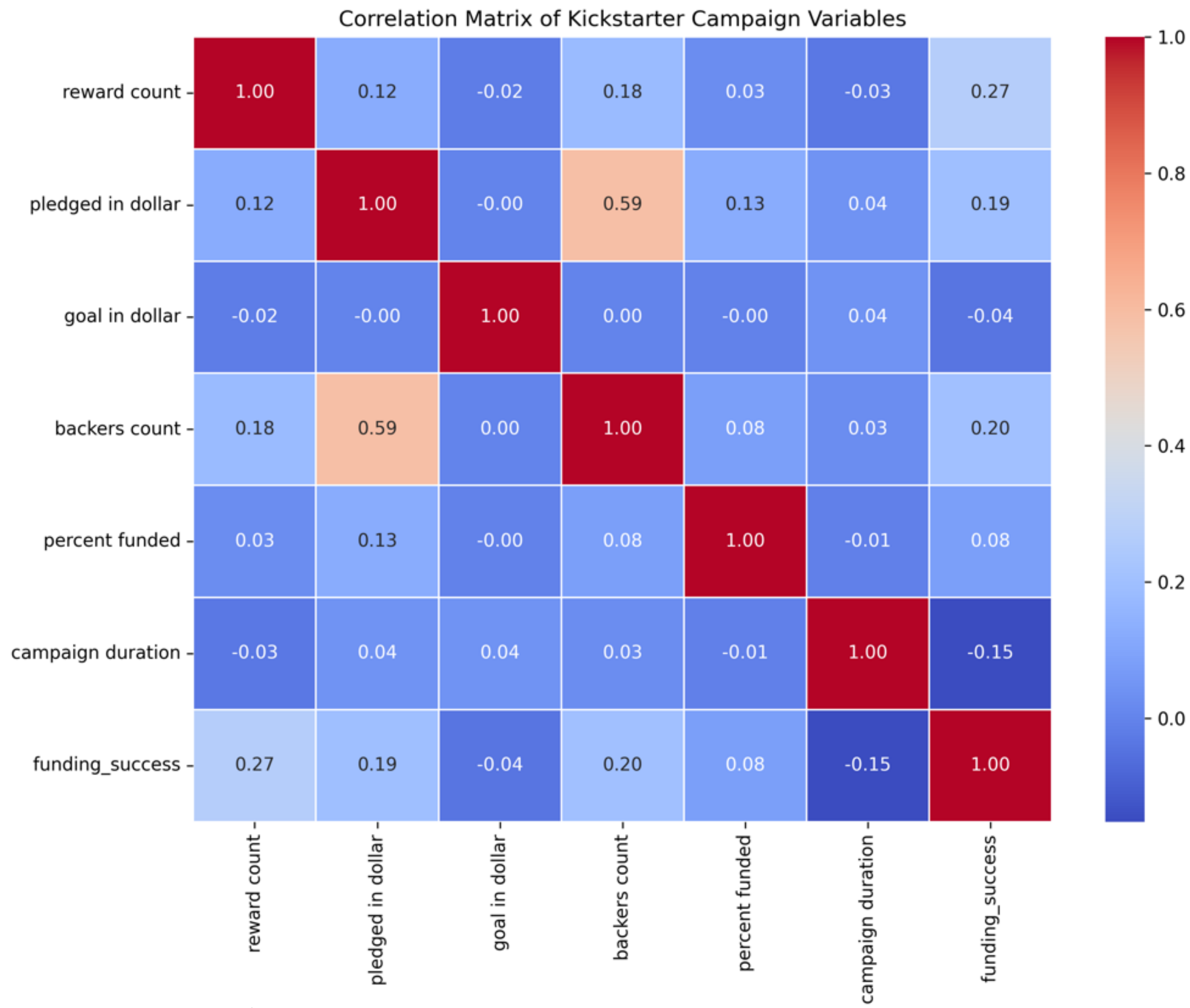


Figure 16: Correlation Matrix of the dataset

campaign duration and success, this is not immediately apparent in the box plot comparison, where successful campaigns have roughly the same median duration as unsuccessful campaigns. This discrepancy is explained by the presence of long-duration outliers among failed campaigns, which pull down the correlation despite having little influence on the median. Furthermore, when grouping campaigns into short, medium, and long duration bins, the data reveals that short campaigns (under 20 days) have the highest success rate (~52.5%), while long campaigns (over 40 days) show a significant drop in success (~36.3%). This nonlinear trend highlights the nuanced relationship between duration and performance: while moderate-length campaigns may perform well, extremely long campaigns are generally detrimental, likely due to backer fatigue or perceived lack of urgency. To confirm this, a violin plot was generated. This is shown in Figure 17, where the unsuccessful campaign duration plot is more top-heavy. Furthermore, 'percent funded' is weakly correlated with everything. Since this is a derived metric (pledged/goal), it is not unexpected. Contradicting to the earlier mentioned literature, 'goal in dollar' does not seem to have any clear relation to any other metric.

Textual Characteristics of Campaigns

In addition to numerical and structural attributes, the textual elements of a Kickstarter campaign—such as its blurb, slug, and reward titles—can play a significant role in influencing its perceived appeal, trustworthiness, and eventual success. This section explores the language patterns associated with both successful and unsuccessful campaigns, as well as the naming conventions used in the most popular reward tiers.

Figure 18 presents two word clouds that visualize the most frequently occurring terms in the blurbs and slugs of successful (left) and unsuccessful (right) campaigns. Clear semantic themes emerge from this comparison. Successful campaigns are characterised by language focused on design, technology, and functionality—frequent words include "design," "smart," "camera," "easy," and "device." This suggests an emphasis on product completeness, usability, and innovation. The appearance of words like "build" and "create" may also indicate the appeal of maker-culture and tangible outcomes, which could resonate with Kickstarter's tech-savvy backer base. In contrast, unsuccessful campaigns show greater usage of more abstract or platform-focused terms such as "platform," "project," and "help." While the word

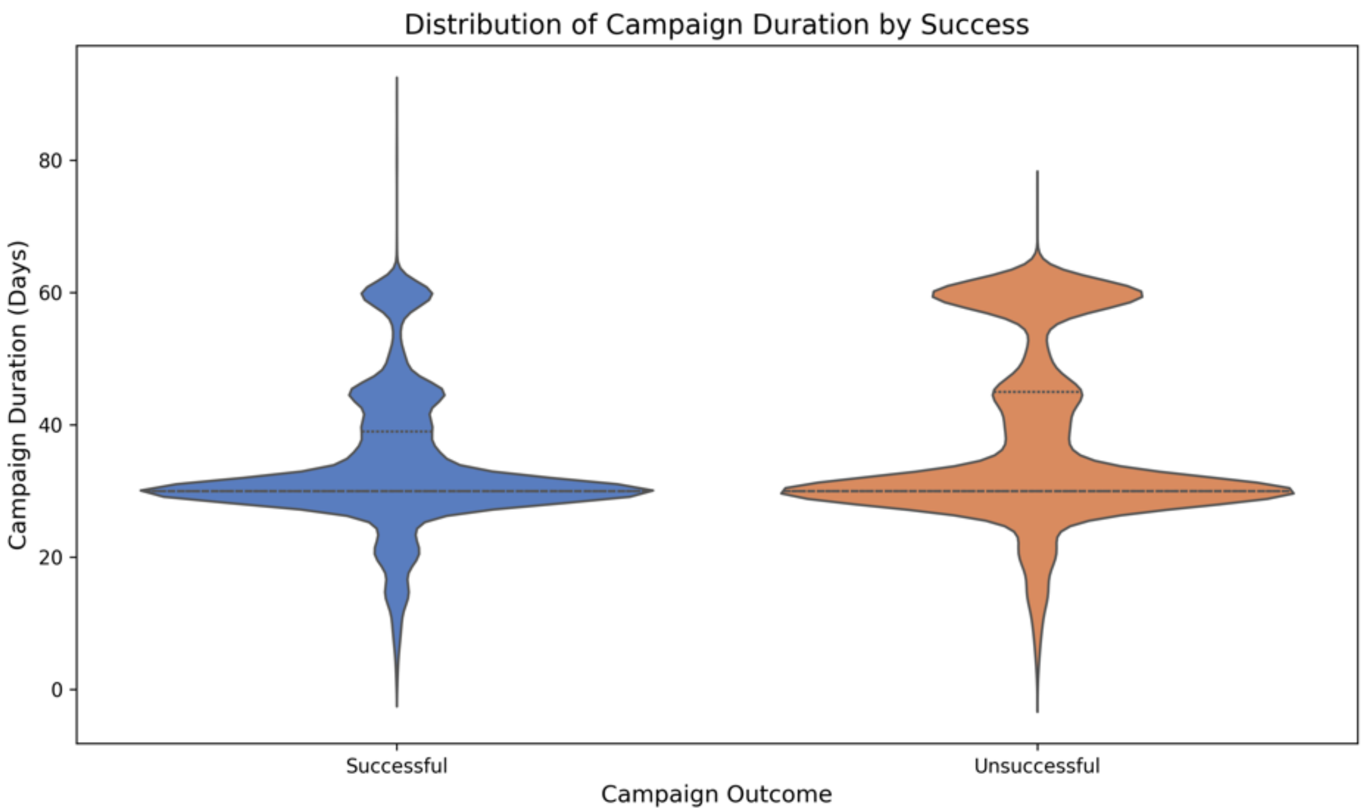


Figure 17: Violin Plots Campaign Duration

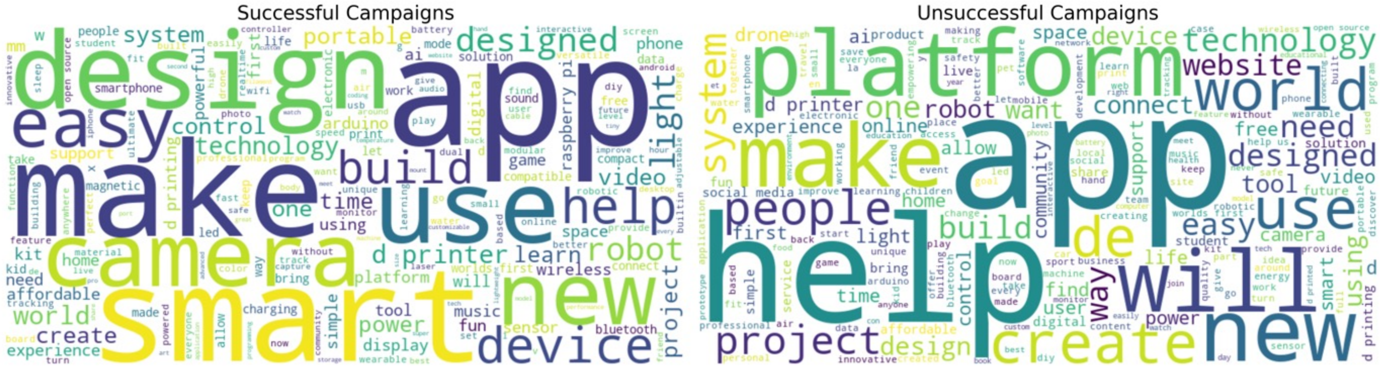


Figure 18: Word clouds of campaign texts

"app" appears prominently in both, the context in which it is framed seems to differ. Unsuccessful campaigns may rely more heavily on aspirational or conceptual language (e.g., "change the world," "connect people") without clearly communicating tangible benefits.

The prominence of words like "will" and "want" may reflect future promises rather than present deliverables. Successful campaigns appear to use clear, concrete, and product-focused language, consistent with prior research emphasising the importance of transparency, clarity, and emotional resonance in campaign narratives (Adamska-Mieruszewska et al., 2021; K. F. Yang et al., 2023; Babayoff & Shehory, 2022). This aligns with findings that a positive, concrete, and persuasive communication style fosters backer trust and increases contribution likelihood ((Q. Li & Wang, 2024; Parhankangas & Renko, 2017).

Beyond campaign summaries, the titles of reward tiers also reveal patterns worth exploring. Figure 19 shows the 20 most backed reward titles across all campaigns. Similarly, the dominance of "early bird" and "special offer" labels among the most backed rewards supports prior evidence that personalization, exclusivity, and scarcity enhance campaign appeal (Butticè et al., 2018; Wessel et al., 2019). These strategies tap into psychological triggers like urgency and perceived value, reinforcing behavioural phenomena such as the phantom effect and scarcity-driven purchases (M. Y. Chen et al., 2021; Y. Lin et al., 2016). Together, these findings emphasise that how a campaign communicates, both semantically and structurally, can be just as critical as what it offers.

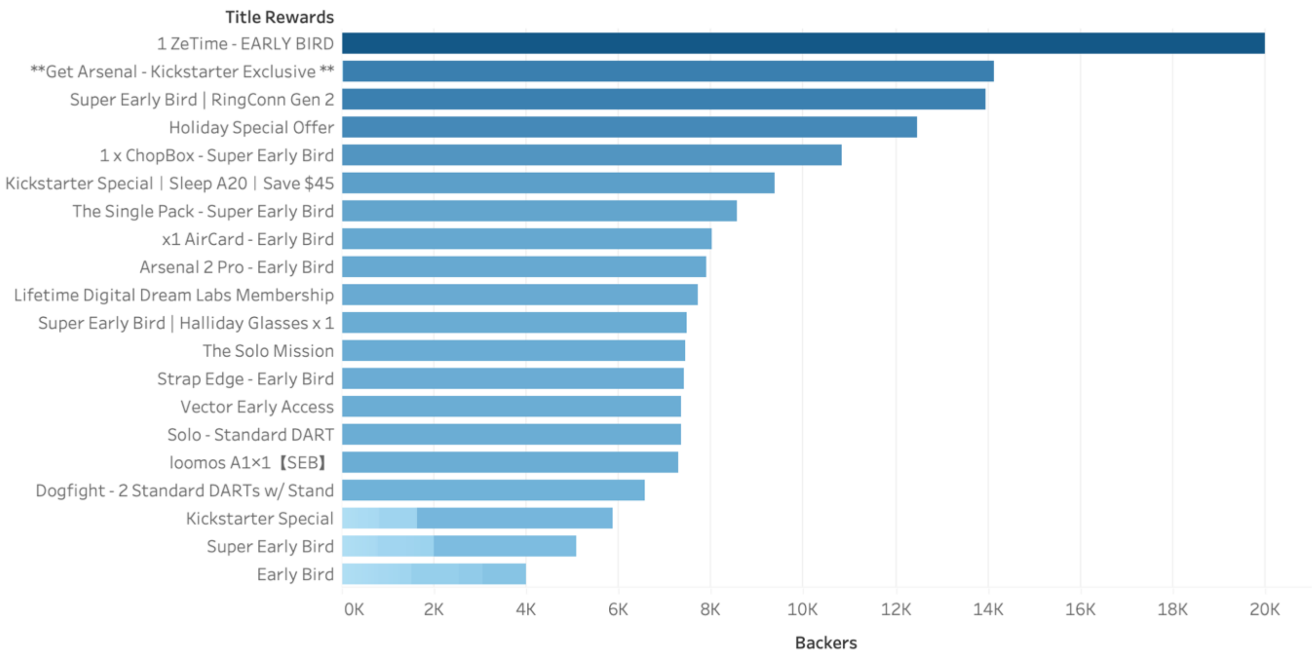


Figure 19: Most successful reward titles

4.4 Data Preparation

To make sure the predictive modelling and optimisation are effective, the dataset undergoes a structured preprocessing pipeline to be able to use it for feature selection and prepare for modelling. This includes selecting informative features, labelling emotional tone, vectorising textual content, splitting the dataset for evaluation, and formatting inputs for efficient hyperparameter search.

Feature Selection

The selected features are divided into three primary groups: reward-tier characteristics, campaign metadata, and textual content. These features capture both quantitative and qualitative aspects of the crowdfunding projects, allowing the models to learn from structural and semantic signals alike.

Reward tier features:

o Reward count: A measure of variety; too many options risk overchoice (Elitzur et al., 2024), and the optimal number of rewards seems to be around 10 (Cai et al., 2021).

o Average reward price: Indicates perceived value and price positioning.

o Early bird presence: Known to drive early engagement (Wessel et al., 2019).

o Sentiment of reward titles: Reflects persuasive or emotional tone (Yosipof et al., 2024).

o Top reward name category: General naming patterns like "Bundle" or "Early Bird" show the strategy used in the reward tiers.

o Exclusivity/limited: Scarcity can create urgency (Y. Lin et al., 2016).

Campaign metadata:

o Goal in dollar: Higher goals reduce success odds unless justified by product value (Ahmad et al., 2017; Kuppuswamy & Bayus, 2018; Mollick, 2014; Ryoba et al., 2021; M. Zhou et al., 2016).

o Campaign duration: As shown, this is tied to urgency and momentum, data reveals that short campaigns (under 20 days) have the highest success rate on average, coherent with literature (Fernandez-Blanco et al., 2020).

o Subcategory: Different categories might behave differently, as shown in our dataset the success percentages differ quite a lot between the subcategories.

o Country: The same principle holds for the different countries, and cultural preferences can have a large effect on certain preferences.

Textual Content:

o Blurb and slug: The campaign's summary pitch.

o Reward titles: Often used to communicate urgency or exclusivity.

o Sentiment label: Indicates tone and engagement potential, generated from the blurb and slug.

o Embeddings (LLaMA or sentence-transformers/all-MiniLM-L6-v2): Represent tone, clarity, and abstract themes. The sentence-transformers/all-MiniLM-L6-v2 are extracted from the blurb and reward titles, and the LLaMA embeddings are exracted from the campaign title, description, goal, duration and reward tiers (saved as prompts).

Excluded from this setup are the full reward descriptions and shipping details due to incompleteness, heterogeneity, and incompatibility with structured optimisation.

Sentiment and Emotion Labelling

To extract emotional tone from campaign texts and reward texts, sentiment analysis is applied to the blurb and slug of each campaign. A transformer-based model from Hugging Face, cardiffnlp/twitter-roberta-base-sentiment, is used to classify each text as positive, neutral, or negative. This model is chosen because it is fine-tuned on large-scale, real-world sentiment datasets derived from social media which closely resembles the tone and brevity of Kickstarter campaign texts. The unified API allows seamless integration of pre-trained models, while the Model Hub facilitates access to task-specific fine-tuned variants such as the one used here. Compared to rule-based sentiment tools like VADER, which are often used as baselines in sentiment analysis but underperform on nuanced or ambiguous expressions, transformer-based models offer superior generalization and deeper contextual understanding

(Medhat et al., 2014), as mentioned in Chapter 3. Before classification, the text inputs are pre-processed through cleaning and tokenisation, which involves converting to lowercase, removing punctuation or noise, and breaking the text into tokens or sub-word units that the model can process (Daniel & Martin, 2024).

Vectorisation of Text Data for Baseline model

To enable ML models to interpret campaign text, the raw language must be converted into a numerical format. This process, known as text vectorisation, transforms words or sentences into embeddings. These embeddings are essential for feeding textual data into ML models as such models cannot operate directly on raw text. For the baseline model we use different embeddings than the LLaMA Embeddings to compare the results. To incorporate abstract linguistic features into the RF model, only select textual fields were transformed into dense vector representations using sentence embeddings. Rather than embedding the entire dataset, two fields were specifically chosen:

- Blurb and slug: These typically reflect tone, urgency, and the persuasion strategy.
- Reward titles: Aggregated across reward tiers, these titles often include promotional phrases, or limited time offers such as "Early Bird", which can signal pricing tactics and appeal.

For this thesis, a transformer-based embedding approach is used, as it offers superior performance in capturing contextual meaning compared to traditional methods like TF-IDF or Word2Vec. Specifically, the sentence-transformers/all-MiniLM-L6-v2 model is chosen for its balance of speed, accuracy, and compactness. This model is part of the Sentence-Transformers framework and is based on the MiniLM architecture—a distilled variant of larger transformer models like BERT—optimised for sentence-level semantic understanding. It generates dense, low-dimensional embeddings (384 dimensions) that encode not just the presence of words, but also their contextual relationships and underlying intent. The preprocessing pipeline consists of cleaning the text (blurb and slug), converting it to lowercase, and applying tokenisation using the sentence-transformers library. The resulting vectors are saved in .npy format or appended to the dataset as additional columns. These embeddings only serve as input for the baseline RF model. This step is critical, as it allows the model to leverage semantic patterns embedded in the text—such as emotional tone, clarity, and thematic framing—which are known to influence crowdfunding success. Without this transformation, the model would be unable to recognise or learn from linguistic patterns in campaign language.

For the baseline Random Forest (RF) model, the dataset was split using a stratified hold-out strategy to maintain the original class distribution of successful and unsuccessful campaigns across all subsets. This method ensures that the model is exposed to a representative sample of the campaign outcomes during training, validation, and testing, preventing any imbalance from skewing results. The splitting process was conducted in two steps, where initially, 90% of the data was reserved for training and validation, while 10% was set aside for final testing. Within the 90% subset, the data was further divided into 70% for training and 20% for validation. The final distribution of the data was as follows: Training set: 63% of the total dataset; Validation set: 18% of the total dataset; Test set: 10% of the total dataset. For BOHB there is a different set-up, described in the next Chapter.

4.5 Concluding words Chapter 4

This chapter has established the foundation for optimising reward-tier configurations through the collection, cleaning, and structuring of a rich Kickstarter dataset. Combining Web Robots IO campaign data with custom webscraping for reward tiers resulted in a uniquely granular dataset containing both campaign-level and reward-level features.

Key preprocessing steps such as deduplication, monetary conversion, and format standardisation ensure data integrity, while exploratory analysis confirmed alignment with known crowdfunding patterns according to literature. The data was first suited to make a first RF baseline model, which is the first model to predict the success of the campaigns given the retrieved data. The baseline model uses embeddings often used in short social media descriptions, suitable for this context, to compare it later on with the LLaMA Embeddings that are generated in the next Chapter as well.

The resulting dataset is fully prepared for use in the RF model. It includes numerical, categorical, and vectorised textual features. This structured and semantically enriched dataset forms the empirical core for the optimisation experiments presented in the following chapters.

5. Model Development

To support this modelling architecture, three targeted research questions were explored in the preceding theoretical section in Chapter 3:

- RQ2.1 demonstrated BOHB's adaptive resource allocation and probabilistic modelling, suitable in a high-dimensional design space while minimising the risk of overfitting.
- RQ2.2 outlined the unique role of textual features—such as clarity, sentiment, and tone—in shaping campaign outcomes, captured in LLaMA embeddings and sentiment scores.
- RQ2.3 clarified how the interaction between reward design and campaign narrative can be jointly modelled using an architecture that supports both dense textual embeddings and hyperparameter search, enabling strategic optimisation of campaign configuration.

This chapter builds on the theoretical foundations presented in Chapter 3 and continues with the implementation. This chapter shows the overall modelling strategy and outlines the rationale behind the decisions while developing the model. The proposed modelling approach consists of four steps:

- A Random Forest (RF) model, that acts as a baseline model, to test a simple prediction of success of campaigns based on campaign and reward attributes.
- Three different ML learning models (RF, MLP, XGBoost) with the LLaMA embeddings, to see if there is a better performance of the success prediction of the model and to see which ML model is best suited to be used in the following steps, of which the ML model is tuned by BOHB.
- A model that uses the preferred ML method (XGBoost) together with BOHB to predict success in the general case (no input of campaign variables)
- The final reward-tier optimiser model using BOHB, aimed at improving reward-tier design by tuning pricing, quantity, and tier structure to increase funding success and minimise financial risk, by giving the campaign category, country, duration, and funding goal.

The model development and training processes for this thesis were conducted using Python 3.8.10 on the Jupyter Notebook server provided by the University of Twente. These servers run on Dell PowerEdge R750 machines, equipped with 72 Cores / 144 Threads, 256 GB Memory, and 2 x Nvidia A16 GPUs, ensuring sufficient computational power. The model in this thesis emerges as a novel contribution in response to the lack of structured reward-tier optimisation methods found in current research. While past studies have either focused on predictive models for campaign success (Ahmad et al., 2017; Elitzur et al., 2023; Ryoba et al., 2021; M. Zhou et al., 2016) or very specific subjects such as general pricing heuristics for rewards (J. Chan et al., 2023; Gong et al., 2021; Keisar & Lev, 2023; Simons et al., 2017), very few have explored a method that both predicts and actively optimises using textual and structured inputs.

5.1 Random Forest Baseline Model

To establish a performance benchmark and validate the quality of the prepared dataset, a baseline model is constructed prior to implementing more complex model structures. This serves two purposes: first, to assess the quality of the dataset and compare the performance to literature; and second, to provide a reference point for comparing the performance of more advanced models later in this thesis. The RF classifier is selected as the baseline model due to its robust performance across a wide range of supervised learning problems. RF offer high interpretability, handle both categorical and numerical features well, and are resilient to overfitting due to their ensemble nature. Moreover, prior research in crowdfunding prediction has shown their strong predictive performance, especially when working with tabular data and mixed feature types (Haitham et al., 2024b; Zhong, 2022a). While the Random Forest model is not used as the final predictive tool, it plays a key role in validating preprocessing choices, understanding feature behaviour, and establishing a reliable baseline for future model comparison. They do lack the ability to understand semantic and contextual relationships in text. Therefore, in this model, the sentence-transformers/all-MiniLM-L6-v2 embeddings are used, to later compare the results with the LLaMA embeddings.

Performance Metrics

The model is trained on the training set and evaluated on both the validation and test sets. Additionally, a 5-fold stratified cross-validation is performed to assess model variance. The evaluation results are summarized in Table 5. These results show that the model is consistent across different data splits, with limited variance, suggesting reliable generalization to unseen data. The high precision indicates that when the model predicts a successful campaign, it is likely correct, which is an important property for applications where false positives, such as investing in a failing campaign, are costly. Given the financial consequences of misclassification, F1 score, and precision are preferred metrics over accuracy alone.

