Explainable Reinforcement Learning (XRL) in Finance and in a Low-Predictability Betting Game

MARINA-ALEXANDRA MURTAZA, University of Twente, The Netherlands

1 ABSTRACT

In dynamic domains such as finance, where predicting outcomes is essential, introducing reinforcement learning (RL) has shown considerable potential but it remains largely unexplored due to the limited number of practical applications available and field limitation as of now. A key challenge in financial decision-making comes from the complexity and low predictability of financial systems, making it difficult to understand, evaluate, and trust these decisions. This is particularly relevant when integrating RL applications, as they mostly operate as black-box models, which lack transparency. The application of Explainable AI (XAI) and its techniques in RL comes as a promising solution. This research will analyze why and how XAI methods are used in the financial field in order to underline benefits and actual progress and results with a focus on their contribution to decision-making, risk assessment and regulatory compliance. Furthermore, it will make use of a simplified betting game as a case study in order to explore how Explainable Reinforcement Learning (XRL) can be used to improve explainability by examining and explaining decisions made by RL agents in an unpredictable financial environment.

KEYWORDS

Reinforcement learning, Explainable AI, XRL, black-box Machine Learning, explainability, low predictability, betting game

2 INTRODUCTION

The process of decision-making across fields such as finance, health care, or even autonomous systems, has been undeniably improved and disrupted by the introduction of artificial intelligence. Introducing AI in finance has led to the reduction of manual work when dealing with fraud detection, risk assessment, personalized financial services or algorithmic trading, while also improving customer support through chat bots or AI-driven models that enhance efficiency and reduce human error in contract analysis. In the field of healthcare, introducing AI led to the possibility of analyzing patient data and suggesting personalized treatments for them. For instance, "Nature Medicine" conducted a study that proved how Google's AI performed better than radiologists when faced with the task of detecting breast cancer [4]. Regarding autonomous systems, Tesla's Full Self-Driving software is possible thanks to the introduction of AI in decision making. It works by collecting and processing real-time data from sensors and cameras and making decisions regarding driving in a split-second. Studies show a significant decrease in accidents caused by human error during the testing of autonomous vehicle testing [20]. The considerable advancements in AI, as well as relevant subfields such as machine learning, have revolutionized these industries, introducing innovative methods that optimize decision-making and address complex problems. A notable outcome of such developments is Reinforcement Learning, an approach to machine learning that is concerned with how an agent makes a decision in an environment in order to maximize the expected outcome and reward [18]. RL has been present and studied for several decades, with its modern form based on the theory of Markov decision processes (MDP) emerging since the 1980s [8]. This theory enables the prediction of all future states and expected rewards through only the current state and action, not taking into consideration previous occurrences [17]. Furthermore, most RL applications either represent a MDP or a partial MDP.

RL has been integrated and used in many financial applications such as market making, portfolio management and optimization, optimal execution, option pricing and hedging [19][7]. This approach aided and benefited this field, where the previously used mathematical models were not able to encompass the complexity of this dynamic and volatile domain. This leads to under-performance and potential financial losses [7]. However, creating other much more intricate methods of approach with the help of reinforcement learning that give better and more efficient results, inevitably leads to the computationally untraceable, unexplainable, and recurring AI problem of the "black box". Data is being fed to the algorithm that will provide an output but how that