To get more information about the model's capabilities, two images are shown in Figure 23 and Figure 24, respectively a Confusion Matrix and a Feature Importance Plot. The confusion matrix below provides a clear view of the model's classification performance on the validation set. It shows that the model correctly classified a substantial number of both successful (979) and unsuccessful (1297) campaigns, while it incorrectly classified 279 unsuccessful and 422 successful ones. In binary classification often false positives are more common. This conservative bias may be justifiable in a crowdfunding context, where overestimating potential success could mislead creators. A Feature Importance Plot ranks the input features according to their influence on model predictions, helping to identify which reward-tier or campaign-level attributes are most impactful. Feature importance is derived using Gini impurity, highlighting which features the model relied on most frequently to make splits and decisions. Unsurprisingly, the funding goal is the most influential feature, reinforcing findings from Mollick (2014) and Kuppuswamy & Bayus (2018) that campaigns with overly ambitious goals are less likely to succeed. The early_bird_flag is also important, aligning with the idea that urgency mechanisms help generate early momentum Wessel et al. (2019). Other high-impact variables include reward count, confirming the notion of choice overload (Elitzur et al., 2024), and campaign category, showing the different performance of categories such as Gadgets and Social Good. Additionally, campaign age contributed moderate predictive power, likely capturing shifts in

	Validation Set	**Test Set**	**5-Fold Cross-Validation (Mean ± Std Dev)**
Accuracy	76.45%	76.96%	76.96% ± 0.94%
F1 Score	73.64%	73.39%	73.60% ± 1.25%
Precision	77.82%	80.44%	79.83% ± 1.12%
Recall	69.88%	67.48%	68.28% ± 1.80%

Table 6: Performance Metrics Baseline Model

platform dynamics or best practices over time. Finally, the reward_name_cluster feature, which groups reward tier naming strategies (e.g., "starter," "bundle"), showed that linguistic packaging plays a role in model decisions, which supports the broader claim that language framing influences perceived value and campaign appeal Yosipof et al. (2024). The emb_reward_reward_title_x is an embedding, thus a fixed-length numerical vector that preserves a certain semantic meaning and linguistic structure. The exact meaning of these embeddings is unknown but can be approached. As an example, the meaning of emb_reward_reward_title_242 is approached later this subchapter.

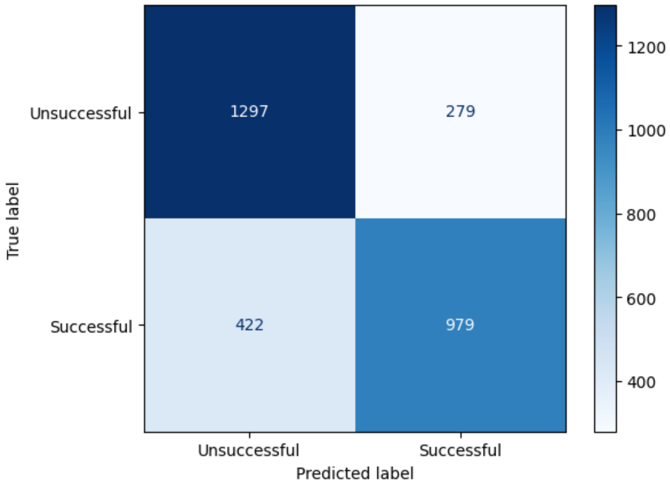


Figure 20: Confusion Matrix baseline model

Embedding-Based Semantics and SHAP Dependence
SHAP (SHapley Additive exPlanations) is applied to get a deeper insight in Feature Importance. SHAP assigns each feature a marginal contribution to the model's output on a per-sample basis, making it particularly suited for uncovering how specific features affect individual predictions (Joseph et al., 2024; Kavzoglu & Teke, 2022). This method provides local interpretability, complementing traditional global feature importance metrics. Interesting enough one can see that the SHAP-based ranking of influential features differed from the feature importances derived from Gini impurity, the default method used by Random Forests. Gini impurity reflects how frequently a feature is used to split data across all trees in the ensemble, providing a measure of aggregate importance. SHAP values assess the actual contribution of each feature to prediction outcomes. This makes SHAP more sensitive to subtle but consistent patterns, particularly useful when working with high-dimensional or correlated features such as embeddings. The SHAP results brought renewed attention to several features. Reward count, previously seen as moderately important, emerged as one of the top predictors under SHAP. This suggests that the model does not only reacts to how often the feature is used in decision trees, but how its specific values influence prediction outcomes. Similarly, category_main_Gadgets showed strong predictive influence. This aligns with earlier findings in

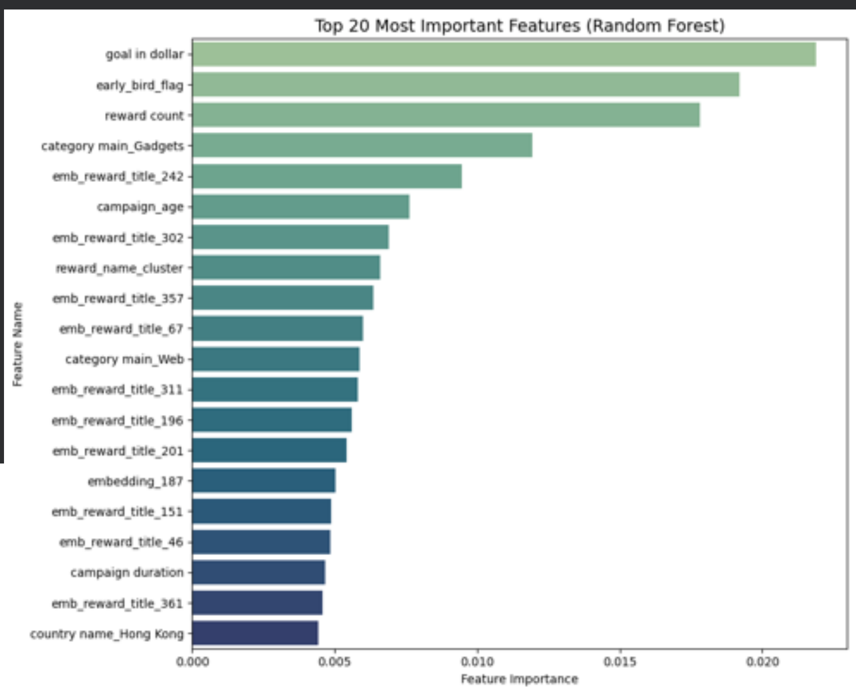


Figure 21: Feature importances baseline model

the dataset that certain campaign domains, particularly those grounded in tangible technology products, are more likely to succeed. One interesting finding is that emb_reward_title_242, one dimension of the dense sentence embeddings derived from reward tier titles, is highly correlated with the success of a campaign, which can be seen in Figure 21. If we zoom in on what this embedding is correlated with in Figure 23, we find that campaigns with high values in emb_reward_title_242 frequently featured reward titles with detailed specifications, technical terminology, and references to electronic components or sensors. In contrast, campaigns with low values tend to be more abstract or service-oriented, often framed around social initiatives or local community engagement. Some of the latter were written in Spanish, suggesting that this embedding may also encode linguistic or cultural markers.

Implications for Modelling and Theory

The analysis of embedding dimensions alongside SHAP values offers critical insight into the model's ability to leverage textual features in a meaningful way, and its importance. This finding supports the growing body of research suggesting that transformer-based language models can extract actionable features from marketing and persuasive text, enabling more accurate outcome prediction (Reimers & Gurevych, 2019). Furthermore, the interpretability of emb_reward_title_242 indicates that learned embedding dimensions, although abstract, can correspond to recognisable communicative strategies, such as technical specificity versus emotional appeal.

🔝 Top 15 campaigns with **high emb_reward_title_242** values:

	emb_reward_title_242	text_for_sentiment
12242	0.174084	Raspberry Pi emulator: WiFi, ETH, I2C, SPI, 2 ...
2750	0.154326	4-20mA & 0-10V In/Out, Opto-Isolated In/Open D...
9991	0.150496	1080P 60FPS/ 2K 30FPS, Sony STARVIS, All Water...
389	0.147973	The Amazing Super Stove is a game changer. Fre...
9702	0.142463	The non-contact smart thermometer that brings ...
4454	0.140853	Learn to control automation interfaces using N...
5604	0.138312	F1.0 aperture | 1/1.8" SONY STARVIS Sensor | B...
8234	0.137730	25 Cylinders|Car Restarter|Inflate More Tires ...
12746	0.137609	The safe and easy-to-use professional-grade DI...
13072	0.137040	Dive deep with Nemo down to 100 meters (328ft)...
14625	0.134731	AI Designs|AI Eyes|20W Diode+2W IR|300*300mm|3...
218	0.133025	The Ultimate Online Private Space & GPS Camera...
8231	0.132044	Satisfy your pet's curiosity with an interacti...
12719	0.131290	All-in-one hydration tool that auto-tracks wat...
10920	0.131152	Hyperdolly is the ultimate solution for accura...

Figure 23: Campaign text with high emb_reard_title_242 values

🔻 Bottom 15 campaigns with **low emb_reward_title_242** values:

	emb_reward_title_242	text_for_sentiment
11566	-0.164206	Udderly is a platform for moms to buy or sell ...
10228	-0.153358	The restaurant industry is always hiring! Our...
2942	-0.149342	Plataforma de gestión educativa, especializada...
3739	-0.146259	Un robot que distribuya contedores a puertos d...
1022	-0.144323	Are Wisconsin elected officials listening to W...
3010	-0.144252	Prototype: www.framespot.com\n\n#SaveYourInter...
4138	-0.143903	App para publicar y buscar artículos o servici...
10197	-0.142935	A website enabling runners to list, book and h...
11889	-0.140020	The tool you did not know you needed multitool...
11775	-0.135204	Arcade-Pi is a mini Arcade game machine which ...
12936	-0.135039	Desarrollando una plataforma para agilizar y a...
8084	-0.131772	The Student Dashboard: Your All-In-One Solutio...
11583	-0.130528	GoodNews allows you to get and share uplifting...
8631	-0.129745	Maple is a project to make use of Junkyard & s...
10229	-0.129540	Accurate Information About Florida Law Firms\r...

Figure 24: Campaign text with low emb_reard_title_242 values

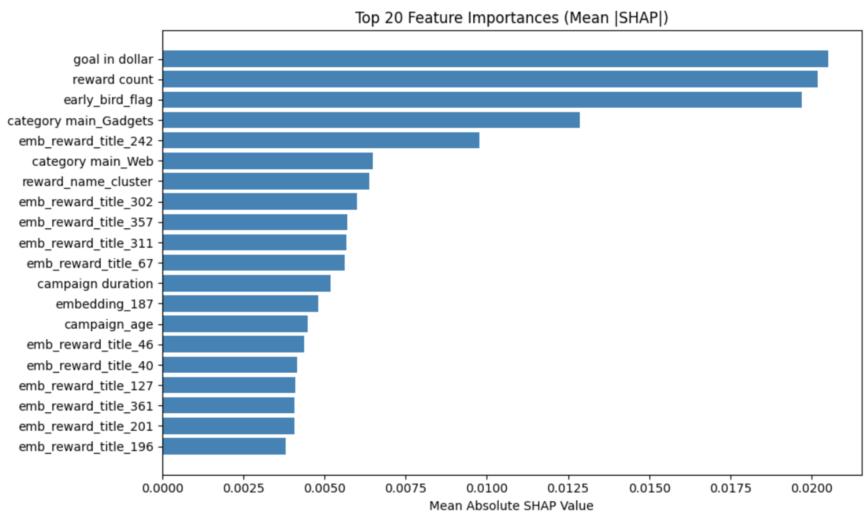


Figure 22: SHAP Feature Importances Baseline Model

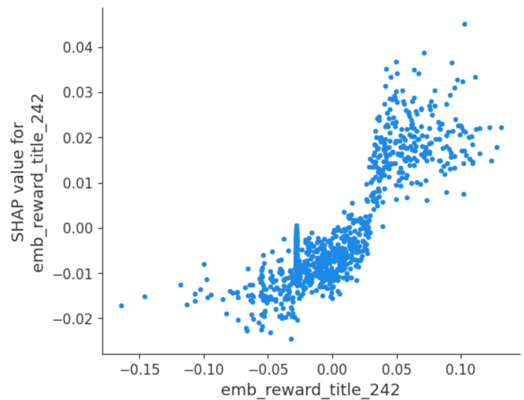


Figure 25: SHAP correlation values emb_reward_title_242

5.2 LLaMA 3.2 Embeddings & Model Comparison

Crowdfunding campaigns are heavily text-driven, where the project descriptions, reward tiers, campaign goals, and timeframes contain important indicators of a campaign its success, as those elements are central in conveying a campaign's value proposition, urgency, and credibility to potential backers. Using only traditional numerical features would not capture the nuanced signals within textual content, such as emotional appeal, clarity of description, or perceived value of rewards. To address this, we employ LLaMA 3.2 - 3B Instruct, a state-of-the-art large language model, to generate high-dimensional embeddings. LLaMA is chosen for the reasons mentioned in Chapter 3: it captures long-term dependencies in textual data, preserves semantic information about reward structures, campaign goals, and project descriptions, and LLaMA's architecture enables it to process varied textual inputs-campaign narratives, reward titles, and descriptions-into a unified vector space. The embeddings are generated using Hugging Face's Transformers library, which facilitates seamless integration with PyTorch. The generation process consists out of merging campaign data with the reward information based on the project slug, where the campaign title, description, goal, duration and reward tiers are saved as prompts. These prompts are used to create the final embeddings.

Data Preparation and Model Training

The generated embeddings are combined with campaign metadata and funding success labels (0 for unsuccessful, 1 for successful). With this structured dataset, we can use different ML models to predict the success of the campaign. The four chosen models are:

- Logistic Regression (LR): One of the most basic models chosen as a benchmark.
- Random Forest (RF): Chosen for its interpretability and ability to handle imbalanced classes.
- XGBoost: Preferred for its gradient boosting mechanism, allowing better handling of structured data and complex decision boundaries.
- Multilayer Perceptron (MLP): Used for capturing non-linear relationships due to its deep learning architecture.

The models are trained and validated using an 80/10/10 split; 80% for model learning, 10% as validation set for hyperparameter adjustments, and 10% as a test set to evaluate its performance. The model training is performed with scikit-learn for LR, RF, and MLP, and XGBoost library for the XGBoost model.

Model Evaluation and Comparative Analysis

A more extended evaluation is presented in the next Chapter, but since one model had to be chosen to continue the modelling process with, the models are already compared on performance. Each model is evaluated based on Accuracy, Precision, Recall, and the F1-Score. A summary of the results is provided in Table x. XGBoost and MLP had both a good performance. Figure 29 shows how they compare on all metrics. There is chosen to optimise the MLP and XGBoost model, since they both generally perform better than the RF model. One can also notice the least complex model, LR, already provides a good performance. This is similar with the study by M. J. Zhou et al. (2015) mentioned in Chapter 2, which obtained accuracy of 73%. Here an accuracy of more than 74% is achieved, most probably since more

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	0.742780	0.716927	0.732558	0.724659	0.815020
Random Forest	0.746138	0.749196	0.677326	0.711450	0.815057
XGBoost	0.763600	0.745614	0.741279	0.743440	0.847130
MLP	0.751511	0.703325	0.799419	0.748299	0.832128

Table 7: Performance Metrics LR, RF, XGBoost, MLP

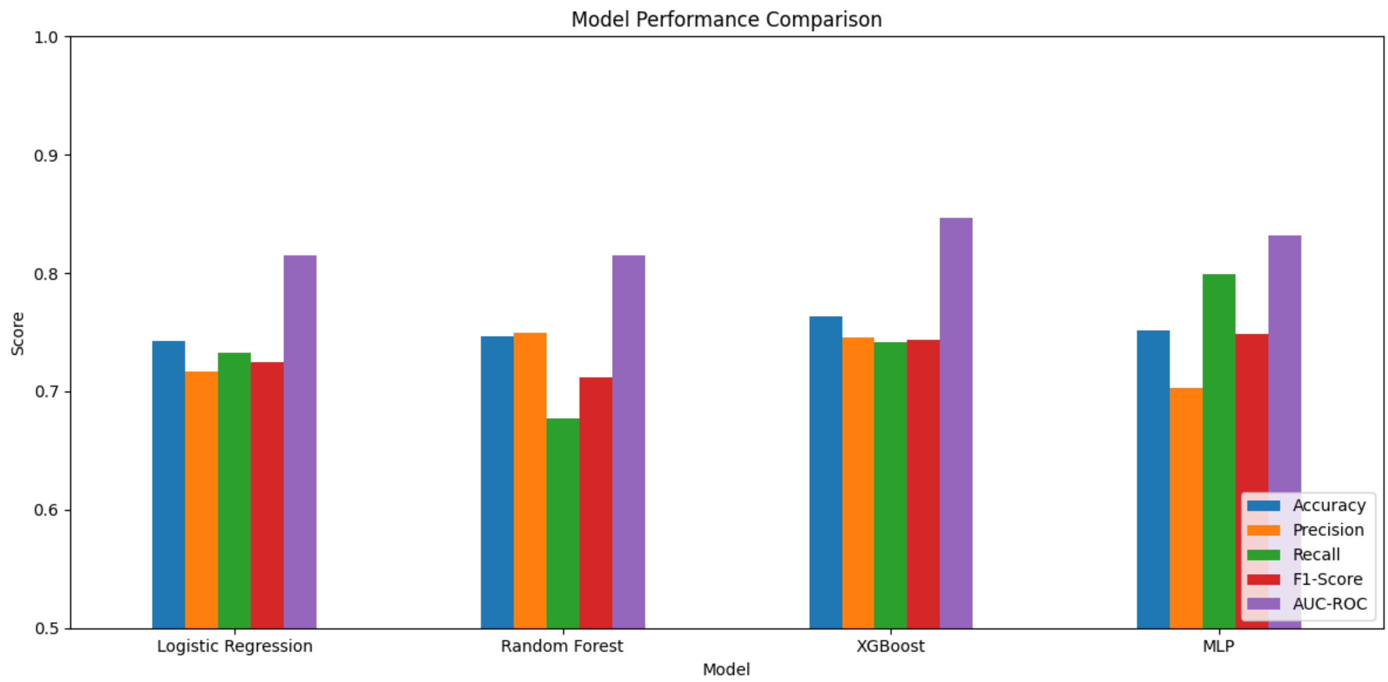


Figure 26: Model Performances Visualised

information, such as the reward information and the embeddings, are used. MLP scored the highest on Recall, or recognising the true positives, but the precision was lower than the other models. The confusion matrices are shown in Figure 28, further showing the performances of the models, made with matplotlib and seaborn.

BOHB Hyperparameter Optimization

We proceed with BOHB to fine-tune the hyperparameters of the models. We explore the hyperparameter space for learning rate, max depth, and number of estimators (XGBoost), as well as hidden layer sizes and activation functions (MLP). We also minimise the cross-validation error during training using ConfigSpace and hpbandster for automated configuration sampling. The outcome of the BOHB optimisation showed an increased performance of the XGBoost and MLP models. We again generate Confusion Matrices (Figure 27) to compare the results.

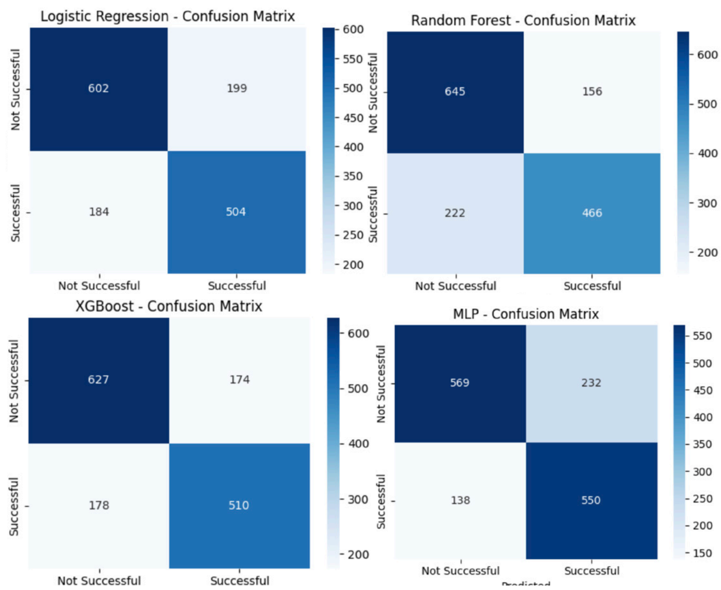


Figure 28: Confusion Matrices of LR, RF, MLP, XGBoost

Mainly the recall and F1-score are improved.

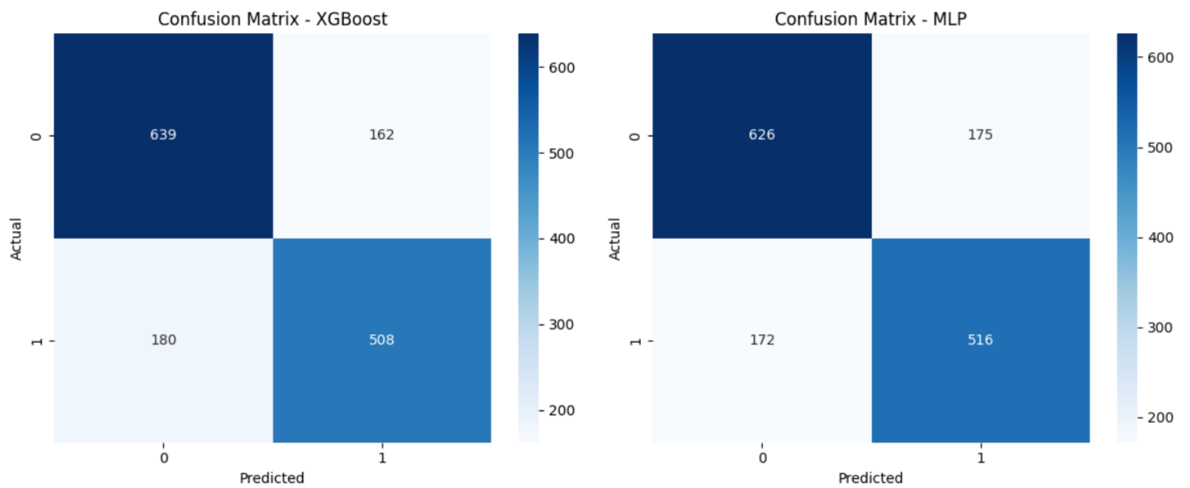


Figure 27: Confusion Matrices of optimised XGBoost and MLP

5.3 Generalized Success Prediction with XGBoost and BOHB

The primary goal of our analysis is to optimise reward strategies and use the success metric to see which reward parameters are most effective. We first simulate the general case, where we do not enter any campaign-specific information. We use the dataset that has standardised information about the Kickstarter campaigns, including funding targets, campaign duration, number of backers, category, and country. Additionally, we use the cleaned reward information with information about reward attributes, such as average reward price, exclusivity, early bird flag, and reward clustering. Moreover, we keep using the LLaMA Embeddings that we generated in the previous sub-chapter. Last, we use the one-hot encodings for campaign categories and countries, ensuring that categorical data is numerically represented. The final dataset contains 14,885 campaigns and 3,119 features. This consolidated dataset served as the input for the modelling and the optimisation process.

Modelling Approach

To predict campaign success, we used the XGBoost Classifier with the optimised settings by BOHB in the previous sub-chapter, combined with Optuna for hyperparameter optimization. The objective is to maximize the F1 score, focusing on optimising predictive accuracy across both successful and unsuccessful campaigns, given their near-even distribution. Despite the balanced classes, F1 score remains the preferred metric as it balances both precision and recall, ensuring that the model effectively captures both successful and unsuccessful campaigns without bias towards one class. The model included the following reward-based features: reward count, average reward price, presence of early bird rewards, presence of exclusive rewards, and the reward name cluster. Furthermore, the model included the following campaign attributes: campaign duration, and funding target. The LLaMA embeddings are also included, and lastly, the categorical encodings: columns for category and country. Optuna optimised all the former mentioned reward-based features, plus the campaign-based features. The model is trained using an 80-20 train-test split and evaluated using the weighted F1 score. The training process involved imputing still missing values using the mean strategy.

Results and Feature Importance

The optimised model achieves a best F1 score of 0.795, indicating a reasonable level of predictive accuracy. Figure 32 depicts the top 20 Feature Importances as derived from the XGBoost model. The feature importance analysis revealed that the most LLaMA embeddings are the most influential feature, suggesting that certain linguistic structures or meanings significantly boost the chances of success. It indicates that textual information about campaigns carries substantial predictive power. Categorical features such as campaign categories (e.g., Gadgets) also carry high importance. It is interesting to see why these LLaMA embeddings have this large influence. This is further explored in the Sensitivity Analysis in Chapter 6. Moreover, the Discussion Chapter reflects on the choice for these embeddings, and the different result that is obtained compared to the sentence-transformers/all-MiniLM-L6-v2 embeddings.

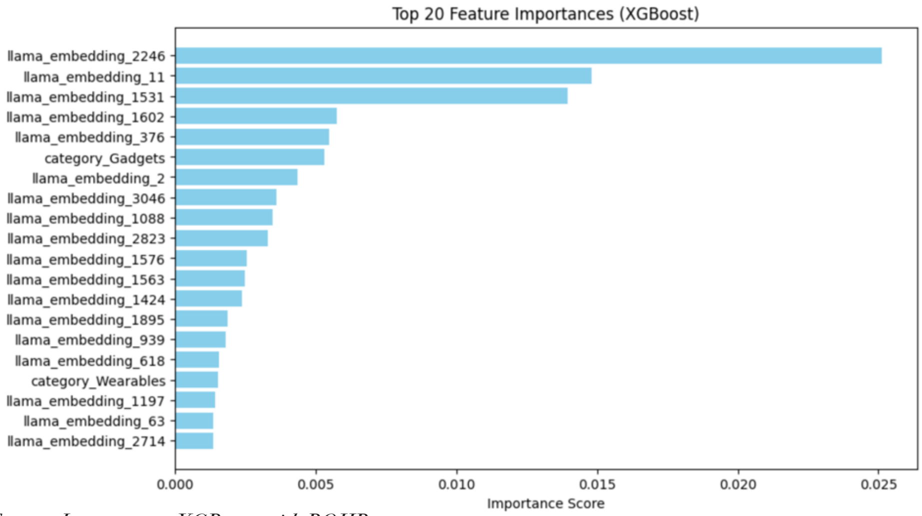


Figure 29: Top 20 Feature Importances XGBoost with BOHB

5.4 Reward-Tier Optimiser with Customisation Strategy

The next step to evolve the model, is to implement it in a real use-case scenario, where it has practical relevance, a tool campaign creators can use to optimise their own campaign. The model introduces several advancements over the previous model presented in the previous sub-chapter. It now also incorporates the average calculated sentiment of the reward tiers, to also inform the campaign creators which tone to use to address the potential backers. The BOHB is applied through Optuna, state-of-the-art hyperparameter optimization framework, as used in the study by Nguyen & Liu (2025). Optuna uses the same reward-based features as in the previous model to optimise, along with the added sentiment score, from -1 for negative, and 1 for positive. Similarly, as in the previous model, the search space had the reward count ranging from 1 to 33, complying with the finding in literature that there is a tipping point starting at 33 reward options (Elitzur et al., 2024). The space for the average reward price spans from 10 to 1000, the early bird and the exclusive flags are binary indicators, and the reward name cluster contains either a 0 for no cluster that is matched, or 1 to 4, which represent 4 different reward clusters with a certain semantic meaning. Similar as before, Optuna aims to maximize the F1-score, for a given number of trials.

Examples of optimisation studies

We now perform five different studies, to discover how the model changes its predictions for different input. We investigate whether the F1-score will change for different input, and how certain parameters can take more specific values tied to different categories. For each study, we take n=20 trials, as the incremental benefit of increasing trials starts to taper off after around 20–30 in many optimisation scenarios. This merely illustrates how different inputs affect the output. Table 8 and 9 show the conducted studies.

The optimisation process also reveals deeper insights into campaign design principles. For instance, while early-bird rewards are traditionally viewed as critical for driving initial momentum, the optimisation indicates that they may not always be necessary, especially when exclusive rewards are present. Similarly, the clustering of reward names hints at thematic preferences among backers that can be strategically targeted for higher engagement. We discuss these findings further in the next Chapter, the Evaluation, and the Discussion Chapter.

Study	Category	Country	Campaign Duration	Funding Goal ($)
1	Apps	The Netherlands	30 days	50,000
2	Gadgets	United States	45 days	50,000
3	Software	United States	60 days	100,000
4	Design	United Kingdom	30 days	25,000
5	Games	Japan	20 days	15,000

Table 8: 5 Different sets of study parameters

Study	F1-score (rounded)	Reward count	Average reward price ($)	Early bird	Exclusive	Sentiment (rounded)	Reward name cluster
1	0.792832	1	479.40	0	0	-0.3722	2
2	0.785858	28	226.76	0	0	-0.2961	0
3	0.785858	27	230.46	0	1	-0.9974	3
4	0.785858	19	208.66	0	1	-0.6292	1
5	0.785513	28	298.19	0	1	-0.0744	4

Table 9: 5 Different outcomes of studies

5.5 Concluding words Chapter 5

In this chapter, we presented the development and optimisation of a multi-step ML pipeline to predict and enhance the success of crowdfunding campaigns. Starting with a Random Forest baseline model, we established a reference for performance and identified key predictive features. Although effective in its simplicity, the model lacked the capacity to capture the semantic richness of campaign descriptions, which led us to integrate LLaMA embeddings in the next phase. These embeddings allowed us to represent textual information more effectively, improving the model's understanding of narrative and emotional tone.

Subsequently, we introduced XGBoost as the primary predictive model due to its superior performance in terms of F1 score and interpretability of feature importance. To further optimise the predictive accuracy, we used BOHB for hyperparameter tuning, refining key parameters related to campaign attributes and reward features. This optimisation process not only improved model performance but also revealed the critical impact of certain features, such as early bird rewards, average reward price, and campaign sentiment, as in agreeance with literature.

The final phase introduced the reward-tier optimiser using Optuna, aimed at optimising reward structures for specific campaign scenarios. By combining sentiment analysis with reward-based features, we created a more nuanced model capable of suggesting optimal configurations for different categories, countries, and funding goals. Several exploratory studies demonstrated how varying these parameters influenced campaign success, providing valuable insights into strategic reward design.

Overall, the model progressed from basic prediction to targeted optimisation, offering a data-driven tool to improve reward strategies in crowdfunding. In the following chapters, we will further evaluate its practical implications, limitations, and potential for future research.

6. Model Evaluation

This final chapter, before the discussion and conclusion, reflects on the performance of the model. First, it is compared with previous literature research. How does the baseline model perform compared to models presented earlier in crowdfunding research, as quite a vast number of models have been tested? Second, the models presented in this thesis are compared. The baseline model, trained with other embeddings than the LLaMA embeddings, are compared. There is reflected on the choice of ML learning models, with XGBoost chosen as preferred model for the other models using BOHB. The inner workings of the model are explored through a deeper analysis, and a sensitivity analysis is applied to certain variables. There is reflected on the robustness of the model, as well as an error analysis. There are many ways to iterate on the model, or to boost its performance. These are discussed in the next Chapter, the Discussion.

This Chapter completes the answers to the last research questions; namely:

- What are the optimal hyperparameters and configurations for BOHB in combination with NLP and other ML models to maximise prediction accuracy, and how do they compare to a baseline model?
- How can AI-driven optimisation of reward-tier configurations enhance crowdfunding campaign success by balancing campaign outcomes, reducing backer choice complexity, and minimising financial risks?

6.1 Comparison with literature

In this section, we systematically compare the findings of our models, including the baseline Random Forest model, LLaMA 3.2 embedding-based models, and the optimised XGBoost and MLP models, with the literature reviewed in Chapter 2. This comparison focuses on two key areas: predictive performance and the influential features.

Performance
First, we look at the baseline model we created. The predictive accuracy of our baseline RF model achieved an accuracy of 76.96%, which is in line with the findings from the literature. For instance, Zhong (2022) demonstrated that RF models consistently outperform other traditional ML methods like logistic regression and SVM for crowdfunding prediction tasks. Similarly, Haitham et al. (2024) observed that RF classifiers maintain stability and generalisation capability even in high-dimensional spaces, making them a good choice for baseline comparisons. When we compare its accuracy with earlier models, it seems to underperform when compared with more recent campaigns, since for example Tran et al. (2016) obtained an accuracy of 81% with RF, and Ahmad et al. (2017) obtained an even higher accuracy of 94%. However, these respectively also incorporated Twitter features, and the other additional information about the creator such as the number of Facebook friends. Thus, we conclude the model performs up to standard.
In contrast, our advanced models leveraging LLaMA 3.2 embeddings and XGBoost gave accuracy scores of 76.36% for XGBoost and 75.15% for MLP. This is consistent with literature emphasising the advantage of gradient-boosted methods and neural networks for handling complex, non-linear relationships in crowdfunding data (Y. Guo et al., 2021; Lee et al., 2018; Yu et al., 2018). Notably, our MLP model, while slightly lower in overall accuracy, excelled in recall, identifying a higher proportion of successful campaigns, aligning with W. Wang et al. (2020), who noted deep learning's strength in capturing intricate patterns missed by ensemble methods.
Furthermore, our fine-tuning with BOHB yielded additional performance gains, particularly for XGBoost and MLP, supporting findings by Ahmad et al. (2017), who showed that hyperparameter optimisation can significantly enhance model performance in crowdfunding contexts. The overall best F1-score achieved was 0.795 with the optimised XGBoost model, reflecting a balanced prediction capacity for both successful and unsuccessful campaigns, resonating with findings from Tran et al. (2016) and Y. Li et al. (2016).

Influential Features
Most of the models highlight the importance of specific features that also were present in literature. The baseline RF model showed the funding goal as the most critical predictor of success, agreeing with findings by Mollick (2014) and Kuppuswamy & Bayus (2018), who emphasised realistic goal setting as essential for backer confidence. Similarly, the inclusion of embeddings in our advanced models revealed that semantic richness in reward descriptions directly influenced campaign success. This observation aligns with the work of Yosipof et al. (2024), who found that positive sentiment and detailed language significantly impact technology-focused campaigns. Interestingly, the SHAP analysis of our baseline model showed that the reward count is a top predictor, reflecting the choice overload hypothesis described by Elitzur et al. (2024), which found a tipping point at 33 reward options. Furthermore, campaign categories, particularly technology and gadgets, are consistently influential, matching findings by Corsini & Frey (2023), who notes that category alignment with backer expectations enhances campaign visibility and funding potential. The dominance of the early bird flag indicates that early incentives are a strong predictor of campaign success and is also mentioned by Wessel et al. (2019), known to drive early engagement.

6.2 Sensitivity Analysis

The Sensitivity Analysis aims to evaluate the robustness and reliability of the model by systematically altering key input features and observing the changes in predicted outcomes. This helps to understand which features have the strongest influence on campaign success and identify any vulnerabilities or biases in the model. We execute this analysis since we noticed some strange values in the Reward-Tier Optimiser with Customisation Strategy Model, some F1-score values were the same, even when entering different values as input. During the different trials, there were not many duplicate F1-score values, but it seems worth to investigate this observation. There could be several reasons causing this phenomenon: XGBoost is a powerful classifier, but it may learn generalised patterns for certain types of campaigns. This could particularly be the case if LLaMA embeddings for these two categories are semantically close. Another reason could be that the hyperparameter space has flat regions, where small changes to some parameters do not affect the outcome. This is a known type of behaviour in complex models. One other reason could be that the number of trials were too little. With performing a Sensitivity Analysis, we can find out how much change one certain parameter in the model makes. For now, we stick to analysing the general model, as the customised model has so many different configurations, analysing it would be an extensive process.

Feature Importances

We already plotted the 20 most important features of the model, by the XGBoost model in Figure 29, and now we add the SHAP analysis in Figure 30. We find that most of the important features, in both plots, are certain LLaMA Embeddings, and certain campaign categories (e.g., Gadgets) – also referred to as subcategories, as the main category remains Technology – carry high importance as well. We look at them more closely to find out what they represent, as they are still abstract representations of language. We take both the LLaMA Embedding number 1602 and 2246, as they arose as the features with the highest influence on success. In Figure 31 the distributions of both variables are plotted for the cases of successful, and unsuccessful campaigns. The correlation with success for the LLaMA Embedding 1602 is -0.3975. The correlation for the LLaMA Embedding 2246 is 0.3748. In Figure 30, the different projects associated with these embeddings can be seen. The projects highly associated with Embedding 1602 seem to have a strong focus on security, artificial intelligence, and automation. For the lower values, there is a shift towards more niche markets and physical consumer products. This suggests that perhaps more niche campaigns have a higher success rate. On the other hand, the LLaMA Embedding 2246 may capture physical technology innovation, particularly around devices and fabrication, and for the negatively associated projects there is a noticeable drop-off in technological hardware projects, and more emphasis is placed on social initiatives or commemorative events. This is similar to the findings of the RF baseline model.

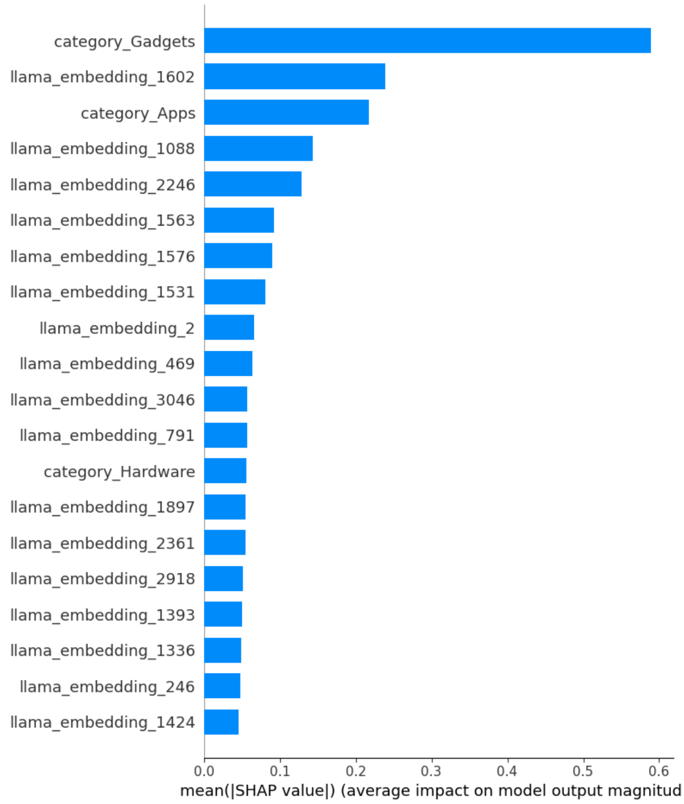


Figure 30: SHAP values XGBoost model with BOHB

Sensitivity Analysis

For the Sensitivity Analysis, we focus on the most influential features identified during the model development phase, and the most influential LLaMA Embeddings derived from the Feature Importance

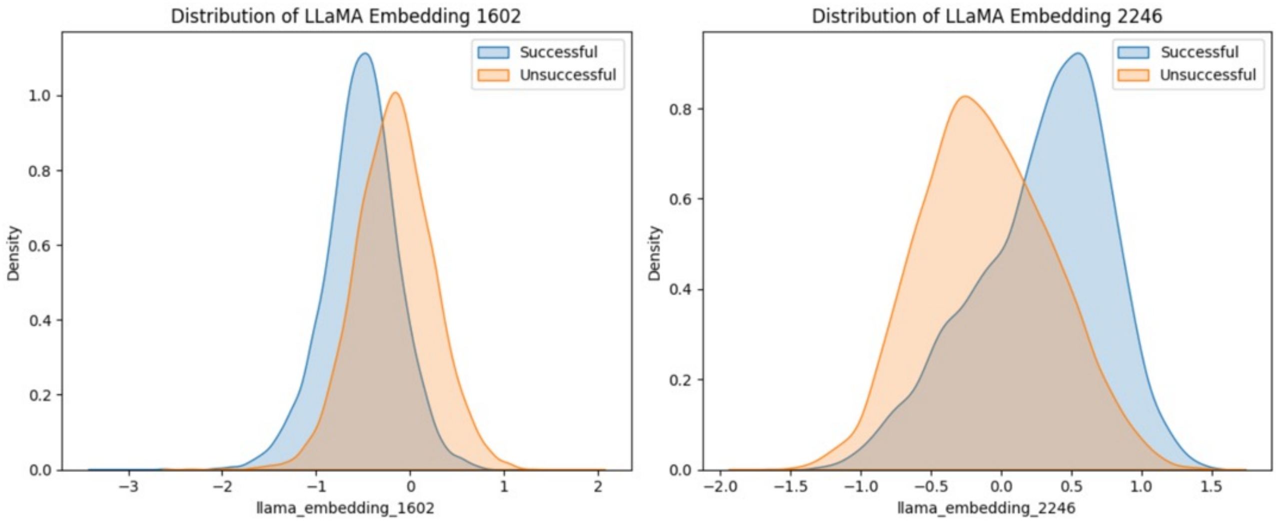


Figure 31: Distribution LLaMA Embeddings with Success labels

analysis and SHAP Analysis. These include the Reward Count, Average Reward Price, Early Bird Flag, Exclusive Flag, Sentiment Score, Reward Name Cluster, Campaign Duration, Funding Goal, and the LLaMA Embeddings 2246, 11, 1531, 1602, 376, 1088, 1563. The represented Sensitivity Analysis is shown in Figure 33. We will now carefully examine the sensitivity scores.

- Reward Count: The F1-score indicates that the model is only sensitive to a very large number of rewards. This contradicts literature and previous assessment of the data, where the optimum was found around 10. Moreover, there should be a higher sensitivity for lower reward numbers as well.
- Average Reward Price: There is a clear upward trend having a high plateau between $300–$900. This observation aligns with literature suggesting that backers are attracted to well-priced rewards that balance value and exclusivity.
- Early Bird Flag: The model's performance remains constant regardless of the early bird flag's value. This suggests that early bird incentives may not strongly

🔍 Top 10 highest values for llama_embedding_1602:

	project_name	llama_embedding_1602
605	ai-bug-bounty-automated-tool	1.868
5635	imprendiio	1.345
13953	virtual-private-safe-vps	1.339
2074	car-prototype-alternative-vehicle-fuel-source	1.230
1093	as-simple-as-possible-deep-learning-tensorflow...	1.187
8124	networking-social-media-and-educational-platfo...	1.179
14557	wwwlawyervaultcom	1.127
1789	brain-computer-music-interface-3d-printed-records	1.120
1381	battledrone	1.114
4112	findit-crea-il-tuo-business-trovando-i-tuoi-cl...	1.107

🔍 Top 10 lowest values for llama_embedding_1602:

	project_name	llama_embedding_1602
4305	fokis-reinventing-how-we-photograph-kids-and-pets	-3.230
13301	touchconnect-pro-smartscreen-with-built-in-wif...	-2.637
14823	zeromouse-coin-size-smart-wireless-mouse-for-w...	-2.633
5920	isteady-q-smart-selfie-stick-tripod-with-360-a...	-2.588
13844	vegetarano	-2.434
7029	magrota-worlds-lightest-and-full-function-filt...	-2.418
3110	dipongo-0	-2.355
1516	big-bang-lab	-2.332
2121	caring-men-global-mobile-app-0	-2.285
8038	narvalo-urban-active-the-first-iot-and-adaptiv...	-2.250

🔍 Top 10 highest values for llama_embedding_2246:

	project_name	llama_embedding_2246
102	360-transpo-camera	1.506
3534	ecubmaker-toydiy2-4-in-1-3d-printer	1.497
5947	iv3-clock-circuitpython	1.492
4356	formaker-4-in-1-cnc-mill-laser-pcb-3-d-printer...	1.461
7192	master-arduino-and-esp-microcontrollers-0	1.451
13151	ticktalk-4	1.448
13526	tx-a-small-powerful-and-affortable-55w-laser-e...	1.444
146	3d-printable-christmas-gnomes-and-ornaments	1.406
10240	rgb3dscanner-enable-your-smartphone-to-3d-scan...	1.398
1203	auto-feeder	1.381

🔍 Top 10 lowest values for llama_embedding_2246:

	project_name	llama_embedding_2246
6763	lockdown-notes	-1.708
12992	the-wei-a-platform-to-unleash-the-leader-in-you	-1.641
13527	tyb-the-worlds-first-hot-air-balloon-powered-b...	-1.635
3757	entertainment-helicopter	-1.499
5281	historic-aircraft-rescue	-1.496
9497	priceless-moments-capture-your-moments-with-yo...	-1.470
9499	prideflight2018	-1.461
304	75-year-commemorative-flight-in-2019	-1.418
14067	vring-stop-swiping-and-start-living	-1.388
13588	ugale-the-first-insole-with-active-ventilating...	-1.374

Figure 32: Projects with a high or low correlation of the selected LLaMA Embeddings

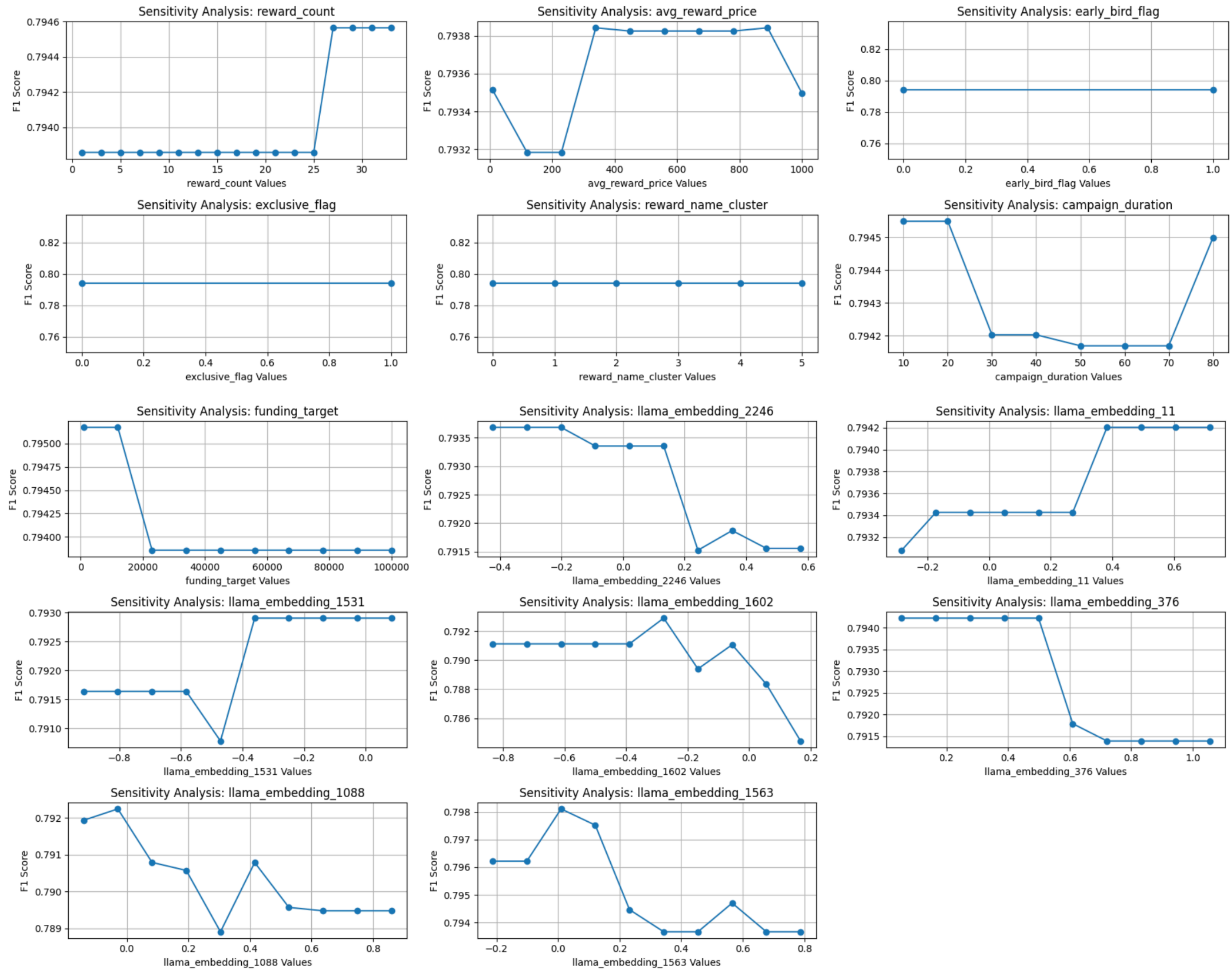


Figure 33: Sensitivity Analysis

influence the model's assessment of campaign success. This is highly unexpected and does not agree with literature and the earlier findings with the top 20 reward names in the dataset.

• Exclusive Flag: Similar to the early bird flag, this is not expected. In literature and the data inspection, it showed making use of exclusivity, increases the chance of campaign success.

• Reward Name Cluster: The model's sensitivity is low across the different clusters, although there are slight variations. This indicates that the linguistic framing of reward names does have some influence, but it is not a primary driver of predictive success, which is also not expected.

• Campaign Duration: There is a sharp peak in performance for campaigns around 10-20 days and 70–80 days. It did show up in the violin plot shown earlier in Chapter 4, there are simply no unsuccessful values for 70 or more days, thus the model may assume this leads to success.

• Funding Target: A steep decline in the F1-score is observed as the funding target increases beyond $20,000. This aligns with crowdfunding literature.

All the LLaMA embeddings cause the F1-score to clearly fluctuate with changes in this embedding's values. This suggests that specific semantic patterns captured by these dimensions, possibly related to technical innovation or specific project descriptors, impact the model's prediction.

6.3 Review Embeddings

The sensitivity analysis points out a strange occurrence: the early bird flag seems of no influence. Aside from that, literature is not in agreeance with the number of rewards, nor the exclusivity flag. One possible explanation could be the use of the LLaMA embeddings. As they are abstract of meaning, it is difficult to see what they represent, and if there is possibly a feature redundancy and overfitting. When inspecting the data that was used to generate these embeddings, it became apparent that the LLaMA-based embeddings were generated from composite prompts that already contained many of these structured variables. The embeddings were based on the campaign titles, the campaign description, the reward tiers, the goal, and duration.
This likely caused feature leakage, where the embedding vectors already encoded information that was separately passed as tabular input. As a result, structured inputs added little value—or even distorted the model's learning process—because the same information was redundantly presented in two different forms.

Testing with Alternative Embeddings
To test this hypothesis, we replaced the LLaMA embeddings with sentence-transformers/all-MiniLM-L6-v2 embeddings. These were generated only from the descriptive campaign texts (campaign blurb and reward titles), excluding numeric features like goal or duration. First, the XGBoost model was evaluated with these new embeddings. The results in Table 10, and the visuals of Figure 34, show how they compare to the RF model that used the same embeddings. The results are still better than the RF model, and they outperform the XGBoost model with the LLaMA Embeddings.

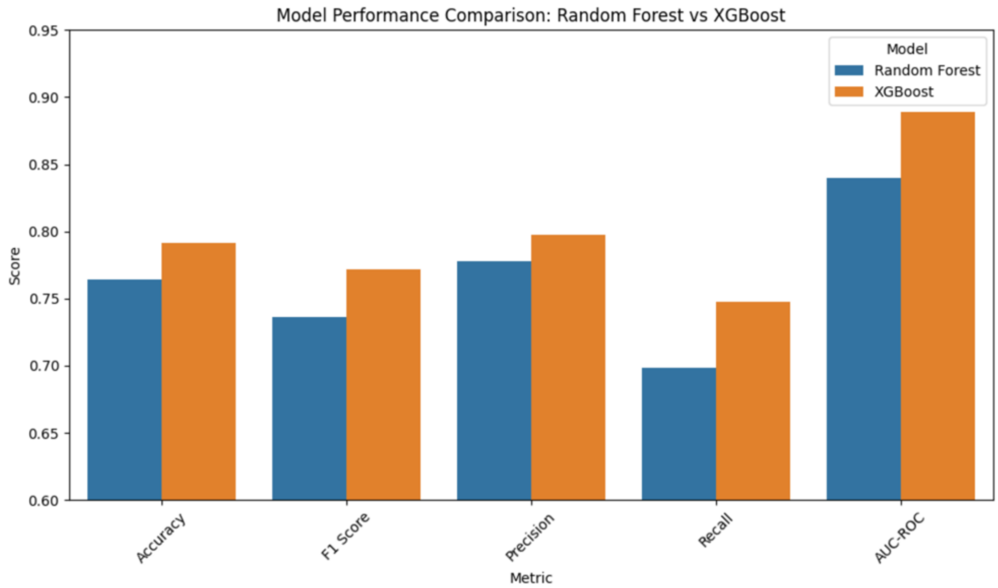


Figure 34: Comparison model performance RF and XGBoost same embeddings

Model	Accuracy	F1 Score	Precision	Recall	AUC-ROC
Random Forest	0.7645	0.7364	0.7782	0.6988	0.8401
XGBoost LLaMA Embeddings	0.7636	0.7434	0.7456	0.7413	0.8471
XGBoost MiniLM Embeddings	0.7918	0.7717	0.7976	0.7475	0.8887

Table 10: Metrics of RF and XGBoost with 2 different embeddings

Compared to earlier results with LLaMA embeddings (see Table 7 in Chapter 5), this configuration shows clear improvements across all metrics, especially in AUC-ROC and F1 Score. This supports the hypothesis that more focused embeddings, which avoid information overlap with structured features, lead to better generalisation. When we look at the feature importance, we do find that certain properties, such as the sub-categories, goal, campaign duration and average reward price are now more prominent in the top 20.

BOHB with New Embeddings
We then applied Bayesian Optimisation with Hyperband (BOHB) using the same refined sentence-transformer embeddings. The objective was to find optimal hyperparameters of the campaign properties for the XGBoost model without tuning campaign structure. For a 20-trial run, we found: Best Parameters: {'reward count': 4, 'avg_reward_price': 570.212868382487, 'early_bird_flag': 1, 'exclusive_flag': 0} Best F1 Score Achieved: 0.7806941628336065.
When we test the complete model for its performance on the test set, we find the results presented in Table 11. These results are as well better than the previous model. It seems the MiniLM Embeddings deliver better results in this specific situation than the LLaMA Embeddings.

Early Bird Flag
As a last step, we explore if the model now resembles more the finding that adding an Early Bird Reward will increase the chance of campaign success.
Unexpectedly, we find that in this model as well, the early bird flag has little significance. Most surprisingly, there is even a slightly higher chance on success without having an early bird presence.
In Figure 36, the red line (median) shows that most campaigns experience 0 effect on flipping the early bird presence. The mean change equals -0.015. On average, enabling the early bird reward decreases the predicted probability of success by ~1.5 percentage points. The fraction where change is larger than 0 equals 0.48: in 48% of the samples, the probability increased when enabling early bird.

Feature	**Importance**	**(Gain)**
12	category main_Gadgets	0.040019
647	emb_reward_title_242	0.008467
762	emb_reward_title_357	0.008319
13	category main_Hardware	0.008137
20	category main_Web	0.006619
707	emb_reward_title_302	0.005853
472	emb_reward_title_67	0.005768
0	goal in dollar	0.005702
766	emb_reward_title_361	0.004926
716	emb_reward_title_311	0.004688
1	campaign duration	0.004235
11	category main_Flight	0.004069
453	emb_reward_title_48	0.003874
7	category main_Apps	0.003819
729	emb_reward_title_324	0.003738
436	emb_reward_title_31	0.003571
2	avg_reward_price	0.003533
607	emb_reward_title_202	0.003387
445	emb_reward_title_40	0.003321
556	emb_reward_title_151	0.003249

Table 12: Feature importance XGBoost other embeddings

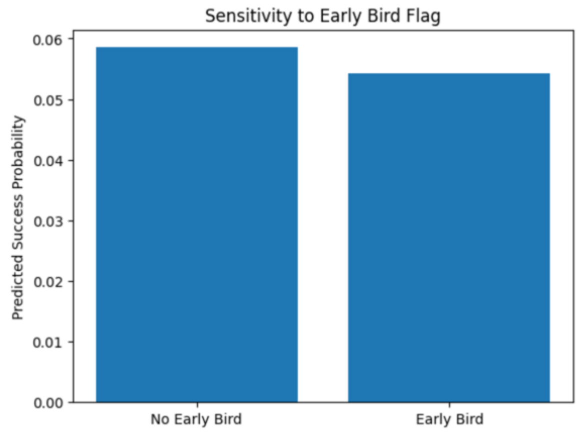


Figure 35: Sensitivity to Early Bird flag

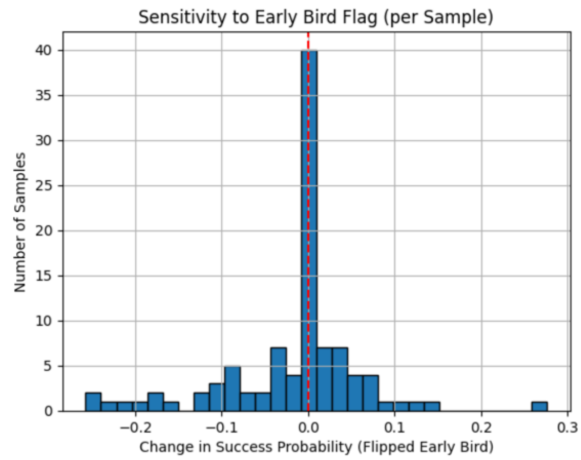


Figure 36: Sensitivity to Early Bird flag per sample

Accuracy	F1 Score	Precision	Recall	AUC-ROC
0.8180	0.8177	0.8182	0.8180	0.9011

Table 11: Metrics BOHB with MiniLM Embeddings

6.4 Model Comparison

In this section, we present a comparative evaluation of the different models developed throughout this thesis. The goal is to analyse their predictive performance in forecasting campaign success, as well as to evaluate the effectiveness of various feature engineering techniques and optimisation strategies. The models evaluated are the LR, RF baseline model, the Advanced ML Models with LLaMA Embeddings (RF, MLP, XGBoost), and Generalized Success Prediction Model with XGBoost and BOHB, as well as the models with the other embeddings, discussed in the previous subchapter.

Evaluation Metrics
To ensure consistent evaluation across models, we used the following performance metrics:

- Accuracy: Measures the overall correctness of predictions.
- Precision: The proportion of true positive predictions among all positive predictions.
- Recall (Sensitivity): The proportion of true positives captured among all actual positives.
- F1-Score: The harmonic mean of precision and recall
- AUC-ROC: The trade-off between sensitivity and specificity.

The F1-Score is emphasised as the primary evaluation metric throughout the thesis, given its capacity to balance precision and recall, especially in cases where false positives and false negatives carry distinct implications for campaign investment decisions.

Model Performance Overview
Below in Table 12, we summarise the key performance metrics obtained for each model.
We see that on most metrics, the RF Baseline model, with the sentence-transformers/all-MiniLM-L6-v2 embeddings, performs better than with the LLaMA embeddings. We notice the XGBoost models perform better with the sentence-transformers/all-MiniLM-L6-v2 embeddings instead of the LLaMA embeddings. Initial LLaMA embeddings may have been too dense and information-rich, leading to redundancy and degraded model interpretability. By refining the source of embeddings and clearly separating textual and structured features, the model achieved higher accuracy, stability, and generalisation.
One can clearly see the XGBoost is the preferred model to use in the continued set-up, as it has a higher accuracy and precision. Both the optimised models with BOHB, outperformed all other models in terms of F1-Score. Its ability to handle structured and unstructured data efficiently makes it well-suited for crowdfunding prediction. The introduction of BOHB improved hyperparameter tuning, yielding a substantial increase in F1-Score. This demonstrates the effectiveness of structured hyperparameter search for optimising campaign configurations.

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Logistic Regression	0.7428	0.7169	0.7326	0.7247	0.8150
Random Forest (Baseline)	0.7696	0.8044	0.6748	0.7339	0.8401
RF with LLaMA Embeddings	0.7461	0.7492	0.6773	0.7114	0.8151
MLP with LLaMA Embeddings	0.7515	0.7033	0.7994	0.7483	0.8321
XGBoost with LLaMA Embeddings	0.7636	0.7456	0.7413	0.7434	0.8471
MLP optimised	0.7670	0.7467	0.7500	0.7484	0.8472
XGBoost optimised	0.7703	0.7582	0.7384	0.7482	0.8466
XGBoost with MiniLM embeddings	0.7918	0.7976	0.7475	0.7717	0.8887
XGBoost + BOHB (General Model)	0.7944	0.8017	0.7491	0.7952	0.8808
XGBoost + BOHB with MiniLM embeddings	0.8180	0.8177	0.8182	0.8180	0.9011

Table 12: All Performance Metrics

6.5 Concluding words Chapter 6

When comparing the performance of the initial models to literature, no strange values are seen. The models boosted by BOHB yielded additional performance gains, which was expected by the studies described in Chapter 2. The baseline model showed the funding goal as the highest predictor of success, which aligned with literature as well. Moreover, the inclusion of embeddings in all our advanced models revealed that semantic richness in reward descriptions directly influenced campaign success. In the SHAP analysis of the baseline model the reward count is also a top predictor, which is in agreeance with literature as well. Other important features discovered in the baseline model are the early bird flag and the category.
Looking at the performance of the model, in comparison to each other, we find that the final model does have the highest scores on all metrics, except the recall, which is highest of the optimised MLP model. It turned out that LLaMA Embeddings are not the most suitable in this context, and the sentence-transformers/all-MiniLM-L6-v2 embeddings performed better. While this chapter continuously shows a strong performance of the model, there are some peculiar observations. The feature importance of the baseline Random Forest model is almost completely different than the feature importance of the XGBoost model in combination with BOHB. With the latter model, the early bird presence is not correlated with a higher chance of success of the campaign. The next Chapter reflects on this finding. The sensitivity analysis reveals that some of the embeddings are the most important features of the model predicting campaign success.

7. Discussion

First, we reflect on the results obtained with the model. We reflect on the design process of the model, improvements that could have been made, and if the outcomes were expected. We inspect the weak points of the model, evident from the Evaluation. Furthermore, the practical relevance of the model is evaluated. Subsequently, we look at inherent weak properties of the model, annotations on certain undesired properties the model components can bring, and the limitations in time and computing power. Last, we discuss possible improvements for future versions of the model by employing a more advanced or extended model.

Model performance and practical implications
We reflect on the most notable observations of the Evaluation; the difference of performance with different embeddings and the difference of feature importance in the baseline Random Forest model and the advanced XGBoost with BOHB model. One area of potential model fragility lies in the choice of embeddings. While the LLaMA 3.2 3B Instruct model offers powerful semantic representations, it is relatively new and has not yet been widely benchmarked in structured prediction tasks. Public leaderboards, such as the MTEB Hugging Face benchmark, demonstrate that lighter-weight models like all-MiniLM-L6-v2 or mpnet-base-v2 still outperform LLaMA variants on various sentence-level retrieval and classification tasks. Thus, testing multiple embedding models, including text-embedding-ada-002, which although paid, offers strong performance, could help validate the robustness of the language features used. This comparative analysis would help ensure the observed model behaviour is not an artifact of a particular embedding strategy. Moreover, the feature importance in the latter model is not always in coherence with findings in literature. In the Random Forest model, traditional reward-based features like reward count, average reward price, and the early bird flag were assigned significant importance, aligning with the literature that we outlined in Chapter 2 and underscores their influence on campaign success (Y. Lin et al., 2016; Wessel et al., 2019) However, in the advanced XGBoost + BOHB model, the early bird flag, which is often cited as a critical mechanism for driving early backer engagement, emerged as almost entirely irrelevant in the sensitivity analysis. This finding directly contradicts established research and general assumptions in crowdfunding literature, which shows that early momentum is a strong predictor of campaign success. Unfortunately, the near-zero importance of the early bird flag is not just an anomaly, it introduces questions about the interpretability and practical relevance of our XGBoost + BOHB model. Since the Kickstarter dataset is widely adapted by researchers, this cannot be caused by data that has different values. If one of the most widely accepted mechanisms for boosting initial backer interest is deemed ineffective by the model, it implies that either:

- The model is not capturing certain temporal or behavioural dynamics correctly.
- There is a possible feature interaction effect not captured by the individual analysis, such as early bird rewards only being effective when paired with specific price points or campaign durations.

Furthermore, the implications for the reward-tier optimisation framework are substantial. The primary aim of the model is to optimise reward-based parameters to enhance campaign success. If critical levers like early bird options are dismissed as ineffective, the actionable insights the model provides may lack practical utility for campaign creators. This diminishes the model's relevance in real-world scenarios where campaign strategists rely on early momentum tactics as a proven method for driving backer engagement.
To address this issue, a deeper temporal analysis can be employed to see if the early bird effectiveness is time-sensitive and is not present in the current model structure. We could also test interaction terms to examine whether early bird flags have conditional dependencies with other parameters, such as campaign duration or average reward price. One other possibility is revisiting feature engineering to ensure that the representation of early bird dynamics is correctly captured in the input data.
Without addressing these aspects, the model's capacity to provide meaningful optimisation for reward tiers sadly remains somewhat limited, as without understanding this, the consequences of placing an Early Bird offer in a campaign are now not clear with this model.

Limitations and Challenges

The ML mechanisms used do not only bring advantages. There is a trade-off for the fast-computing method that BOHB uses. Generalisation is a common concern for multi-fidelity algorithms, like BOHB, since they extract subsets to represent the entire dataset (L. Yang & Shami, 2020). Aside from this, it can be difficult to correctly apply NLP as it is so complex. As mentioned, LLMs can exhibit social biases and toxicity during the generation process. This can give biased outputs. They also have limitations in using the most up to date data, they cannot incorporate real-time or dynamic information. Moreover, LLMs can be sensitive to adversarial prompts: subtle input manipulations that cause erratic outputs (Chang et al., 2023). One other crucial consideration in introducing the LLaMA embeddings is how they interact with the existing structured features, such as campaign duration, goal amount, and category. In the current setup, LLaMA embeddings are integrated alongside these structured features without explicit cross-referencing or interaction modelling. This design choice allows the model to independently learn the contribution of semantic content versus numerical and categorical data. However, it also opens up potential dependencies that could influence predictions in unintentional ways. For instance, if certain narrative styles are more prevalent in specific categories or are historically linked to particular goal ranges, the embeddings might inadvertently capture these biases, amplifying correlations that may not generalise well. This possibly could have caused the strange occurrences we just observed. Possibly, the model could be inspected without the integration of the LLaMA Embeddings to see if this causes any significant shifts in feature importance or model performance. This comparative analysis would help isolate the true contribution of LLaMA embeddings versus structured numerical and categorical features, revealing whether their inclusion introduces noise, amplifies biases, or genuinely improves predictive capacity. Furthermore, it would allow for a deeper dive of the embeddings' role, whether they primarily capture narrative quality, sentiment, or merely reflect latent patterns tied to campaign categories or funding goals.

Future Enhancements and Optimisations

While working on this thesis, the many different options to tackle the presented problem became evident. There are many different approaches one could take; one evident option would be to use BOHB to tune a LLaMA model, or similar LLM. There are several possibilities to do this. One could fine-tune LLaMA (or a smaller model like BERT) on a task (e.g., success prediction or sentiment classification). Or combining other tuning methods with BOHB; an alternative approach could involve parameter-efficient fine-tuning methods such as delta-tuning. Delta-tuning techniques, such as LoRA and adapters, have been shown to significantly reduce computational costs by modifying only a small subset of model parameters instead of fine-tuning the entire model. This could be a promising direction for future research, especially in resource-constrained environments. Another addition could be to include parameters such as more reward information, if there is a physical reward or digital reward, if it includes an experience, the number of limited-edition tiers, or the percentage of tiers with emotional language. Moreover, it would be interesting to use prompt engineering to create better reward descriptions.

The conclusion that can be derived from this thesis, is that we have proven the use of language has a substantial effect on the outcome of the campaign in the category Technology, as we continuously saw that certain embeddings scored high in feature importance. This matches with earlier findings in literature by Yosipof et al. (2024). Therefore, it would be interesting to explore the options even more to predict crowdfunding success with NLP, as this is a relatively undiscovered research area. Moreover, it would be very interesting to combine it with the analysis of media. LLaMA is currently working on multimodal extensions to the models, enabling image recognition, video recognition, and speech understanding capabilities (Grattafiori et al., 2024), and when this is available, could support the crowdfunding research.

8. Conclusion

This thesis presented an integration of Bayesian Optimization HyperBand (BOHB) and LLaMA 3.2 Embeddings to predict and optimise the success of Kickerstarter Technology crowdfunding campaigns, with a particular focus on reward-tier configurations. This is a novel contribution to the existing crowdfunding literature. The study began with a comprehensive data collection process that combined publicly available Kickstarter data with custom-scraped reward-tier information. After this data was pre-processed into structured datasets and an exploratory analysis was performed, a baseline Random Forest (RF) model was created. This basic model was used to compare the findings and its performance with literature and form a reference point for the other models that were developed. The basic model scored 0.7428 on accuracy, 0.7169 on precision, 0.7326 on recall, 0.7247 on the F1-Score and 0.8150 on AUC-ROC. The best performing model, the XGBoost BOHB model with all-MiniLM-L6-v2 Embeddings scored 0.8180 on accuracy, 0.8177 on precision, 0.8182 on recall, 0.8180 on the F1-Score, and 0.9011 on the AUC-ROC.

Building upon that baseline model, we introduced LLaMA 3.2 Embeddings to capture the semantic richness of campaign descriptions and reward titles. These Embeddings were integrated into three different Machine Learning models: RF, MLP, and XGBoost. While the RF model with LLaMA Embeddings did not improve its performance, we continued to pursue this strategy as the LLaMA embeddings encompass far more dimensions than the embeddings used in the baseline model. The advanced models demonstrated improvements in capturing textual nuances and campaign features, with XGBoost emerging as the best-performing model in terms of F1-score, accuracy, and interpretability. BOHB was used to tune the hyperparameters of the XGBoost and MLP models, improving performance significantly, confirming its efficacy in high-dimensional, multi-fidelity settings.

During Evaluation, the all-MiniLM-L6-v2 showed improved results over using LLaMA Embeddings, as the initial LLaMA embeddings may have been too dense and information-rich, leading to redundancy and degraded model interpretability.

The primary innovation of this study is the Reward-Tier Optimiser, which applies BOHB to fine-tune reward parameters, including price, exclusivity, and sentiment, based on campaign characteristics. This optimiser not only predicted campaign success but also suggested optimal reward configurations for maximum backer engagement. Five case studies demonstrated the model's ability to adapt its predictions based on campaign type, country, duration, and funding goal, highlighting its possible real-world applicability.

Despite its successes in configuring the different reward-tier configurations, the study revealed some limitations, particularly in the interpretability of LLaMA embeddings and the unexpected insignificance of early bird flags—contradicting established crowdfunding literature. This anomaly underscores the need for further exploration of temporal dynamics and feature interactions in crowdfunding campaigns. Moreover, the sensitivity analysis showed that certain reward-based features were underutilised by the model, indicating potential for refinement in feature representation and interaction modelling.

From a practical standpoint, this thesis validates the integration of advanced NLP techniques with HPO methods like BOHB for reward optimisation in crowdfunding. It paves the way for campaign creators to utilise Machine Learning not only to predict campaign outcomes but also to strategically design reward structures that enhance backer engagement. The findings also open avenues for future research in multi-modal analyses, where LLaMA's evolving capabilities could integrate visual and textual campaign elements for even greater predictive power.

Ultimately, this research shows potential for creating models that can assist us in decision-making, which can capture invisible semantic meaning. It makes a significant step toward data-driven campaign design and successful crowdfunding outcomes.

9. References

About – Kickstarter. (n.d.). Retrieved October 9, 2024, from https://www.kickstarter.com/about?ref=global-footer

Abualigah, L., Al-Ajlouni, Y. Y., Daoud, M. Sh., Altalhi, M., & Migdady, H. (2024). Fake news detection using recurrent neural network based on bidirectional LSTM and GloVe. Social Network Analysis and Mining, 14(1), 1–16. https://doi.org/10.1007/S13278-024-01198-W/TABLES/8

Adamska-Mieruszewska, J., Mrzygłód, U., Suchanek, M., & Fornalska-Skurczy□ska, A. (2021). Keep it simple. The impact of language on crowdfunding success. Economics & Sociology. https://doi.org/10.14254/2071-789X.2021/14-1/9

Aggarwal, M., & Murty, M. N. (2021). Deep Learning. SpringerBriefs in Applied Sciences and Technology, 35–66. https://doi.org/10.1007/978-981-33-4022-0_3/FIGURES/12

Ahmad, F. S., Tyagi, D., & Kaur, S. (2017). Predicting crowdfunding success with optimally weighted random forests. 2017 International Conference on Infocom Technologies and Unmanned Systems: Trends and Future Directions, ICTUS 2017, 2018-Janua, 770 – 775. https://doi.org/10.1109/ICTUS.2017.8286110

Ahmed, M., & Abdel-Aty, M. (2013). A data fusion framework for real-time risk assessment on freeways. Transportation Research Part C: Emerging Technologies, 26, 203–213. https://doi.org/10.1016/J.TRC.2012.09.002

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2623–2631. https://doi.org/10.1145/3292500.3330701

Aleksina, A., Akulenka, S., & Lublóy, Á. (2019). Success factors of crowdfunding campaigns in medical research: perceptions and reality. Drug Discovery Today. https://doi.org/10.1016/J.DRUDIS.2019.05.012

Allison, T. H., Davis, B. C., Webb, J. W., & Short, J. C. (2017). Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2017.09.002

all-MiniLM-L6-v2. (n.d.). Https://Huggingface.Co/Sentence-Transformers/All-MiniLM-L6-V2.

Alomari, E. A. (2024). Unlocking the Potential: A Comprehensive Systematic Review of ChatGPT in Natural Language Processing Tasks. In CMES - Computer Modeling in Engineering and Sciences (Vol. 141, Issue 1, pp. 43–85). Tech Science Press. https://doi.org/10.32604/cmes.2024.052256

Andhika Viadinugroho, R. A., & Rosadi, D. (2023). A weighted metric scalarization approach for multiobjective BOHB hyperparameter optimization in LSTM model for sentiment analysis. Information Sciences, 644. https://doi.org/10.1016/j.ins.2023.119282

Ansel, J., Yang, E., He, H., Gimelshein, N., Jain, A., Voznesensky, M., Bao, B., Bell, P., Berard, D., Burovski, E., Chauhan, G., Chourdia, A., Constable, W., Desmaison, A., Devito, Z., Ellison, E., Feng, W., Gong, J., Gschwind, M., … Chintala, S. (2024). PyTorch 2: Faster Machine Learning Through Dynamic Python Bytecode Transformation and Graph Compilation. International Conference on Architectural Support for Programming Languages and Operating Systems - ASPLOS, 2, 929–947. https://doi.org/10.1145/3620665.3640366

Attard Cortis, P., & Muir, F. (2022). Twelve tips for using Pattern Matching in data analysis for qualitative medical education research. Medical Teacher, 44(11), 1209–1213. https://doi.org/10.1080/0142159X.2021.1937587

Azadi, S., Okabe, Y., & Carvelli, V. (2025). Bayesian-optimized 1D-CNN for delamination classification in CFRP laminates using raw ultrasonic guided waves. Composites Science and Technology, 264. https://doi.org/10.1016/j.compscitech.2025.111101

Aziz, S., & Dowling, M. (2019). Machine Learning and AI for Risk Management. Palgrave Studies in Digital Business and Enabling Technologies, 33–50. https://doi.org/10.1007/978-3-030-02330-0_3/TABLES/1

Aziz, S., Nazir, M. R., Nazir, M. I., & Gazali, S. (2023). Crowdfunding A bibliometric analysis and future research Agenda. Heliyon, 9(12), e22981. https://doi.org/https://doi.org/10.1016/j.heliyon.2023.e22981

Babayoff, O., & Shehory, O. (2022). The role of semantics in the success of crowdfunding projects. PLoS ONE. https://doi.org/10.1371/JOURNAL.PONE.0263891

Baek, S., Kim, J., Yu, H., Yang, G., Sohn, I., Cho, Y., & Park, C. (2023). Intelligent Feature Selection for ECG-Based Personal Authentication Using Deep Reinforcement Learning. Sensors, 23(3). https://doi.org/10.3390/s23031230

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. http://arxiv.org/abs/2010.12421

Barnes, S. J. (2024). Understanding the effects of socialness and color complexity in listing images on crowdfunding behavior. Industrial Management & Data Systems. https://doi.org/10.1108/IMDS-07-2023-0517

Baudat, G., & Anouar, F. (2000). Generalized discriminant analysis using a kernel approach. Volume 12, Issue 10, Pages 2385 - 2404, 12(10), 2385–2404. https://doi.org/10.1162/089976600300014980

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2013.07.003

Bento, N., Gianfrate, G., & Thoni, M. H. (2019). Crowdfunding for sustainability ventures. Journal of Cleaner Production. https://doi.org/10.1016/J.JCLEPRO.2019.117751

Bi, S., Liu, Z., & Usman, K. (2017). The influence of online information on investing decisions of reward-based crowdfunding. Journal of Business Research, 71, 10–18. https://doi.org/https://doi.org/10.1016/j.jbusres.2016.10.001

BioLite FirePit. (2019, December 2). Https://Www.Kickstarter.Com/Projects/Biolite/Biolite-Firepit-See-Fire-Not-Smoke/Description.

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(2). https://doi.org/10.1002/widm.1484

Blanchard, S. J., Noseworthy, T. J., Pancer, E., & Poole, M. (2022). Extraction of visual information to predict crowdfunding success. Production and Operations Management. https://doi.org/10.1111/POMS.14083

Brazdil, P., van Rijn, J. N., Soares, C., & Vanschoren, J. (2022). Metalearning for Hyperparameter Optimization. In Cognitive Technologies (pp. 103–122). https://doi.org/10.1007/978-3-030-67024-5_6

Buscemi, A., & Proverbio, D. (2024). ChatGPT vs Gemini vs LLaMA on Multilingual Sentiment Analysis. http://arxiv.org/abs/2402.01715

Butler, L., Yigitcanlar, T., & Paz, A. (2021). Barriers and risks of Mobility-as-a-Service (MaaS) adoption in cities: A systematic review of the literature. Cities, 109. https://doi.org/10.1016/J.CITIES.2020.103036

Butticè, V., Franzoni, C., Rossi-Lamastra, C., & Rovelli, P. (2018). The Road to Crowdfunding Success: A Review of the Extant Literature. Oxford University Press EBooks. https://doi.org/10.1093/OSO/9780198816225.003.0005

Cai, Z., Zhang, P., & Han, X. (2021). The inverted U-shaped relationship between crowdfunding success and reward options and the moderating effect of price differentiation. China Finance Review International. https://doi.org/10.1108/CFRI-11-2019-0152

Calic, G., Arseneault, R., & Ghasemaghaei, M. (2021). The Dark Side of Machiavellian Rhetoric: Signaling in Reward-Based Crowdfunding Performance. Journal of Business Ethics. https://doi.org/10.1007/S10551-021-04984-0

Candogan, S. T., Cornelius, P. B., Gokpinar, B., Körpeoglu, E., & Tang, C. S. (2024). Product Development in Crowdfunding: Theoretical and Empirical Analysis. https://doi.org/10.1287/MSOM.2022.0344

Cappa, F., Franco, S., Ferrucci, E., & Maiolini, R. (2021). The Impact of Product and Reward Types in Reward-Based Crowdfunding. IEEE Transactions on Engineering Management. https://doi.org/10.1109/TEM.2021.3058309

Carradını, S., & Nystrom, E. C. (2024). Asking is Hard: Appeals Language in Kickstarter Crowdfunding Campaigns. Journal of Technical Writing and Communication. https://doi.org/10.1177/00472816241262245

Cavalcanti, G. D. C., Mendes-Da-Silva, W., Felipe, I. J. dos S., & Santos, L. A. (2024). Recent advances in applications of machine learning in reward crowdfunding success forecasting. Neural Computing & Applications (Print). https://doi.org/10.1007/S00521-024-09886-6

Chan, C. S. R., Pethe, C., & Skiena, S. (2021). Natural language processing versus rule-based text analysis: Comparing BERT score and readability indices to predict crowdfunding outcomes. Journal of Business Venturing Insights, 16, e00276. https://doi.org/https://doi.org/10.1016/j.jbvi.2021.e00276

Chan, J., Huang, Z., Liu, D., & Cai, Z. (2023). Better to Give Than to Receive: Impact of Adding a Donation Scheme to Reward-Based Crowdfunding Campaigns. Information Systems Research. https://doi.org/10.1287/ISRE.2023.1224

Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., Chen, H., Yi, X., Wang, C., Wang, Y., Ye, W., Zhang, Y., Chang, Y., Yu, P. S., Yang, Q., & Xie, X. (2023). A Survey on Evaluation of Large Language Models. http://arxiv.org/abs/2307.03109

Chemla, G., & Tinn, K. (2020). Learning Through Crowdfunding. Management Science, 66(5), 1783–1801. https://doi.org/10.1287/mnsc.2018.3278

Chen, M., Liu, Z., Ma, C., & Gong, X. (2021). A distinctive early bird price in reward-based crowdfunding. Electronic Commerce Research. https://doi.org/10.1007/S10660-019-09356-5

Chen, M. Y., Chang, J.-R., Chen, L.-S., & Shen, E. L. (2021). The key successful factors of video and mobile game crowdfunding projects using a lexicon-based feature selection approach. Journal of Ambient Intelligence and Humanized Computing. https://doi.org/10.1007/S12652-021-03146-4

Chen, M.-Y., Chang, J.-R., Chen, L.-S., & Chuang, Y.-J. (2022). Identifying the key success factors of movie projects in crowdfunding. Multimedia Tools and Applications. https://doi.org/10.1007/S11042-022-12959-0

Cheng, C., Tan, F., Hou, X., & Wei, Z. (2019). Success Prediction on Crowdfunding with Multimodal Deep Learning. International Joint Conference on Artificial Intelligence. https://doi.org/10.24963/IJCAI.2019/299

Cho, H., Kim, Y., Lee, E., Choi, D., Lee, Y., & Rhee, W. (2020). Basic Enhancement Strategies When Using Bayesian Optimization for Hyperparameter Tuning of Deep Neural Networks. IEEE Access, 8, 52588–52608. https://doi.org/10.1109/ACCESS.2020.2981072

Coakley, J., Lazos, A., & Liñares-Zegarra, J. M. (2021). Equity Crowdfunding Founder Teams: Campaign Success and Venture Failure. British Journal of Management. https://doi.org/10.1111/1467-8551.12494

Cordova, A., Dolci, J., & Gianfrate, G. (2015). The Determinants of Crowdfunding Success: Evidence from Technology Projects. Procedia - Social and Behavioral Sciences, 181, 115–124. https://doi.org/https://doi.org/10.1016/j.sbspro.2015.04.872

Corsini, F., & Frey, M. (2023). Crowdfunding sustainable products with the product search matrix: niche markets vs. mass markets. Electronic Commerce Research. https://doi.org/10.1007/S10660-023-09674-9

Cradock-Henry, N. A., Flood, S., Buelow, F., Blackett, P., & Wreford, A. (2019). Adaptation knowledge for New Zealand's primary industries: Known, not known and needed. Climate Risk Management, 25, 100190. https://doi.org/10.1016/J.CRM.2019.100190

Crescenzo, V. De, Bonfanti, A., Castellani, P., & Mota, L. (2022). Effective entrepreneurial narrative design in reward crowdfunding campaigns for social ventures. The International Entrepreneurship and Management Journal. https://doi.org/10.1007/S11365-021-00774-2

Cumming, D., Lan, Y., Shan, Y. G., & Zhang, J. (2024). Discretionary tone in reward-based crowdfunding: Do project owners talk their way to success? The British Accounting Review, 101433. https://doi.org/10.1016/J.BAR.2024.101433

Daniel, J., & Martin, J. H. (2024). Speech and Language Processing.

de Holanda, W. D., e Silva, L. C., & de Carvalho César Sobrinho, Á. A. (2024). Machine learning models for predicting hospitalization and mortality risks of COVID-19 patients. Expert Systems with Applications, 240, 122670. https://doi.org/10.1016/J.ESWA.2023.122670

Ding, N., Qin, Y., Yang, G., Wei, F., Yang, Z., Su, Y., Hu, S., Chen, Y., Chan, C. M., Chen, W., Yi, J., Zhao, W., Wang, X., Liu, Z., Zheng, H. T., Chen, J., Liu, Y., Tang, J., Li, J., & Sun, M. (2023). Parameter-efficient fine-tuning of large-scale pre-trained language models. Nature Machine Intelligence, 5(3), 220–235. https://doi.org/10.1038/s42256-023-00626-4

Elitzur, R., Katz, N., Muttath, P., & Soberman, D. (2023). Machine Learning Methods and the Analysis of Threshold and Goldilocks Effects in Reward-Based Crowdfunding. https://doi.org/10.2139/SSRN.4674493

Elitzur, R., Muttath, P., & Soberman, D. (2024). Crowdfunding and too much choice: A recipe for disappointment. Journal of Business Venturing Insights, 21, e00436. https://doi.org/10.1016/J.JBVI.2023.E00436

Elrashidy, Z., Haniffa, R., Sherif, M., & Baroudi, S. (2024). Determinants of reward crowdfunding success: Evidence from Covid-19 pandemic. Technovation, 132, 102985. https://doi.org/10.1016/J.TECHNOVATION.2024.102985

Etter, V., Grossglauser, M., & Thiran, P. (2013). Launch hard or go home!: predicting the success of kickstarter campaigns. Conference on Online Social Networks. https://doi.org/10.1145/2512938.2512957

Falcão, G., Fernandes, A. M., Garcia, N., Aidos, H., & Tomás, P. (2023). Stacking deep learning models for early detection of wildfire smoke plumes. European Signal Processing Conference, 1370–1374. https://doi.org/10.23919/EUSIPCO58844.2023.10289811

Falkner, S., Klein, A., & Hutter, F. (2018). BOHB: Robust and Efficient Hyperparameter Optimization at Scale. 35th

International Conference on Machine Learning, ICML 2018, 4, 2323–2341.

Fernandez-Blanco, A., Villanueva-Balsera, J. M., Rodríguez-Montequín, V., & Morán-Palacios, H. (2020). Key Factors for Project Crowdfunding Success: An Empirical Study. Sustainability. https://doi.org/10.3390/SU12020599

Frydrych, D., Bock, A. J., Kinder, T., & Koeck, B. (2014). Exploring entrepreneurial legitimacy in reward-based crowdfunding. Venture Capital: An International Journal of Entrepreneurial Finance. https://doi.org/10.1080/13691066.2014.916512

Genc, E., Abdelgawad, L., Morari, V., & Kluegl, P. (2020). Convolutional neural networks for classification of German blurbs. Proceedings of the 15th Conference on Natural Language Processing, KONVENS 2019, 323 – 327. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85103443794&partnerID=40&md5=6c289d3c5ac1ccc112787d25d5fef60f

Ghaffarian, S., van der Voort, M., Valente, J., Tekinerdogan, B., & de Mey, Y. (2022). Machine learning-based farm risk management: A systematic mapping review. Computers and Electronics in Agriculture, 192, 106631. https://doi.org/10.1016/J.COMPAG.2021.106631

Giro555. (n.d.). Https://Giro555.Nl.

Goay, C. H., Ahmad, N. S., & Goh, P. (2021). Transient Simulations of High-Speed Channels Using CNN-LSTM with an Adaptive Successive Halving Algorithm for Automated Hyperparameter Optimizations. IEEE Access, 9, 127644–127663. https://doi.org/10.1109/ACCESS.2021.3112134

Gong, J., Pavlou, P. A., & Zheng, Z. (Alvin). (2021). On the Use of Probabilistic Uncertain Rewards on Crowdfunding Platforms: The Case of the Lottery. Information Systems Research. https://doi.org/10.1287/ISRE.2020.0963

Gove, R., & Faytong, J. (2012). Machine Learning and Event-Based Software Testing: Classifiers for Identifying Infeasible GUI Event Sequences. Advances in Computers, 86, 109–135. https://doi.org/10.1016/B978-0-12-396535-6.00004-1

Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Vaughan, A., Yang, A., Fan, A., Goyal, A., Hartshorn, A., Yang, A., Mitra, A., Sravankumar, A., Korenev, A., Hinsvark, A., ... Ma, Z. (2024). The Llama 3 Herd of Models. http://arxiv.org/abs/2407.21783

Greenberg, M., Pardo, B., Hariharan, K., & Gerber, E. (2013). Crowdfunding support tools: predicting success & failure. CHI Extended Abstracts. https://doi.org/10.1145/2468356.2468682

Gündüz, H. (2024). Comparative analysis of BERT and FastText representations on crowdfunding campaign success prediction. PeerJ Computer Science. https://doi.org/10.7717/PEERJ-CS.2316

Guo, J., & Li, X. (2019). Prediction of index trend based on LSTM model for extracting image similarity feature. ACM International Conference Proceeding Series, 335–340. https://doi.org/10.1145/3349341.3349427

Guo, Q., Zhou, S., Wu, Y., Peng, Z., & Ma, X. (2022). Understanding and Modeling Viewers' First Impressions with Images in Online Medical Crowdfunding Campaigns. International Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3491102.3501830

Guo, Y., Zhou, X., Zhan, C., Zeng, Y., & Zhong, L. (2021). Prediction and Analysis of Success on Crowdfunding Projects. Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering, 785–789. https://doi.org/10.1145/3443467.3443854

Haasbroek, F., & Ungerer, M. (2020). Theory versus practise: Assessing reward-based crowdfunding theory through a South African case study. https://doi.org/10.4102/SAJBM.V51I1.2078

Haitham, H., Amgad, N., Tarek, H., & Solayman, M. (2024). From Data to Decisions: Enhancing Crowdfunding Strategies AI-Driven and Business Intelligence Techniques. 2024 Intelligent Methods, Systems, and Applications (IMSA), 475–481. https://doi.org/10.1109/IMSA61967.2024.10652782

Han, S., Eom, H., Kim, J., & Park, C. (2020). Optimal DNN architecture search using Bayesian Optimization Hyperband for arrhythmia detection. IEEE.

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3–12. https://doi.org/10.1002/ASMB.2209

Hoegen, A., Steininger, D. M., & Veit, D. (2018). How do investors decide? An interdisciplinary review of decision-making in crowdfunding. Electronic Markets. https://doi.org/10.1007/S12525-017-0269-Y

Hoque, Md. M. (2024). Crowdfunding for innovation: a comprehensive empirical review. Future Business Journal. https://doi.org/10.1186/S43093-024-00387-5

Hörisch, J., & Tenner, I. (2020). How environmental and social orientations influence the funding success of investment-based crowdfunding: The mediating role of the number of funders and the average funding amount. Technological Forecasting and Social Change. https://doi.org/10.1016/J.TECHFORE.2020.120311

Hossain, M., & Oparaocha, G. O. (2017). Crowdfunding: Motives, Definitions, Typology and Ethical Challenges. 7(2). https://doi.org/doi:10.1515/erj-2015-0045

How, D. N. T., Hannan, M. A., Lipu, M. S. H., Ker, P. J., Mansor, M., Sahari, K. S. M., & Muttaqi, K. M. (2022). SOC Estimation Using Deep Bidirectional Gated Recurrent Units with Tree Parzen Estimator Hyperparameter Optimization. IEEE Transactions on Industry Applications, 58(5), 6629–6638. https://doi.org/10.1109/TIA.2022.3180282

Hu, Y. H., Lo, C. L., & Shih, S. P. (2014). Developing early warning systems to predict students' online learning performance. Computers in Human Behavior, 36, 469–478. https://doi.org/10.1016/J.CHB.2014.04.002

Hui, J., Greenberg, M., & Gerber, E. (2014). Understanding the role of community in crowdfunding work. CSCW. https://doi.org/10.1145/2531602.2531715

Hussain, N., Pietro, F. Di, & Rosati, P. (2023). Crowdfunding for Social Entrepreneurship: A Systematic Review of the Literature. Journal of Social Entrepreneurship. https://doi.org/10.1080/19420676.2023.2236637

Illustrated Guide to LSTM's and GRU's: A step by step explanation | by Michael Phi | Towards Data Science. (n.d.). Retrieved April 1, 2024, from https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21

Im, C., Lee, J., Kim, D., Jun, S. C., & Seo, H. (2025). Seeking Optimal Montage for Single-Pair Transcranial Direct Current Stimulation Using Bayesian Optimization and Hyperband—A Feasibility Study. Neuromodulation, 28(1), 86–94. https://doi.org/10.1016/j.neurom.2024.09.475

Inman, J., Khandait, T., Pedrielli, G., & Sankar, L. (2023). Parameter Optimization with Conscious Allocation (POCA). Proceedings - Winter Simulation Conference, 3436–3447. https://doi.org/10.1109/WSC60868.2023.10407962

Janku, J., Kucerová, Z., & Darena, F. (2023). Behavioral Insights from Crowdfunding Financing: What Do Social Media Ties, Emotional Cues and Sentiment Tell Us? Ekonomicky Casopis. https://doi.org/10.31577/EKONCAS.2023.02.01

John Tukey. (1977). Exploratory Data Analysis (Frederick Mosteller, Ed.). Addison-Wesley.

Joseph, L. P., Deo, R. C., Casillas-Pérez, D., Prasad, R., Raj, N., & Salcedo-Sanz, S. (2024). Short-term wind speed forecasting using an optimized three-phase convolutional neural network fused with bidirectional long short-term memory network model. Applied Energy, 359. https://doi.org/10.1016/j.apenergy.2024.122624

Kankanamge, N., Yigitcanlar, T., Goonetilleke, A., & Kamruzzaman, M. (2019). Can volunteer crowdsourcing reduce disaster risk? A systematic review of the literature. International Journal of Disaster Risk Reduction, 35, 101097. https://doi.org/10.1016/J.IJDRR.2019.101097

Kavzoglu, T., & Teke, A. (2022). Advanced hyperparameter optimization for improved spatial prediction of shallow landslides using extreme gradient boosting (XGBoost). Bulletin of Engineering Geology and the Environment, 81(5). https://doi.org/10.1007/s10064-022-02708-w

Keisar, B., & Lev, O. (2023). Follow the money: Techniques for pricing rewards in crowdfunding projects. Electronic Commerce Research and Applications, 62, 101324. https://doi.org/https://doi.org/10.1016/j.elerap.2023.101324

Khemani, B., Patil, S., Kotecha, K., & Tanwar, S. (2024). A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. Journal of Big Data, 11(1), 1–43. https://doi.org/10.1186/S40537-023-00876-4/TABLES/13

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. Multimedia Tools and Applications, 82(3), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

Kim, S., Park, J., Chung, W., Adams, D., & Lee, J. H. (2024). Techno-economic analysis for design and management of international green hydrogen supply chain under uncertainty: An integrated temporal planning approach. Energy Conversion and Management, 301. https://doi.org/10.1016/j.enconman.2023.118010

Kindler, A., Golosovsky, M., & Solomon, S. (2019). Early prediction of the outcome of Kickstarter campaigns: is the success due to virality? Palgrave Communications. https://doi.org/10.1057/S41599-019-0261-6

Klein, A., Falkner, S., Bartels, S., Hennig, P., & Hutter, F. (2017). Fast bayesian hyperparameter optimization on large datasets. Electronic Journal of Statistics, 11(2), 4945–4968. https://doi.org/10.1214/17-EJS1335SI

Klein, A., Falkner, S., Springenberg, J. T., & Hutter, F. (2017). Learning Curve Prediction with Bayesian Neural Networks. International Conference on Learning Representations.

Korzynski, P., Haenlein, M., & Rautiainen, M. (2021). Impression management techniques in crowdfunding: An analysis of Kickstarter videos using artificial intelligence. European Management Journal, 39(5), 675–684. https://doi.org/10.1016/J.EMJ.2021.01.001

Kraus, S., Richter, C., Brem, A., Cheng, C.-F., & Chang, M.-L. (2016). Strategies for reward-based crowdfunding campaigns. Journal of Innovation & Knowledge, 1(1), 13–23. https://doi.org/https://doi.org/10.1016/j.jik.2016.01.010

Kuchera, B. (2015). No one is getting rich from Exploding Kittens' $8.7 million Kickstarter - Polygon. https://www.polygon.com/2015/2/25/8102751/exploding-kittens-kickstarter-rich

Kumar, S., & Pandey, M. (2015). The Impact of 9-Ending Pricing Strategy on the Consumers' Attitude and Buying Behavior: An Exploration. 2(2), 93–98. http://www.krishisanskriti.org/aebm.html

Kunz, M. M., Bretschneider, U., Erler, M., & Leimeister, J. M. (2017). An empirical investigation of signaling in reward-based crowdfunding. Electronic Commerce Research, 17(3), 425–461. https://doi.org/10.1007/s10660-016-9249-0

Kuppuswamy, V., & Bayus, B. (2018). Crowdfunding Creative Ideas: The Dynamics of Project Backers in Kickstarter. https://doi.org/10.2139/SSRN.2234765

Kuppuswamy, V., & Bayus, B. L. (2017). Does my contribution to your crowdfunding project matter. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2016.10.004

Le, N. T. K., Hadiprodjo, N., El-Alfy, H., Kerimzhanov, A., & Teshebaev, A. (2023). The Recent Large Language Models in NLP. 22nd International Symposium on Communications and Information Technologies, ISCIT 2023, 93–98. https://doi.org/10.1109/ISCIT57293.2023.10376050

Learn About Indiegogo | Indiegogo. (n.d.). Retrieved October 9, 2024, from https://www.indiegogo.com/about/our-story

Lee, S., Lee, K., & Kim, H. (2018). Content-based Success Prediction of Crowdfunding Campaigns: A Deep Learning Approach. CSCW Companion. https://doi.org/10.1145/3272973.3274053

Lei, S., Xu, K., Huang, Y., & Sha, X. (2020). An Xgboost based system for financial fraud detection. E3S Web of Conferences, 214, 02042. https://doi.org/10.1051/E3SCONF/202021402042

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., Zettlemoyer, L., & Ai, F. (n.d.). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. https://huggingface.co/transformers

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2017). Hyperband: Bandit-based configuration evaluation for hyperparameter optimization. 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings.

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2018). Hyperband: A novel bandit-based approach to hyperparameter optimization. Journal of Machine Learning Research, 18, 1–52.

Li, L., Yang, L., Zhao, M., Liao, M., & Cao, Y. (2022). Exploring the success determinants of crowdfunding for cultural and creative projects: An empirical study based on signal theory. Technology in Society, 70. https://doi.org/10.1016/j.techsoc.2022.102036

Li, Q., & Wang, N. (2024). Fundraiser engagement, third-party endorsement and crowdfunding performance: A configurational theory approach. PLoS ONE. https://doi.org/10.1371/JOURNAL.PONE.0308717

Li, X., Yan, D., Wang, K., Weng, B., Qin, T., & Liu, S. (2019). Flood Risk Assessment of Global Watersheds Based on Multiple Machine Learning Models. Water 2019, Vol. 11, Page 1654, 11(8), 1654. https://doi.org/10.3390/W11081654

Li, Y., Rakesh, V., & Reddy, C. K. (2016). Project Success Prediction in Crowdfunding Environments. Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, 247–256. https://doi.org/10.1145/2835776.2835791

Li, Y., Shen, Y., Jiang, H., Zhang, W., Li, J., Liu, J., Zhang, C., & Cui, B. (2022). Hyper-Tune: Towards Efficient Hyper-parameter Tuning at Scale. Contemporary Mathematics, 15(6), 1256–1265. https://doi.org/10.14778/3514061.3514071

Li, Y., Shen, Y., Jiang, J., Gao, J., Zhang, C., & Cui, B. (2021). MFES-HB: Efficient Hyperband with Multi-Fidelity Quality Measurements. 35th AAAI Conference on Artificial Intelligence, AAAI 2021, 10A, 8491–8500. https://doi.

org/10.1609/aaai.v35i10.17031

Li, Y., Wang, S., Ding, H., & Chen, H. (2023). Large Language Models in Finance: A Survey. ICAIF 2023 - 4th ACM International Conference on AI in Finance, 374–382. https://doi.org/10.1145/3604237.3626869

Liang, W., Luo, S., Zhao, G., & Wu, H. (2020). Predicting Hard Rock Pillar Stability Using GBDT, XGBoost, and LightGBM Algorithms. Mathematics 2020, Vol. 8, Page 765, 8(5), 765. https://doi.org/10.3390/MATH8050765

Lin, L., Ma, X., Chen, C., Xu, J., & Huang, N. (2024). Imbalanced Industrial Load Identification Based on Optimized CatBoost with Entropy Features. Journal of Electrical Engineering and Technology, 19(8), 4817–4832. https://doi.org/10.1007/s42835-024-01933-5

Lin, Y., Lee, W.-C., & Chang, C.-C. H. (2016). Analysis of rewards on reward-based crowdfunding platforms. 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 501–504. https://doi.org/10.1109/ASONAM.2016.7752281

Ling, Y., Huang, T., Yue, Q., Shan, Q., Hei, D., Zhang, X., Shi, C., & Jia, W. (2022). Improving the estimation accuracy of multi-nuclide source term estimation method for severe nuclear accidents using temporal convolutional network optimized by Bayesian optimization and hyperband. Journal of Environmental Radioactivity, 242. https://doi.org/10.1016/j.jenvrad.2021.106787

Liu, C., Yang, S., Hao, T., & Song, R. (2022). Service risk of energy industry international trade supply chain based on artificial intelligence algorithm. Energy Reports, 8, 13211–13219. https://doi.org/10.1016/J.EGYR.2022.09.182

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM Computing Surveys, 55(9). https://doi.org/10.1145/3560815

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11–26. https://doi.org/10.1016/J.NEUCOM.2016.12.038

Llama-3.2-3B-Instruct. (n.d.). Https://Huggingface.Co/Meta-Llama/Llama-3.2-3B-Instruct.

Lukkarinen, A., Teich, J. E., Wallenius, H., & Wallenius, J. (2016). Success drivers of online equity crowdfunding campaigns. Decision Support Systems. https://doi.org/10.1016/J.DSS.2016.04.006

Ma, Z., Wang, X., & Hao, Y. (2023). Development and application of a hybrid forecasting framework based on improved extreme learning machine for enterprise financing risk. Expert Systems with Applications, 215, 119373. https://doi.org/10.1016/J.ESWA.2022.119373

Maiolini, R., Franco, S., Cappa, F., & Hayes, D. (2023). Optimizing Reward-Based Crowdfunding. IEEE Engineering Management Review. https://doi.org/10.1109/EMR.2023.3243148

Marinova, T. (2019). Virtual Community Characteristics as Success Factors for Crowdfunding Projects.

Marneni, D., & Vemula, S. (2022). Analysis of Covid-19 using machine learning techniques. Statistical Modeling in Machine Learning: Concepts and Applications, 37–53. https://doi.org/10.1016/B978-0-323-91776-6.00006-3

Martínez-Cháfer, L., Molina-Morales, F., & Peiró-Palomino, J. (2021). On the drivers of successful crowdfunding: The case of the platform Verkami. Business Research Quarterly. https://doi.org/10.1177/2340944420985862

Masrour, T., Cherrafi, A., & El Hassani, I. (Eds.). (2021). Artificial Intelligence and Industrial Applications. 1193. https://doi.org/10.1007/978-3-030-51186-9

McCarthy, P. X., Gong, X., Braesemann, F., Stephany, F., Rizoiu, M. A., & Kern, M. L. (2023). The impact of founder personalities on startup success. Scientific Reports 2023 13:1, 13(1), 1–12. https://doi.org/10.1038/s41598-023-41980-y

Mckinney, W. (2010). Data Structures for Statistical Computing in Python.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 5(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Min, B., Ross, H., Sulem, E., Veyseh, A. P. Ben, Nguyen, T. H., Sainz, O., Agirre, E., Heintz, I., & Roth, D. (2023). Recent Advances in Natural Language Processing via Large Pre-trained Language Models: A Survey. ACM Computing Surveys, 56(2). https://doi.org/10.1145/3605943

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large Language Models: A Survey. http://arxiv.org/abs/2402.06196

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey Wolf Optimizer. Advances in Engineering Software, 69, 46–61. https://doi.org/10.1016/J.ADVENGSOFT.2013.12.007

Mitra, T., & Gilbert, E. (2014). The language that gets people to give: phrases that predict success on kickstarter. CSCW. https://doi.org/10.1145/2531602.2531656

Mittal, G., Liu, C., Karianakis, N., Fragoso, V., Chen, M., & Fu, Y. (2020). HyperStar: Task-aware hyperparameters for deep networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 8733–8742. https://doi.org/10.1109/CVPR42600.2020.00876

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2013.06.005

Moy, N., Chan, H. F., Septianto, F., Mathmann, F., & Torgler, B. (2024). Confidence is Good? too Much, not so Much: Exploring the effects on crowdfunding success. Journal of Business Research, 182, 114711. https://doi.org/10.1016/J.JBUSRES.2024.114711

MTEB. (n.d.). Https://Huggingface.Co/Spaces/Mteb/Leaderboard.

Murugan, M. S., & T, S. K. (2023). Large-scale data-driven financial risk management & analysis using machine learning strategies. Measurement: Sensors, 27, 100756. https://doi.org/10.1016/J.MEASEN.2023.100756

Neuhaus, J., Isaak, A. J., & Bostandzic, D. (2021). Million dollar personality: a systematic literature review on personality in crowdfunding. Management Review Quarterly. https://doi.org/10.1007/S11301-021-00242-9

Nguyen, X. D. J., & Liu, Y. A. (2025). Methodology for hyperparameter tuning of deep neural networks for efficient and accurate molecular property prediction. Computers and Chemical Engineering, 193. https://doi.org/10.1016/j.compchemeng.2024.108928

Niyogisubizo, J., Liao, L., Nziyumva, E., Murwanashyaka, E., & Nshimyumukiza, P. C. (2022). Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization. Computers and Education: Artificial Intelligence, 3, 100066. https://doi.org/10.1016/J.CAEAI.2022.100066

Oduro, M. S., Yu, H., & Huang, H. (2022). Predicting the Entrepreneurial Success of Crowdfunding Campaigns Using Model-Based Machine Learning Methods. International Journal of Crowd Science. https://doi.org/10.26599/IJCS.2022.9100003

Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language

Processing. IEEE Transactions on Neural Networks and Learning Systems, 32(2), 604–624. https://doi.org/10.1109/TNNLS.2020.2979670

Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2016.11.001

Passos, D., & Mishra, P. (2021). An automated deep learning pipeline based on advanced optimisations for leveraging spectral classification modelling. Chemometrics and Intelligent Laboratory Systems, 215. https://doi.org/10.1016/j.chemolab.2021.104354

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. http://arxiv.org/abs/1912.01703

Pati, R., & Garud, N. (2021). Social Interaction and Crowdfunding Project Success: Moderating Roles of Product Development Stage and Product Innovativeness. IEEE Transactions on Engineering Management. https://doi.org/10.1109/TEM.2021.3061532

Peng, L., Cui, G., Bao, Z., & Liu, S. (2022). Speaking the same language: the power of words in crowdfunding success and failure. Marketing Letters, 33(2), 311–323. https://doi.org/10.1007/s11002-021-09595-3

Pham, B. T., Luu, C., Phong, T. Van, Nguyen, H. D., Le, H. Van, Tran, T. Q., Ta, H. T., & Prakash, I. (2021). Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam. Journal of Hydrology, 592, 125815. https://doi.org/10.1016/J.JHYDROL.2020.125815

Pinkow, F., & Emmerich, P. (2021). Re-Examining Crowdfunding Success: How the Crowdfunding Goal Moderates the Relationship of Success Factors and Crowdfunding Performance. Central European Business Review. https://doi.org/10.18267/J.CEBR.263

Popescul, D., Radu, L.-D., Pavaloaia, V., & Georgescu, M. (2020). Psychological Determinants of Investor Motivation in Social Media-Based Crowdfunding Projects: A Systematic Review. Frontiers in Psychology. https://doi.org/10.3389/FPSYG.2020.588121

Rama, A., Jiang, C., Johan, S., Liu, H., Mai, Y., Rama, A., Jiang, C., Johan, S., Liu, H., & Mai, Y. (2022). Religious and social narratives and crowdfunding success. https://doi.org/10.1016/J.INTFIN.2022.101595

Rawte, V., Sheth, A., & Das, A. (2023). A Survey of Hallucination in Large Foundation Models. http://arxiv.org/abs/2309.05922

Regner, T., & Crosetto, P. (2021). The experience matters: participation-related rewards increase the success chances of crowdfunding campaigns. Economics of Innovation and New Technology. https://doi.org/10.1080/10438599.2020.1792606

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. http://arxiv.org/abs/1908.10084

Repede, S. E., & Brad, R. (2024). LLaMA 3 vs. State-of-the-Art Large Language Models: Performance in Detecting Nuanced Fake News. Computers, 13(11). https://doi.org/10.3390/computers13110292

Rose, S., Wentzel, D., Hopp, C., & Kaminski, J. (2021). Launching for success: The effects of psychological distance and mental simulation on funding decisions and crowdfunding performance. Journal of Business Venturing. https://doi.org/10.1016/J.JBUSVENT.2020.106021

Rubio-Martín, S., García-Ordás, M. T., Bayón-Gutiérrez, M., Prieto-Fernández, N., & Benítez-Andrades, J. A. (2024). Enhancing ASD detection accuracy: a combined approach of machine learning and deep learning models with natural language processing. Health Information Science and Systems, 12(1), 1–12. https://doi.org/10.1007/S13755-024-00281-Y/FIGURES/5

Ryoba, M. J., Qu, S., & Zhou, Y. (2021). Feature subset selection for predicting the success of crowdfunding project campaigns. Electronic Markets. https://doi.org/10.1007/S12525-020-00398-4

Saluja, O. B. (2024). Crowd-funding to promote women entrepreneurship: a systematic literature review of the empirical studies conducted in the last decade (2012–2023). International Journal of Gender and Entrepreneurship. https://doi.org/10.1108/IJGE-09-2023-0225

Samaddar, M., Roy, R., De, S., & Karmakar, R. (2021). A comparative study of different machine learning algorithms on Bitcoin value prediction. Proceedings of the 2021 1st International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies, ICAECT 2021. https://doi.org/10.1109/ICAECT49130.2021.9392629

Sani, U. S., Lai, D. T. C., & Malik, O. A. (2021). Investigating Automated Hyper-Parameter Optimization for a Generalized Path Loss Model. Frontiers in Artificial Intelligence and Applications, 345, 283–291. https://doi.org/10.3233/FAIA210413

Sauermann, H., Franzoni, C., & Shafi, K. (2019). Crowdfunding scientific research: Descriptive insights and correlates of funding success. PloS One. https://doi.org/10.1371/JOURNAL.PONE.0208384

ScienceDirect.com | Science, health and medical journals, full text articles and books. (n.d.). Retrieved December 10, 2024, from https://www.sciencedirect.com/

Scopus search | Elsevier. (n.d.). Retrieved December 10, 2024, from https://www.elsevier.com/products/scopus/search

Sendra-Pons, P., Garzon, D., & Revilla-Camacho, M. (2024). Untangling signalling strategies contributing to overfunding in reward-based crowdfunding. European Research on Management and Business Economics. https://doi.org/10.1016/J.IEDEEN.2024.100251

Sewaid, A., García-Cestona, M., & Silaghi, F. (2021). Resolving information asymmetries in financing new product development: The case of reward-based crowdfunding. https://doi.org/10.1016/J.RESPOL.2021.104345

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & Freitas, N. de. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE, 104(1), 148–175. https://doi.org/10.1109/JPROC.2015.2494218

Sheth, V., Tripathi, U., & Sharma, A. (2022). A Comparative Analysis of Machine Learning Algorithms for Classification Purpose. Procedia Computer Science, 215, 422–431. https://doi.org/10.1016/J.PROCS.2022.12.044

Simons, A., Weinmann, M., Tietz, M., & Brocke, J. (2017). Which Reward Should I Choose? Preliminary Evidence for the Middle-Option Bias in Reward-Based Crowdfunding. HICSS. https://doi.org/10.24251/HICSS.2017.526

Soni, A., Rasool, A., Dubey, A., & Khare, N. (2022). Data Mining based Dimensionality Reduction Techniques. 2022 International Conference for Advancement in Technology, ICONAT 2022. https://doi.org/10.1109/ICONAT53423.2022.9725846

Sun, S., & Liu, H. (2013). Particle Swarm Algorithm: Convergence and Applications. Swarm Intelligence and Bio-Inspired Computation: Theory and Applications, 137–168. https://doi.org/10.1016/B978-0-12-405163-8.00006-5

Swaminatha Rao, L. P., & Jaganathan, S. (2024). Adaptive Bayesian contextual hyperband: A novel hyperparameter optimization approach. IAES International Journal of Artificial Intelligence, 13(1), 775–785. https://doi.org/10.11591/ijai.v13.i1.pp775-785

Tajvarpour, M. H., & Pujari, D. (2022). The influence of narrative description on the success of crowdfunding campaigns: The moderating role of quality signals. Journal of Business Research, 149, 123–138. https://doi.org/10.1016/J.JBUSRES.2022.05.012

Tang, X., Yao., X., Dai., R., & Wang, Q. (2024). Does green matter for crowdfunding? International evidence. Journal of International Financial Markets, Institutions, and Money. https://doi.org/10.1016/J.INTFIN.2024.101950

Tang, Z., Yang, Y., Li, W., Lian, D., & Duan, L. (2023). Deep Cross-Attention Network for Crowdfunding Success Prediction. IEEE Transactions on Multimedia, 25, 1306–1319. https://doi.org/10.1109/TMM.2022.3141256

The Ocean Cleanup. (n.d.). Https://Theoceancleanup.Com/Help-Clean/.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models. http://arxiv.org/abs/2302.13971

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Ferrer, C. C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., … Scialom, T. (2023). Llama 2: Open Foundation and Fine-Tuned Chat Models. http://arxiv.org/abs/2307.09288

Tran, T., Dontham, M., Chung, J., & Lee, K. (2016). How to Succeed in Crowdfunding: a Long-Term Study in Kickstarter. https://doi.org/10.48550/arXiv.1607.06839

Troise, C., Battisti, E., Christofi, M., van Vulpen, N. J., & Tarba, S. (2023). How Can SMEs Use Crowdfunding Platforms to Internationalize? The Role of Equity and Reward Crowdfunding. Management International Review, 63(1), 117–159. https://doi.org/10.1007/s11575-022-00493-y

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. http://arxiv.org/abs/1706.03762

Vu, A. N., & Christian, J. (2024). UK Equity Crowdfunding Success: The Impact of Competition, Brexit and Covid-19. British Journal of Management, 35(1), 321–344. https://doi.org/https://doi.org/10.1111/1467-8551.12714

Wang, J., Luo, J., & Zhang, X. (2022). How COVID-19 Has Changed Crowdfunding: Evidence From GoFundMe. Frontiers in Computer Science, 4. https://doi.org/10.3389/fcomp.2022.893338

Wang, N., Li, Q., Liang, H., Ye, T., & Ge, S. (2018). Understanding the importance of interaction between creators and backers in crowdfunding success. Electronic Commerce Research and Applications. https://doi.org/10.1016/J.ELERAP.2017.12.004

Wang, W., He, L., Wu, Y. J., & Goh, M. (2021). Signaling persuasion in crowdfunding entrepreneurial narratives: The subjectivity vs objectivity debate. Computers in Human Behavior, 114, 106576. https://doi.org/https://doi.org/10.1016/j.chb.2020.106576

Wang, W., Liu, H., & Wu, Y. (2024). How do reward personalization options influence the public's willingness to participate in innovation projects? Insights from crowdfunding in Industry 5.0. European Journal of Innovation Management. https://doi.org/10.1108/EJIM-07-2023-0584

Wang, W., Zheng, H., & Wu, Y. J. (2020). Prediction of fundraising outcomes for crowdfunding projects based on deep learning: a multimodel comparative study. Soft Computing, 24(11), 8323–8341. https://doi.org/10.1007/s00500-020-04822-x

Wang, W., Zhu, K., Wang, H., & Wu, Y. (2017). The Impact of Sentiment Orientations on Successful Crowdfunding Campaigns through Text Analytics. IET Softw. https://doi.org/10.1049/IET-SEN.2016.0295

Wang, Y., Wang, B., & Yan, Y. (2022). Does network externality affect your project? Evidences from reward-based technology crowdfunding. Technological Forecasting & Social Change. https://doi.org/10.1016/J.TECHFORE.2022.121667

Wang, Y., Wang, B., Yan, Y., & Zhang, J. (2021). The Clearer Timelines, the Better? Evidence From Kickstarter. IEEE Transactions on Engineering Management. https://doi.org/10.1109/TEM.2021.3099686

Wang, Z., Yang, L., & Hahn, J. (2022). Winner Takes All? The Blockbuster Effect on Crowdfunding Platforms. Information Systems Research. https://doi.org/10.1287/ISRE.2022.1154

Wangchuk, P. (2021). Common types of Crowdfunding Models, Related Concepts and Its Impact on Business: A Brief Literature Review. Asian Journal of Economics, Business and Accounting. https://doi.org/10.9734/AJEBA/2021/V21I1430471

Web Robots IO. (2025, March 12). Https://Webrobots.Io/Kickstarter-Datasets/.

Weber, C. E., Steigenberger, N., & Wilhelm, H. (2023). After successful fundraising: how overfunding and category spanning affect the release and audience-perceived quality of crowdfunded products. Small Business Economics. https://doi.org/10.1007/S11187-022-00721-7

Wessel, M., Adam, M., & Benlian, A. (2019). The impact of sold-out early birds on option selection in reward-based crowdfunding. Decision Support Systems, 117, 48–61. https://doi.org/https://doi.org/10.1016/j.dss.2018.12.002

Willette, A. A., Willette, S. A., Wang, Q., Pappas, C., Klinedinst, B. S., Le, S., Larsen, B., Pollpeter, A., Li, T., Mochel, J. P., Allenspach, K., Brenner, N., & Waterboer, T. (2022). Using machine learning to predict COVID-19 infection and severity risk among 4510 aged adults: a UK Biobank cohort study. Scientific Reports 2022 12:1, 12(1), 1–11. https://doi.org/10.1038/s41598-022-07307-z

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., … Rush, A. M. (2020). Transformers: State-of-the-Art Natural Language Processing. https://github.com/huggingface/

Xiao, S., & Yue, Q. (2018). Investors' inertia behavior and their repeated decision-making in online reward-based crowdfunding market. Decision Support Systems, 111, 101–112. https://doi.org/https://doi.org/10.1016/j.dss.2018.05.005

Xiong, Y., Ma, Y., Ruan, L., Li, D., Lu, C., & Huang, L. (2022). Comparing different machine learning techniques for predicting COVID-19 severity. Infectious Diseases of Poverty, 11(1), 1–9. https://doi.org/10.1186/S40249-022-00946-4/FIGURES/4

Yaloveha, V., Podorozhniak, A., & Kuchuk, H. (2022). CONVOLUTIONAL NEURAL NETWORK HYPERPARAMETER OPTIMIZATION APPLIED TO LAND COVER CLASSIFICATION. Radioelectronic and Computer Systems, 2022(1), 115–128. https://doi.org/10.32620/reks.2022.1.09

Yang, J., Kim, J., Ryu, H., Lee, J., & Park, C. (2024). Predicting Car Rental Prices: A Comparative Analysis of Machine Learning Models. Electronics (Switzerland), 13(12). https://doi.org/10.3390/electronics13122345

Yang, K. F., Lin, Y. R., & Chen, L. S. (2023). Discovering Critical Factors in the Content of Crowdfunding Projects.

Algorithms, 16(1). https://doi.org/10.3390/a16010051

Yang, K.-F., Lin, Y.-R., & Chen, L.-S. (2023). Discovering Critical Factors in the Content of Crowdfunding Projects. Algorithms. https://doi.org/10.3390/A16010051

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295–316. https://doi.org/https://doi.org/10.1016/j.neucom.2020.07.061

Yang, L., Wang, Z., & Hahn, J. (2020). Scarcity Strategy in Crowdfunding: An Empirical Exploration of Reward Limits. Information Systems Research. https://doi.org/10.1287/ISRE.2020.0934

Yang, X., Ma, H., & Wang, M. (2023). Research on Rumor Detection Based on a Graph Attention Network With Temporal Features. Https://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/IJDWM.319342, 19(2), 1–17. https://doi.org/10.4018/IJDWM.319342

Yang, X.-S. (2021). Genetic Algorithms. Nature-Inspired Optimization Algorithms, 91–100. https://doi.org/10.1016/B978-0-12-821986-7.00013-5

Yasar, B., Sevilay Yılmaz, I., Hatipoglu, N., & Salih, A. (2022). Stretching the success in reward-based crowdfunding. Journal of Business Research, 152, 205–220. https://doi.org/10.1016/J.JBUSRES.2022.07.053

Yosipof, A., Drori, N., Elroy, O., & Pierraki, Y. (2024). Textual sentiment analysis and description characteristics in crowdfunding success: The case of cybersecurity and IoT industries. Electronic Markets. https://doi.org/10.1007/S12525-024-00712-4

Yu, P.-F., Huang, F.-M., Yang, C., Liu, Y.-H., Li, Z.-Y., & Tsai, C.-H. (2018). Prediction of Crowdfunding Project Success with Deep Learning. 2018 IEEE 15th International Conference on E-Business Engineering (ICEBE), 1–8. https://doi.org/10.1109/ICEBE.2018.00012

Zhai, Y., & Shen, W. (2024). Signaling theory in charity-based crowdfunding: Investigating the effects of project creator characteristics and text linguistic style on fundraising performance. Heliyon, 10(4), e25756. https://doi.org/10.1016/J.HELIYON.2024.E25756

Zhang, J., Wang, T., Wang, B., Chen, C., & Wang, G. (2023). Hyperparameter optimization method based on dynamic Bayesian with sliding balance mechanism in neural network for cloud computing. Journal of Cloud Computing, 12(1). https://doi.org/10.1186/s13677-023-00482-y

Zhang, Q., Ye, T., Essaidi, M., Agarwal, S., Liu, V., & Loo, B. T. (2017). Predicting Startup Crowdfunding Success through Longitudinal Social Engagement Analysis. CIKM. https://doi.org/10.1145/3132847.3132908

Zhang, X., Lyu, H., & Luo, J. (2021). What Contributes to a Crowdfunding Campaign's Success? Evidence and Analyses from GoFundMe Data. Journal of Social Computing. https://doi.org/10.23919/JSC.2021.0010

Zhang, Y., Li, H., Ge, L., Zheng, L., Tang, Z., & Zhao, F. (2024). A deep learning model optimized by Bayesian Optimization with Hyperband for fast prediction of the elastic properties of 3D tubular braided composites at different temperatures. Polymer Composites. https://doi.org/10.1002/pc.29171

Zhang, Y., Tong, J., Wang, Z., & Gao, F. (2020). Customer Transaction Fraud Detection Using Xgboost Model. Proceedings - 2020 International Conference on Computer Engineering and Application, ICCEA 2020, 554–558. https://doi.org/10.1109/ICCEA50009.2020.00122

Zhao, C., Zhang, Y., & Liu, B. (2019). Sales Effort Investment and the Success of Online Product Crowdfunding. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2908716

Zhao, Y., Harris, P., & Lam, W. (2019). Crowdfunding industry—History, development, policies, and potential issues. Journal of Public Affairs, 19(1), e1921. https://doi.org/https://doi.org/10.1002/pa.1921

Zhong, C. (2022). SUCCESS PREDICTION OF CROWDFUNDING CAMPAIGNS WITH PROJECT NETWORK: A MACHINE LEARNING APPROACH.

Zhou, M. J., Du, Q., Zhang, X., Qiao, Z., Wang, G., & Fan, W. (2015). Money talks: A predictive model on crowdfunding success using project description.

Zhou, M., Lu, B., Fan, W., & Wang, G. A. (2016). Project description and crowdfunding success: an exploratory study. Information Systems Frontiers. https://doi.org/10.1007/S10796-016-9723-1

Appendix A:
Overview of common ML methods

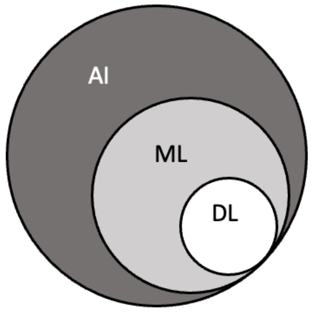
Artificial intelligence is commonly characterized by the capability of machines to exhibit cognitive functions akin to human intelligence. Furthermore, there is strived to reach artificial superintelligence, machines that can outperform humans. Often, machine learning is seen as the same as AI. However, AI is an overarching term that encompasses additional techniques and requirements besides ML. A full AI solution would have automated data identification, testing, and decision-making. ML involves manual data identification, testing by hand, and human decision-making. Because it is difficult to implement a full working system on AI, most focus in research is on machine learning (Aziz & Dowling, 2019).

Machine Learning

ML is thus a branch of AI, in which algorithms (computer programs) use data to automatically improve themselves through experience and learning. Therefore, it aids in detection, recognition, and prediction with the help of historical data. The performance is dependent on the quality and quantity of the data that is available, and the type of the algorithms in use. There are two types of data, labelled and unlabelled. The first has both input and output information, the latter only has output information (Ghaffarian et al., 2022).

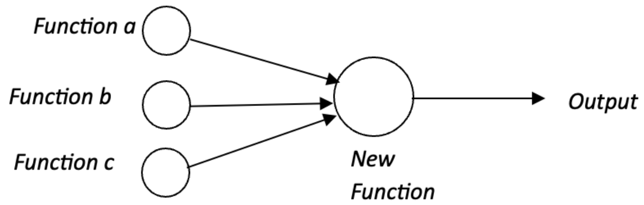
ML has two broader categories: supervised, and unsupervised machine learning. In supervised learning, labelled data is used. A function between the input and output is determined. It is similar to traditional statistics where a relationship is determined between certain variables. They are mainly used for regression and classification tasks. Unsupervised ML you analyse the data to learn more about the structure. This technique only uses unlabelled data. This approach is employed in the absence of comprehensive data knowledge and when the data do not exhibit a discernible pattern, facilitating analysis without manual intervention. This type of ML is used for address clustering, data reduction and anomaly detection tasks (Aziz & Dowling, 2019; Ghaffarian et al., 2022). Furthermore, there is a category that crosses supervised and unsupervised learning, namely deep learning (Aziz & Dowling, 2019). Lastly, there are the categories semi-supervised learning and reinforcement learning. The first needs a small portion of labelled input data and the rest can be unlabelled, this requires less manual work. The latter uses trial and error-based learning and has a feedback mechanism to update the previous status. This is mainly used in decision making (Ghaffarian et al., 2022).

In the following section different ML methods are described, as a supplement to the literature review.



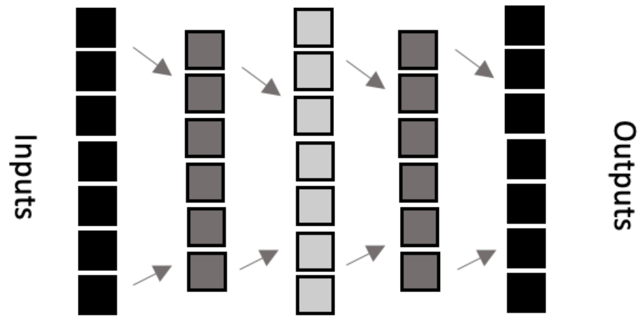
Deep Learning (DL)

LLaMA is a form of Deep Learning. In the Figure above it is depicted what the relation is between the different terms. DL can better mimic human decision-making, as it can model complex relationships more accurately (Aziz & Dowling, 2019). Within the field of Financial Engineering & Management, there are often large data sets with complex interactions, which cannot be specified in a full economic model. DL is a good method to employ in this field, since it can detect and exploit interactions in the data that are currently not visible to any existing financial economic theory (Heaton et al., 2017). In the last few years, the development of nonlinear theory and machine learning algorithms took place, and the academics realized that multivariate linear discriminant analyses could not match the complexity of enterprise financing risks and therefore artificial intelligence models were used to predict the risks (Ma et al., 2023a). DL is sometimes used interchangeably with Neural Networks (NN). This consists of many neurons: simple processors which are connected. They produce a sequence of activations, which happen due to sensors perceiving the environment, or weighted connections from previous neurons The first version was called Artificial Neural Networks (ANN), as it was supposed to mimic biological learning as of in the brain. They are seen as an adaptive information processing system. It is composed of many interconnected processing units, which have the attributes of nonlinearity, non-limitation, they are very qualitive and are non-convex (C. Liu et al., 2022). Deep Learning can be both supervised or unsupervised, and models complex relationships between variables and tries to mimic human decision-making. Since there are multiple 'hidden layers', there are many and combined influences between the input variables, and when they go through the model this process is repeated, and this way all these notes turn into newer factors. This is abstractedly depicted in the figure on the right, in reality all single blocks on the left are connected with every block on the right, repeated in other columns. While this mimics the human brain, it works as a 'black box', and therefore it is not always clear how the inputs are combined to create the output (Aziz & Dowling, 2019).



Types of NN
There are several types of neural networks developed throughout history, getting more and more complex. Here, the mostly used types in Financial Engineering & Management are described: a Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), autoencoder, and Graph Neural Network (GNN).
The first development which caused a chain of new algorithms used in deep learning, was a perceptron. This is one of the simplest, but most important milestones (Aggarwal & Murty, 2021).

Multilayer perceptron (MLP)
One layer of perceptrons cannot deal with classes that are not linear. This is resolved by having multiple layers, which each layer containing multiple perceptrons. This type of feedforward NN is called a Multilayer Perceptron (MLP). One of the advantages of this type of network is that it gives the possibility to learn the required representation from the input data. The learning of the model takes place by continuously replacing the weights assigned to the previous nodes. However, there are some disadvantages: the initial weights and the activation functions can have an influence on the outcome, if the training data is small and the number of hidden layers is large and overfitting can occur, when details of the



data is learned which negatively influences the outcome of the model. Furthermore, the model expects inputs of which the size is already determined, which limits its applications, and sometimes the data is connected to previous data (such as with sentence classifications, where words depend on each other), but the MLP cannot process this sequence data, as it assumes independence (Aggarwal & Murty, 2021).

Convolutional Neural Network (CNN)
CNNs are seen as the state-of-the-art tool for classification and prediction (Aggarwal & Murty, 2021), and the first real successful architecture that is considered deep learning. It uses convolutional layers to identify small local receptive fields, which are combined to find higher level patterns: hierarchical representations (Khemani et al., 2024; W. Liu et al., 2017; Samaddar et al., 2021). Its architecture is made up out of several convolution layers, followed by a pooling layer, which then again follows to a convolution layer. Therefore, the output of the convolution layers will be input for the pooling layer, then the output of the pooling layer will be the input for the other convolution layer. The final output layer is a fully connected layer that is connected to all of the neurons in the previous layer. Thus, it is also a feedforward NN which can be trained by backpropagation as well, similar to the MLP. It is used successfully in large-scale applications where the number of training patterns and the dimensionality of the data are large, and similar to the MLP, it only works well when the amount training data is large. It is mainly used in image processing and in speech processing (Aggarwal & Murty, 2021). These are grid-like structures, and generally well-suited for tasks where spatial relationships between neighbouring elements are important (Khemani et al., 2024). It reduces the parameters used in ANN, and therefore is useful for solving complex tasks (Samaddar et al., 2021). Nowadays, it can be used for handwriting recognition, face detection, behaviour recognition, speech recognition, recommender systems, image

classifications and natural language processing. An advantage is that it does not require a fair amount of pre-processing (W. Liu et al., 2017).

Recurrent Neural Network (RNN)

MLP and CNN models have the limitations that they cannot handle sequences, nor problems with an unknown size of the input data. However, other NNs that are developed, such as RNNs, LSTM, and GRUs are able to solve these issues (Aggarwal & Murty, 2021). The first type is called Recurrent Neural Network (RNN), this DL approach already exists since the 1980s. They are commonly used as there were made advancements in the computing power and the large amount of data they can access. This type of NN operates by having feed-forward sequential input operations, with a variable length, which goes through the recurring hidden layer, whereof the activity differentiates, depending on what was used in the previous time (Abualigah et al., 2024). Its key feature is that it uses recurrent connections established in a network (Rubio-Martín et al., 2024). Thus, previous outputs determine the next outputs and are great for natural language processing (NLP). It uses the same set of parameters for all time steps of an input. This learns the dependencies between elements at different time steps but also avoids overfitting (Aggarwal & Murty, 2021). When this model is trained, it is trained by the reverse training algorithm, and a minimum error function is retrieved. This way it learns from its previous mistakes, however, there is one issue. The RNN suffers from the disappearance of a problem, as it 'forgets' the mistakes from longer ago. This is called the vanishing gradient problem, where contributions of the context geometrically decay over time (Rubio-Martín et al., 2024). The use of Long Short-Term Memory (LSTM) could resolve this problem (Abualigah et al., 2024).

Long Short-Term Memory (LSTM)

This NN is a sophisticated type of RNN, which solves the issues previously discussed. LSTM can develop algorithms and statistical models that can analyse patterns. It has short-term memory and long-term memory that study and learn from sequential data. This type can for example predict the course of a certain stock (Masrour et al., 2021). The LSTM keeps a cell state over time, which can give information earlier obtained to later ones. Therefore, it keeps track of earlier dependencies, over a longer period of time. To get a bit deeper into the working of this NN, the three types of gates it uses are explained. There is the input gate, which controls the amount of information that will be stored and thus forms a filter. Then, there is the forget gate, which will determine the percentage of the previous let through data is kept. Last, there is the output gate, which controls what part of the internal cell state is exposed to the next layers in the network (Rubio-Martín et al., 2024). The largest difference between RNN and LSTM is the cell structure. The latter uses a more complicated structure, it uses the aforementioned gates and four neural network layers. They are similar in a way that they both have hidden states (Aggarwal & Murty, 2021).

Gated Recurrent Unit (GRU)

The GRU has a similar function as the LSTM, it can analyse sequential data and filter information to prevent the vanishing gradient problem. They also make use of internal gates, but the GRU only has two gates: a reset gate and an update gate. The latter is functions similar to the forget and input gate used in LSTM, it determines which information should be saved. Then, the reset gate determines how much previous information is forgotten. They do not make use of cell states but hidden states to transfer information. Since they only have two operators, they are a bit faster to train than LSTM, but it depends on the application which one is better to use. They are used for similar type of applications as LSTM: speech recognition, synthesis, and text generation (Illustrated Guide to LSTM's and GRU's: A Step by Step Explanation | by Michael Phi | Towards Data Science, n.d.). However, it has various applications. In one research they used GRUs to explore temporal dependencies between tweets, and anywhere the capture of temporal correlations is needed, they can be useful (X. Yang et al., 2023).

Autoencoder

This is an unsupervised model used to learn representations in a low dimensional space. It contains an encoder and decoder. It uses non-linear transformations on the input to compress it, so the original data is reconstructed in a lower dimensional representation. This is the work of the encoder. Then, the decoder decompresses the input back to its original. Applications of the autoencoder are dimensionality reduction and representation learning, and recently as generative models (Aggarwal & Murty, 2021).

Graph Neural Network (GNN)

GNNs draw inspiration from CNNs, however, CNNs nor RNNs are not suitable to efficiently handle data structured as graphs. Where CNNs are excellent in processing images, RNNs in handling sequences, and

GNNs in processing information from graphs. The technique Graph Convolution is used. It can process graph-structured data directly and thus leads to a graph neural network. Graphs are representations for nodes which are connected by edges, which perfectly enables them to model relationships and dependencies in complex systems. The GNNs perform operations on the input that sums information from neighbouring nodes to update features of a central node (Khemani et al., 2024).

Traditional Machine Learning Algorithms

DL is a more advanced branch within ML, but there are many more algorithms which are used in Financial Engineering & Management. Below, we list older, more traditional machine learning algorithms: Logistic Regression (LR), Linear Discriminant Analysis (LDA), Decision Trees (DT), Support Vector Machine (SVM), Naive Bayes, and K-Nearest Neighbours (KNN).

Logistic Regression (LR)

Logistic regression can be used for classification and uses a logistic function to model a binary dependent variable (Y. Zhang et al., 2020). The predictions it can make are based on historical values. It differentiates itself by making use of another type of distribution function than just linear, which makes it suitable for credit rating issues, for example. When the dependent variable consists of concrete values this is a suitable method. The logistics function is defined as the following:

$$f(u) = \frac{1}{1 + e^{-u}}$$

With a range between 0 and 1, and only real numbers R in its domain. It will determine the relationship between previous input independent variables, the results of LR can be used as a probabilistic measure and can be regularised to avoid overfitting (Murugan & T, 2023).

Linear Discriminant Analysis (LDA)

A Linear Discriminant Analysis is also a supervised ML regression-like classification technique, which is used to find the best linear combination of multiple predictive variables that are used to maximally distinguish between different groups (Willette et al., 2022). It is mainly used for classification. The method is based on an eigenvalue resolution and gives the precise solution of the maximum of the inertia. There must be noted that this method can only be used for linear problems. Some use the LDA to make a GDA (Generalized Discriminant Analysis) by mapping the input space into a high dimensional feature space (Baudat & Anouar, 2000). It is also used for dimensionality reduction, where it can minimize the distance between two data points of the same class and maximize the inter-class distance. The advantages of this method with respect to its latter application are that besides the dimension reduction, the separation of classes is better than with a PCA. However, the downsides are that labelled data is required since it is supervised ML, and sometimes classes overlap due to athematic mean usage. And the outliers will affect the accuracy of the model (Soni et al., 2022).

Decision Trees (DT)

A Decision Tree is supervised ML, which is used for classification and regression as well. It has a structure similar to a tree, with a root node, branch nodes and leaf nodes, where each of those nodes represents a certain characteristic or attribute (Sheth et al., 2022). This ML model is mainly used in situations where there are clear nonlinear relationships between variables. They are specifically good for mapping capabilities, where data is visualized in a data map. It builds decision trees, with an algorithm that keeps splitting the data into the smallest sequences possible, which puts a focus on some parts of the data. This process is repeated until the stop requirements are achieved (Rubio-Martín et al., 2024). The benefits of using a DT are that it is effective for both regression and classification, easy to understand, it can complete certain missing information in the data, and it can handle categorical and quantitative values. Plus, it is very efficient due to the tree traversal algorithm. However, there are also some downsides. Over-fitting could occur, but this could be resolved by deploying Random Forest, explained later in this paragraph. Furthermore, it can be unstable, difficult to manage in size, prone to errors in sampling, and providing the local optimal best answer instead of the global ideal solution (Sheth et al., 2022).

Support Vector Machine (SVM)

Another type of ML is called a Support Vector Machine (SVM). This is also a supervised machine learning algorithm (Ma et al., 2023b). It is used for classification and regression problems. It can handle linear and non-linear problems. With training data, it can create a line or a hyper plane which will separate the data points into classes (Marneni & Vemula, 2022). The values that are closest to the classification margin are known as support vectors. The algorithm tries to maximize the margin that is present between the support vectors and the hyperplane. Support Vectors are often considered as one of the best classifiers, and many environments and toolboxes implement them (Gove & Faytong, 2012). It is based on statistical learning theory and is a class of generalized linear classification techniques, it classifies the input data by looking for decision boundaries (Xiong et al., 2022). There are some disadvantages to this model. It takes long to train and therefore performs poorly when working with large data sets or when the data set has a lot of noise. Plus, the probability calculations are not provided by the SVM and therefore it can be difficult to interpret the final SVM model (Sheth et al., 2022).

Naive Bayes

This type of ML is constructed using the Bayes rule, and generates Naive Bayes classifiers, which are a group of classification methods. It makes use of several algorithms, which all use the same principle of having each pair of features to be independent. This is the Bayes rule that it implements:

$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

where y is a two-outcome class variable, and X is an n-dimensional dependent feature vector. The class Y with the highest probability is sought, and the output values are used to determine the error of the procedure. If these are categorical, then it is expressed as an error rate, which is the proportion of how many times the prediction was wrong. This error rate of the Bayes procedure is the lowest that any classifier with a random outcome can provide. Other advantages of this procedure are that it is easy to set up, it scales proportionally to the number of predictors and data points, it does not require a lot of training data, it can deal with both discrete and continuous data, binary and multi-class classification problems and it will make stochastic recommendations. The data can be processed continuous or discontinuous and is not affected by irrelevant variables, while it assumes conditional independence. However, there are some negative sides. The models tend to be simplistic, and models that are trained for a longer time and are more complex, generally outperform them. It is also difficult to implement continuous variables, there is no online option so all the data must be stored, when there are more than 100.000 attributes, the model will not scale anymore, and it takes a lot of time to run the model and also requires more memory than SVM or LR (Sheth et al., 2022).

K-Nearest Neighbours (KNN)

This supervised ML algorithm can be used for classification and regression problems. The algorithm finds the k number of nearest data points of the sample and uses the data of the neighbours to classify or predict the value of that specific sample. This method does not have any presumptions on the data, and the choice of the value of k is an important choice as it will determine the result. If the value of k is bigger, the boundary gets smoother (Samaddar et al., 2021). This model does not learn, and is therefore not classified as DL, as it just holds a copy of the data used as training (Rubio-Martín et al., 2024). The advantages of the model are that it is a simple technique and has a quick and inexpensive implementation, the model is ideal for multi-modal classes and sometimes just the most effective way. The disadvantages are that it can be costly to classify unknown records, and when the training set grows, it will take more computation time. When noisy or irrelevant variables are used as input the accuracy of the result will be of less quality (Sheth et al., 2022). This method has a large use in text applications, one given example is a system that categorized public complaints (Murugan & T, 2023).

Ensemble Methods and Advanced Techniques

Here, ensemble methods are described. This implies that multiple models are stacked. This is the case for a Random Forest (RF), and Gradient Boosting Machines (GBM). Furthermore, some advanced techniques which build on these models are described: Adaptive Boosting (AdaBoost), Xtreme gradient boosting (XGB), and Stacked Generalization.

Random Forest (RF)

A Random Forest is a classifier which as is in the name, contains multiple decision trees. The output of this algorithm consists of the voting of these multiple tree (Xiong et al., 2022). It is thus an ensemble of decision trees. Each tree is generated with a random vector which is sampled independently. This technique can generate a training set by randomly choosing examples in a certain configuration (Y. Zhang et al., 2020). This type of ML is excellent for classification and regression problems, and it produces accurate results for data of higher dimensions (Samaddar et al., 2021).

Gradient Boosting Machines (GBM)

A Gradient Boosting Machine is also an ensemble method. It uses a committee-like approach where it trains multiple models in a sequence. Every model is represented as t, and its goal is to rectify the mistakes made in the previous models by going back to t-1 (de Holanda et al., 2024). It works particularly well with small datasets. Plus, overfitting can be prevented to some extent (Liang et al., 2020). It can handle a lot of mixed predictors, quantitative and qualitative. It is therefore popular to use for real-time traffic and weather forecasting, as it can handle the data without preprocessing of rescaling or transformation and thus can be used directly as input. It can also handle rough data which has not been cleaned and has missing values. However, one disadvantage is that it is not that stable and has trouble predicting over a longer period of time. Boosting is an optimization technique that will minimize a loss function by adding a new tree which can reduce the loss function. When the model becomes bigger, the existing trees stay the same but new ones are added, where each weight is newly calculated (Ahmed & Abdel-Aty, 2013).

Adaptive Boosting (AdaBoost)

AdaBoost is an iterative algorithm. It will train different weak classifiers of a training set and will them combine them into a stronger one (X. Li et al., 2019). It encourages a new classifier to learn from previous mistakes spotted in other classifiers, by assigning larger weights. Subsequently it uses a weighted majority vote to make forecasts. It is a quite simple model but deemed very effective to achieve a greater classification accuracy (Hu et al., 2014). It is one of the most popular methods to solve binary classification problems. It also uses the decision tree model as a construct (Pham et al., 2021).

Xtreme Gradient Boosting (XGB)

Xtreme Gradient Boosting is based on RFs and GBMs, but it contains multiple optimizations. It works by focusing on a selected group of variables and repeating this process multiple times. Furthermore, just as with the GBM, every tree considers the models of the previous trees and highlighting the misclassified cases. When then a new tree is created, the error is calculated which assists in minimizing the error in following trees. This model is very efficient and uses different techniques that prevent overfitting (Rubio-Martín et al., 2024). This model is used often in ML and is highly scalable. It outperforms many other ML algorithms due to its speed and accuracy (Lei et al., 2020). It gives insights of data compression, sharding (breaking data up in smaller parts) and cache access patterns (the strategies of storing data most efficient) (Murugan & T, 2023; Y. Zhang et al., 2020).

Stacked Generalization

This is an ensemble modelling technique to merge certain classification models by using a meta-classifier. It uses non-linear weights for low-level predictors, which lowers the generalization error rate, and it will enlarge the prediction accuracy (Niyogisubizo et al., 2022). It can improve single-model detection. This method is also called stacked ensemble learning. The meta-classifier can be a pooling operation or a more complex neural network and makes a final decision about the predictions (Falcão et al., 2023).

Optimization and Feature Selection Techniques:

Genetic Algorithm (GA)

This model is developed in the 1960s and 1970s, and it is loosely based on the natural selection theory of Darwin (X.-S. Yang, 2021). The algorithm finds the optimal solution by changing the process of solving certain problems in other domains into processes like chromosome crossover and mutation in the biological field (C. Liu et al., 2022). It uses genetic operators as recombination, mutation, and selection of adaptive and artificial systems. One advantage of using the GA is that it can deal with very complex problems and with parallelism. It can handle different types of optimizations, where the objective function can be stationary or non-stationary (it will change with time), linear or non-linear, continuous, or discontinuous, or with random noise. As the 'children' in a population can act like independent agents, the group can explore the search space in multiple directions at the same time, which makes it perfect to parallelize the algorithms for implementation. Thus, manipulation of different parameters or groups of encoded strings can happen at the same time. However, the algorithm can produce meaningless results if the wrong choices are made. It can also make it difficult for the algorithm to converge. Therefore, it is vital that the formulation of the fitness function, use of population size, and important parameters such as the rate of mutation and crossover, selection criteria of the new population are carried out carefully. Still, GA's are often used for nonlinear optimization (X.-S. Yang, 2021).

Particle Swarm Algorithm (PSA)

This is a random search algorithm that finds the best solution by imitating the group behaviour of birds (C. Liu et al., 2022). It can be implemented and used easily to solve different types of optimization problems. Its main strength is its fast convergence, and it compares well with many other global optimization algorithms (Sun

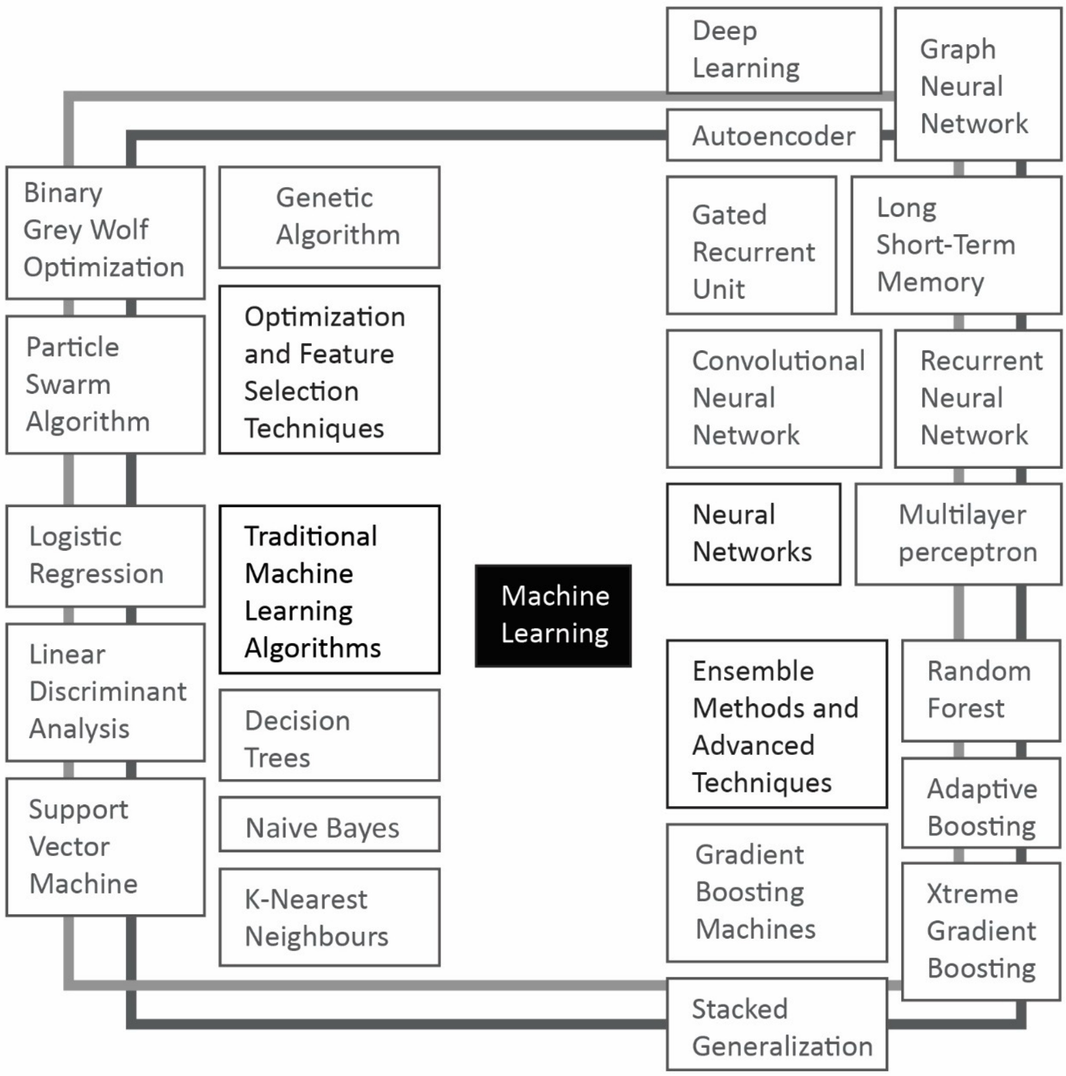
& Liu, 2013).

Binary Grey Wolf Optimization (GWO)

This method is very innovative and serves as an example of the large range of new algorithms available. Not many papers make use of this yet, one paper uses it for a feature selection module (Ma et al., 2023a). It is mainly used for continuous-type optimization problems. It is very complex but based on a principle in nature. It is inspired by grey wolves, as it mimics the leadership hierarchy and the hunting mechanism. There are four types of grey wolves: alpha, beta, delta, and omega. Plus, the three steps of hunting: searching for prey, encircling prey, and attacking prey are used (Mirjalili et al., 2014). It is used for continuous-type optimization problems because it allows arbitrary movement of the grey wolf position in the search space. However, in the binary space those positions are different. Therefore, a correlation must be established through a conversion function (Ma et al., 2023c). There are many more optimization algorithms, such as the Gravitational Search Algorithm (GSA), Differential Evolution (DE), Evolutionary Programming (EP), and Evolution Strategy (ES) (Mirjalili et al., 2014).

Combining algorithms

Often, papers incorporate various techniques to get the most optimal model. One paper discusses the training of the NN, by using different kind of algorithms. They use three AI algorithms, ANN, the genetic algorithm, and the particle swarm algorithm. In this study they made an early-warning model regarding the financial services of the international trade supply chain of the energy industry (C. Liu et al., 2022). Another paper deploys three other algorithms: the cluster based KNN, cluster based LR, and cluster based XG Boost. These were used for their ability to predict loan defaults and their occurrence of likelihood (Murugan & T, 2023). In another study, where they examined the flood risk in global watersheds, they used four machine learning algorithms in this study: logistic regression, naive Bayes, AdaBoost, and random forest (X. Li et al., 2019). As one notices, often a NN with an ensemble method and/or optimization algorithm are combined, dependent on the goal of the model.



Appendix B:
Table of papers used in SR

Paper Title	Writer(s), Year	Crowdfunding Campaigns	Reward-Tier campaigns	Machine Learning (ML)	Textual Analysis	LLM & Similar (e.g., BERT)	Bayesian Optimization & Similar
Community activism, Social ties and ESG campaign success	Akhil Raju, Vijaya B. Marisetty (2024)	✓	✓	✓	✓		
Generative AI for Research Data Processing	Modhurita Mitra et al. (2024)	✓		✓	✓	✓	
An Analysis on the Importance of Persuasion Strategies in Environmental-Oriented Online Crowdfunding Projects	Yun Biao et al. (2023)	✓		✓	✓		
Signaling persuasion in crowdfunding entrepreneurial narratives: The subjectivity vs objectivity debate	Wei Wang et al. (2021)	✓	✓	✓	✓		✓

A Swarm Enhanced Light Gradient Boosting Machine for Crowdfunding Project Outcome Prediction	Shuang Geng et al. (2020)	✓		✓	✓		✓
Effectiveness of BERT Model with the Weaker Loss Function for Chinese Keyword Extraction on Crowdfunding Projects	Qi Li et al. (2022)	✓		✓	✓	✓	
The determinants of crowdfunding success: A semantic text analytics approach	Hui Yuan et al. (2016)	✓	✓	✓	✓		
Mining and investigating the factors influencing crowdfunding success	Yang Song et al. (2019)	✓	✓	✓	✓		

The relationship between a charity crowdfunding project's contents and donors' participation: An empirical study with deep learning methodologies	DongIl Lee, JaeHong Park (2020)	✓		✓	✓		
The power of machine learning methods to predict crowdfunding success: Accounting for complex relationships efficiently	Ramy Elitzur et al. (2024)	✓	✓	✓	✓		
Financing sustainable entrepreneurship: Unpacking the role of campaign information and risk disclosure in reward-based crowdfunding	Christian Hopp et al. (2025)	✓		✓	✓	✓	
Emotional brightness and crowdfunding performance	Jonathan Sitruk et al. (2024)	✓		✓	✓		
How signal portfolios affect success in equity-based crowdfunding: Evidence from the Chinese hotel industry	Liqing La, Seongsoo (Simon) Jang (2024)	✓		✓	✓	✓	

Risk disclosure and entrepreneurial resource acquisition in crowdfunding digital platforms: Evidence from digital technology ventures	Hong Huo et al. (2024)	✓	✓	✓	✓		
Exploring the subjective nature of crowdfunding decisions	Jie Ren et al. (2021)	✓	✓		✓		
Emphasizing the entrepreneur or the idea? The impact of text content emphasis on investment decisions in crowdfunding	Wei Wang et al. (2020)	✓		✓	✓		
An integrated model of prosocial crowdfunding decision: Three utility components and three informational cues	Jenny Jeongeun Yoo et al. (2023)	✓		✓	✓		
Exploring the effects of social capital on crowdfunding performance: A holistic analysis from the empirical and predictive views	Xi Zhang et al. (2022)	✓	✓	✓	✓		

Choose your words carefully: Harnessing the language of crowd-funding for success	Aaron H. Anglin, Robert J. Pidduck (2022)	✓	✓	✓	✓		
Who said what: Mining semantic features for success prediction in reward-based crowdfunding	Liqian Bao et al. (2022)	✓	✓	✓	✓		
What social issues do people invest in? An examination based on the empathy–altruism hypothesis of prosocial crowdfunding platforms	Koichi Nakagawa, Genjiro Kosaka (2022)	✓		✓	✓		
Can you hear me now? Engendering passion and preparedness perceptions with vocal expressions in crowdfunding pitches	Thomas H. Allison (2022)	✓		✓	✓		

Informational cues or content? Examining project funding decisions by crowdfunders	Yan Lin, Wai Fong Boh (2021)	✓		✓	✓		
Natural language processing versus rule-based text analysis: Comparing BERT score and readability indices to predict crowdfunding outcomes	C.S. Richard Chan et al. (2023)	✓	✓	✓	✓	✓	
The Success of Cancer Crowdfunding Campaigns: Project and Text Analysis	Xupin Zhang et al. (2023)	✓		✓	✓		
A comparative analysis of the effect of initiative risk statement versus passive risk disclosure on the financing performance of Kickstarter campaigns	Wei Wang et al. (2024)	✓		✓	✓	✓	

Confidence is Good? too Much, not so Much: Exploring the effects on crowd-funding success	Naomi Moy et al. (2024)	✓	✓		✓		
Digital identities of female founders and crowd-funding performance: an exploration based on the LDA topic model	Yalin Wang, et al. (2022)	✓		✓	✓		
Multimodal dynamic graph convolutional network for crowdfunding success prediction	Zihui Cai et al. (2024)	✓		✓	✓	✓	
Signaling theory in charity-based crowdfunding: Investigating the effects of project creator characteristics and text linguistic style on fundraising performance	Yimeng Zhai, Wang-bing Shen (2024)	✓		✓	✓		

Investors' responses to linguistic and communication factors in crowdfunding projects: An investigation of foodservice startups	Chunsheng (Jerry) Jin et al. (2024)	✓	✓		✓		
Predicting the fundraising performance of environmental crowdfunding projects: An interpretable machine learning approach	Zhanyu Liu, Saiquan Hu (2024)	✓		✓	✓		
Emotion regulation and cheap talk as signaling strategies: Evidence from crowdfunding for Ukraine	Alexandra Fischmann, Larry Zhiming Xu (2024)	✓		✓	✓		✓
New model of utility analysis and performance prediction in crowdfunding: A perspective of behavior-related decision	Ju Wei et al. (2024)	✓		✓	✓		

Green oriented crowd-funding campaigns: Their character-istics and diffusion in different institutional settings	Vincenzo Butticè et al. (2019)	✓		✓	✓		
Proximal language predicts crowdfund-ing success: Behavioural and exper-imental evidence	Xun Zhu (2022)	✓		✓	✓		
Monetizing entrepre-neur re-sponse to crowdfund-ing with text analytics	Wei Wang et al. (2024)	✓		✓	✓		
Impact of Image Content on Medical Crowdfund-ing Success: A Machine Learning Approach	Renwu Wang et al. (2024)	✓		✓	✓		
Linguistic under-standability, signal ob-servability, funding op-portunities, and crowd-funding campaigns	Wei Wang et al. (2022)	✓	✓	✓	✓		

The impact of soft information extracted from descriptive text on crowdfunding performance	Cuixia Jiang et al. (2020)	✓	✓	✓	✓		
Will a digital camera cure your sick puppy? Modality and category effects in donation-based crowdfunding	Larry Zhiming Xu (2018)	✓		✓	✓		
Emotional Intensity-based Success Prediction Model for Crowdfunded Campaigns	Stefano Faralli et al. (2021)	✓	✓	✓	✓		
The merits of a sentiment analysis of antecedent comments for the prediction of online fundraising outcomes	Wei Wang et al. (2022)	✓		✓	✓		

Appendix C:
BOHB in more detail

This Appendix shows the ML mechanisms used in this thesis in more detail.

Bayesian Optimization (BO)
The underlying theory of Bayesian Optimization is based on Bayes' Theorem. According to Thomas Bayes (1763) given the dependent variable y, and the independent variable x, the posterior probability P(y | x), can be written as:

$$P(y|x) = \frac{P(x|y) \cdot P(y)}{P(x)}$$

Thus, one notices that the posterior probability P(y | x) is proportional with the Likelihood P(x | y) times the previous probability P(y), written as:

$$P(y|x) \propto P(x|y) \cdot P(y)$$

This concept of updating prior knowledge with new observations forms the foundation of Bayesian Optimization. The mechanism is using a prior belief over the possible objective function(s) by sequentially refining that model as data are observed via Bayesian posterior updating (Shahriari et al., 2016). A certain objective function f, which is most likely a black-box function without a closed form, is approached. In our case, this is the validation F1-score achieved by an XGBoost model trained with a specific set of hyperparameters This is done by observing data points D={(x_0,y_0),…,(x_(i-1),y_(i-1))} (Falkner et al., 2018). The mathematical function which describes the goal of BO is finding the true global maximiser (or minimiser) with the following function:

$$x^* = \arg\max_{x=X} f(x)$$

In our case, the ttings (reward count,avg reward price,early bird flag,exclusive flag) that need to give the highest f(x). To approach this, two functions are used: a surrogate function and an acquisition function (Shahriari et al., 2016). A surrogate function consists of a prior distribution that shows the belief about the objective function(s). This surrogate function is often modelled by Gaussian Processes (GP), a Tree-structured Parzen Estimator (TPE), or a Random Forest (RF) regressor (Cho et al., 2020a; Swaminatha Rao & Jaganathan, 2024; L. Yang & Shami, 2020). The choice of this surrogate model has a great influence on the performance (Cho et al., 2020). An acquisition function is needed to generate the next hyperparameter configuration θ. Some acquisition functions are the Upper Confidence Bound (UCB), Entropy Search (ES), Predictive Entropy Search (PES) Probability of Improvement (PI), and the Expected Improvement (EI) (Klein, Falkner, Bartels, et al., 2017). UCB, PI and EI are most often used (Cho et al., 2020). SMAC, Auto-WEKA, and Auto-sklearn are automated machine learning frameworks that use Expected Improvement (EI) as the acquisition function for optimizing hyperparameter configurations. It tries to minimize the loss, and select a high uncertainty, shown in the following equation:

$$I_{L_{min}(\theta)} = \max\{L_{min} - L(\theta), 0\}$$

Since the value of L(θ) is unknown, the expectation is calculated by:

$$\mathbb{E}_{L_{min}}\,[L_{min}(\theta)] = \int_{-\infty}^{L_m} \max\{L_{min} - L(\theta), 0\} \cdot\, p_{M_L}(L|\theta)\; d\theta$$

The type of surrogate function determines the exact form of the acquisition function (Brazdil et al., 2022). Often, GP are chosen for p(f), used in Spearmint (L. Li et al., 2018), as they have a high descriptive power and analytic traceability. This is the default surrogate model of classical Bayesian optimization (J. Zhang et al., 2023). It is defined as a collection of random variables, where every finite subset follows a multivariate normal distribution. It has the following characteristics:

Mean function m(x) (typically set to m(x)=0 for simplicity)

Covariance function (kernel) k(x,x^'), which determines how observations influence the predictions

It follows from the given observations D_n={(x_j,y_j)}_(j=1)^n=(X,y) and a Gaussian likelihood p(y|X,f(X)), the posterior distribution p(f|D_n) is also a GP with a analytically tractable mean and covariance functions. Specifically for BO, the Matérn 5/2 kernel is often used, in a form called the Automatic Relevance Determination (ARD), which is expressed as:

$$k_{\frac{5}{2}}(x,x') = \theta\left(1+\sqrt{5}d_\lambda(x,x') + \frac{5}{3}d_\lambda^2(x,x')\right)e^{-\sqrt{5d_\lambda(x,x')}}$$

In this equation, θ and λ are hyperparameters that control the behaviour of the GP, and additional noise covariance parameter accounts for noisy observations. This is all based on the Mahalanobis distance: d_ λ (x,x^')=(x-x^')^T diag(λ)(x-x') (Klein, Falkner, Bartels, et al., 2017).

Another surrogate function is the TPE, which models the probability distributions of configurations rather than modelling the objective function f directly. BO-TPE is tree based and supports conditional hyperparameters(Swaminatha Rao & Jaganathan, 2024). It uses a kernel density estimator to model the following densities:

$$l(x) = p(y < \alpha | x, D)$$
$$g(x) = p(y < \alpha | x, D)$$

To select the next candidate x_new for evaluation, TPE maximizes the ratio l(x)/g(x) which is mathematically equivalent to maximizing the expected improvement (EI). Unlike GP, TPE efficiently handles both continuous and discrete search spaces and scales linearly with the number of data points, making it computationally advantageous compared to the cubic-time scaling of GPs in many Bayesian optimization applications (Falkner et al., 2018).

Lastly, RF is applied in SMAC (L. Li et al., 2018). This model has the advantages of a concise form, and it is less prone to over-fitting than GP (J. Zhang et al., 2023). It handles discrete, continuous, categorical, and conditional HPs (Swaminatha Rao & Jaganathan, 2024). RF computes a predictive mean μ_θ and variance σ_θ, modelled as a Gaussian distribution N(μ_θ,σ_θ) based on the frequentist estimates p(L|θ). In this case, the EI is derived from a closed-form expression:

$$\mathbb{E}\left[I_{L_{min(\theta)}}\right] = \sigma_\theta \cdot [u \cdot \Phi(u) + \varphi(u)]$$

Where the probability function is represented by $u = \frac{(L_{min}-\mu_\theta)}{\sigma_\theta}, \varphi$ and the cumulative density function by Φ of a normal distribution (Brazdil et al., 2022).

The pseudocode of a BO model is shown in Algorithm 1, adapted from Andhika Viadinugroho & Rosadi (2023); Bischl et al. (2023); Klein, Falkner, Bartels, et al. (2017); Shahriari et al. (2016).

Bayesian Optimization is often used for A/B testing, Recommender systems, robotics and reinforcement learning, environmental monitoring and sensor networks, preference learning and interference learning, automatic Machine Learning and Hyperparameter Tuning, Combinatorial Optimization, and Natural Language Processing and Text (Shahriari et al., 2016).

Hyperband (HB)

Another HPO method is Hyperband. This is a hedging strategy which divides the whole problem into several halving problems, as described previously. These successive halving problems are also referred to as brackets, where they all look for a unique value of η for a B. Beforehand, two variables need to be defined: R, the maximum resource for a singular configuration in one iteration, and λ or η, the elimination factor (Goay et al., 2021). HB starts with the configuration, in an outer loop, which prioritizes exploration with the constraint that at least one configuration has the allocation of R resources and then evert next bracket reduces the factor n by approximately η and continues until every configuration is allocated R resources. Therefore, it works well where adaptive allocation is desired, but also where more conservative configurations are required. The inner loop uses the successive halving mechanism (L. Li et al., 2018).

The pseudocode of Hyperband is shown in Algorithm 2, adapted from (Falkner et al., 2018; L. Li et al., 2018;

Algorithm 1 Bayesian Optimization

1: Initialize data D_0 using an initial design (e.g. a random sampling of HP configurations)
2: **for** n 1; 2; . . . ; **do**
3: select new x_{n+1} by optimizing acquisition function α
$x_{n+1} = \arg\max_x \alpha(x; D_n)$
4: conduct sampling on objective function $y_1 = f(x_{n+1})$
5: augment data $D_{n+1} = \{D_n, (x_{n+1}, y_{n+1})\}$
6: update statistical model by adding newly observed data point
7: **end for**

Algorithm 2 Hyperband

Initialize maximum budget R, discard proportion η, and hyperparameter space X
compute $s_{max} = \lfloor \log_{\eta}, R \rfloor$, total budget $B = (s_{max} + 1)R$
for $s = s_{max}, s_{max} - 1, \dots, 0$ do:
    determine number of configurations
    $n_1 = \left\lceil \frac{B}{R} \frac{\eta^s}{s+1} \right\rceil$
    $r_1 = R\eta^{-s}$
    sample n_1 configurations $\{x_1, x_2, \dots, x_{n_1}\}$ randomly from X
    for $i = 0, \dots, s$ do:
        determine number of surviving configurations
        $\eta_i = \lfloor \eta_1 \eta^{-i} \rfloor$
        determine resource allocation for this iteration
        $r_i = r_1 \eta^i$
        evaluate each configuration x_j using r_i resources
        $y_j = f(x_j, r_i)$
        select top $\lfloor \frac{n_i}{\eta} \rfloor$configurations for the next iteration
    end for
end for
return configuration with best evaluation result

Y. Li et al., 2021). Since HB is so versatile it can be combined with any hyperparameter sampling approach, which is the reason BOHB could be created (L. Li et al., 2018). One study from (Kavzoglu & Teke, 2022) shows Hyperband is a reliable HPO method, as it scored higher than GA, BO-GP, and BO-TPE, and computed the results faster.

Bayesian Optimization and Hyperband
Hyperband on itself has the downside that it uses a method like random search to select the new configurations, where no information is passed on to optimize this process. This is where BO is introduced, to substitute this random process and use the resources more efficiently. (Passos & Mishra, 2021). The convergence to the optimal solution is improved by combining the ability of Hyperband to dynamically allocate resources over different configurations and the ability of guided sampling in Bayesian Optimization. This way it improves the convergence speed compared to the individual models (Falkner et al., 2018).
In most cases, the surrogate model used in BOHB

Algorithm 3 Tree Parzen Estimator

Initialize data D_0 using an initial design (e.g. a random sampling of HP configurations)
for n 1; 2; . . . ; **do**
    define promising set D_l containing top $\lceil \gamma \cdot |D_n| \rceil$ best-performing configurations
    define non-promising set $D_g = D_n \backslash D_l$
    fit density models $l(x)$ using D_l and $g(x)$ using D_g
    sample candidate configurations
    $C = \{x_j \sim l(x) \big| \, j = 1, \dots, n_c\}$
    select new x_{n+1} by optimizing EI
    $x_{n+1} = \arg max_{x \in C} EI_{y^+}(x)$
    conduct sampling on objective function $y_{n+1} = f(x_{n+1})$
    augment data $D_{n+1} = \{D_n, (x_{n+1}, y_{n+1})\}$
    update statistical model by adding newly observed data point
end for

is TPE, which uses the EI as acquisition function.
As mentioned, TPE is suitable since it handles both continuous and discrete search spaces and scales linearly with the number of data points, making it computationally advantageous (Falkner et al., 2018). TPE uses a KDE to model good and bad hyperparameters separately, which gives a PDE for hyperparameters given their observed performance (Kim et al., 2024) The pseudocode of a TPE model is shown in Algorithm 3, adapted from (Andhika Viadinugroho & Rosadi, 2023).

The Hyperband model uses the pruning strategy, which refers to the technique of the early stopping mechanism (Andhika Viadinugroho & Rosadi, 2023). However, the random sampling in the HB is now replaced by the BO model, the TPE (Y. Zhang et al., 2024). The pseudocode of a BOHB model is shown in Algorithm 4, adapted from (Andhika Viadinugroho & Rosadi, 2023; Falkner et al., 2018; Y. Zhang et al., 2024).

Algorithm 4 Bayesian Optimization and Hyperband (BOHB)

Initialize data D_0 using an initial design (e.g. a random sampling of HP configurations)
compute $s_{max} = \lfloor \log_\eta, R \rfloor$, total budget $B = (s_{max} + 1)R$
for $s = s_{max}, s_{max} - 1, \dots, 0$ do:
  determine number of configurations
  $n_1 = \left\lceil \frac{B}{R} \frac{\eta^s}{s+1} \right\rceil$
  $r_1 = R\eta^{-s}$
  sample n_1 configurations $\{x_1, x_2, \dots, x_{n_1}\}$ randomly from X
  for $i = 0, \dots, s$ do:
    determine number of surviving configurations
    $\eta_i = \lfloor \eta_1 \eta^{-i} \rfloor$
    determine resource allocation for this iteration
    $r_i = r_1 \eta^i$
    if $rand() < \rho$ then
      select new x_{n+1} using random configuration
    else
      find best density model
      $b = \arg\max \{\mathrm{D_b}: |D_b| \geq N_{min} + 2\}$
      if $b = \emptyset$ then
        select new x_{n+1} using random configuration
      else
        compute number of good and bad samples
        $N_{b,l} = \max(N_{min}, q \cdot N_b)$,
        $N_{b,g} = \max\left(N_{min}, N_b - N_{b,l}\right)$
      fit density models $l(x)$ and $g(x)$ using good and bad samples
      adjust bandwidth $l'(x) = b_w \cdot l(x)$
      sample N_s candidate configurations from $l'(x)$
      select new x_{n+1} by optimizing EI
      $x_{n+1} = \arg max_{x \in N_s} l(x)/g(x)$
    end if
    conduct sampling on the objective function $y_{n+1} = f(x_{n+1})$
    augment data $D_{n+1} = \{D_n, (x_{n+1}, y_{n+1})\}$
  end for
end for
return configuration with the best evaluation result

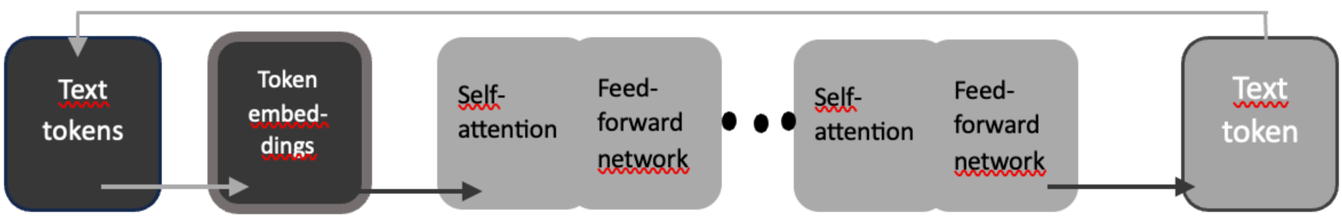
Appendix D: Transformer Architecture LLaMA

This section gives more detail on the working of LLaMA, a Transformer Model.

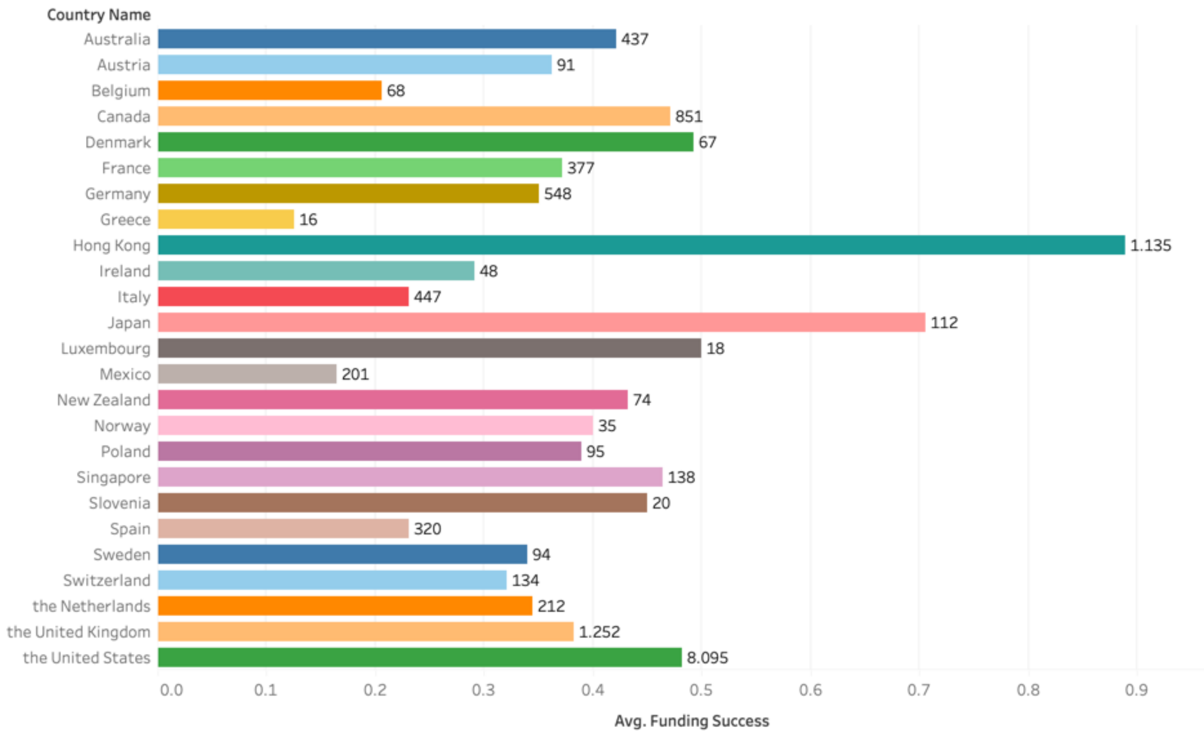
Transformer Architecture
In many LLMs the core component is a self-attention module in Transformer. A standard Transformer consists of two parts: an encoder and a decoder. The encoder is composed of a stack of N = 6 identical Transformer layers, each with two sub-layers: (1) a multi-head self-attention layer, and (2) a position-wise fully connected feed-forward network. The decoder mirrors this structure but adds a third sub-layer that performs multi-head attention over the encoder's output. However, as a decoder-only model, LLaMA removes the encoder and relies entirely on self-attention to model sequential dependencies; LLaMA is an autoregressive language model. Given a context sequence X, LLaMA tries to predict the next token y, as shown below. The model is trained by maximizing the probability of the given token sequence $(x_1, x_2, \ldots, x_{(t-1)})$ in a specific context: $P(y|X)=P(y|x_1,x_2,\ldots,x_{(t-1)})$, where t is the current position. Incorporating the chain rule, the conditional probability can be computed with the product of probabilities in every position: (Chang et al., 2023)

$$P(y|X) = \prod_{t=1}^{T} P(y|x_1, x_2, \ldots, x_{t-1})$$

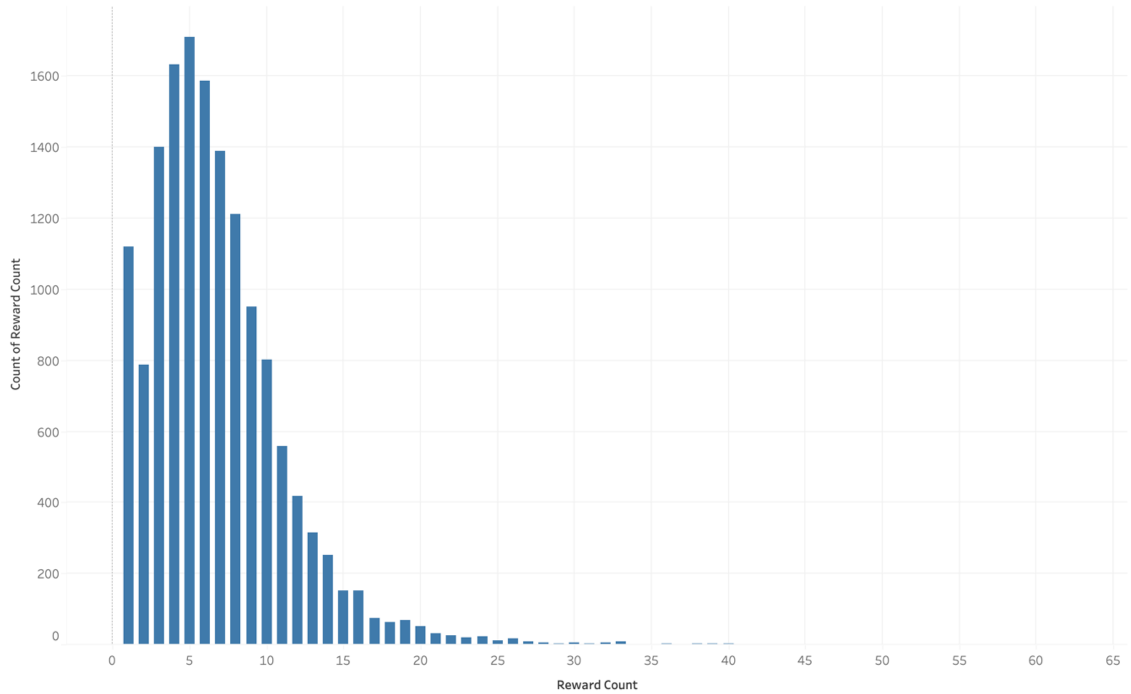
Where T is the sequence length. LLaMA adopts the same decoder-only architecture as GPT-3 but introduces key modifications for improved efficiency and performance. It employs SwiGLU activation instead of ReLU, Rotary Positional Embeddings (RoPE) rather than absolute positional encodings, and Root Mean Square Layer Normalization (RMSNorm) instead of standard layer normalization. These architectural adjustments help improve training stability, enhance contextual understanding, and reduce computational cost. These are some key architectural components:
- LLaMA's self-attention mechanism efficiently processes text by dynamically assigning attention weights to different tokens, allowing it to learn contextual relationships across long sequences. Instead of relying on a single attention function, multi-head self-attention is used, where each attention head captures different aspects of token relationships. Queries, keys, and values are projected into smaller dimensions before computing attention, which improves the model's ability to capture complex patterns while reducing computational overhead.
- To retain word order information, LLaMA incorporates Rotary Positional Embeddings (RoPE), which differ from traditional absolute positional encodings. RoPE encodes both absolute and relative positional information, improving the model's ability to generalize across longer sequences and maintain coherence in text generation.
- Another key architectural feature is efficient tokenization. This implies to convert a sequence of text into smaller parts. LLaMA uses SentencePiece Encoding, a subword-based tokenization approach that enhances vocabulary flexibility and reduces out-of-vocabulary issues. Unlike word-based tokenization, which struggles with rare or unseen words, subword tokenization enables the model to process words more efficiently, improving generalization across multiple languages.
- Additionally, data preprocessing and cleaning play a critical role in model performance. LLaMA benefits from noise removal, deduplication, and outlier handling during training. Deduplication prevents overfitting to repeated patterns and ensures diverse learning. Studies have shown that removing redundant data improves the model's ability to generalize to new inputs.
Overall, LLaMA's transformer-based architecture, combined with optimized data preprocessing, RoPE-based positional embeddings, and efficient tokenization, enables it to achieve state-of-the-art performance in various NLP tasks, including crowdfunding text analysis. This section is primarily based on the work of Minaee et al. (2024), which outlines LLaMA's architectural components, including its self-attention mechanism, positional embeddings, and tokenization strategies. The figure below shows all the components of the architecture and training. More information from Meta can be found in the paper from Grattafiori et al. (2024).



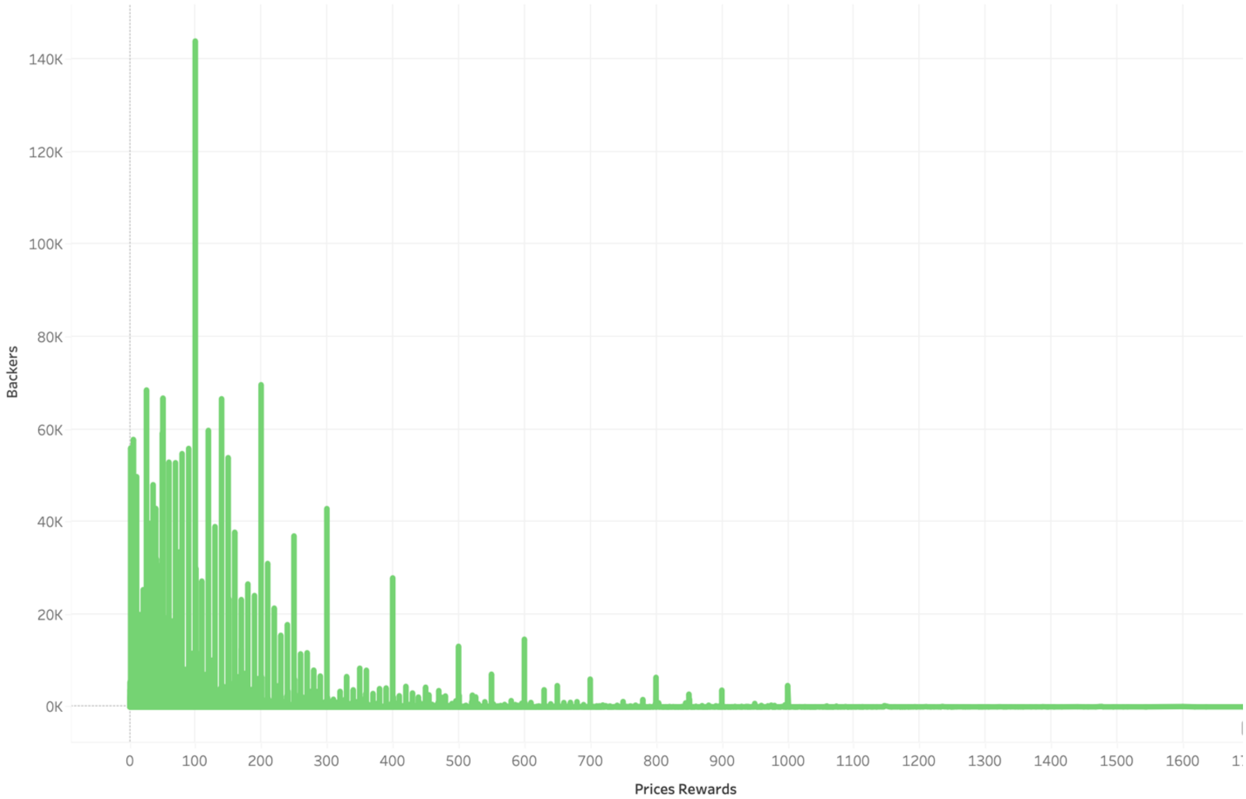
Appendix E:
Extra Figures Data Visualisation



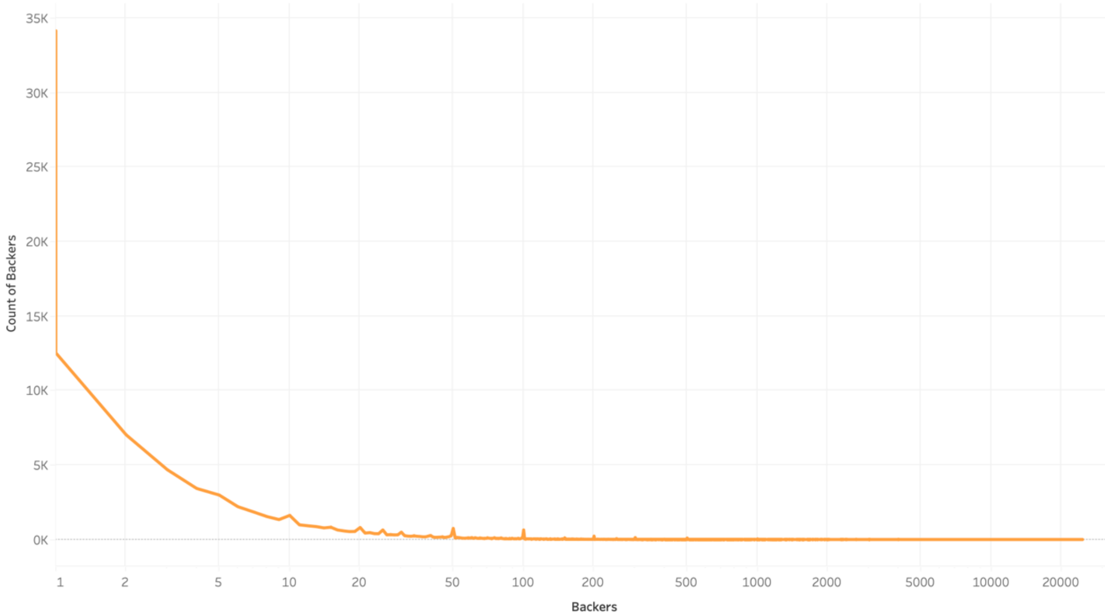
Average Funding Success per County. Colours are decorative



The number of rewards distribution



The backers per prices of rewards



The distribution of backers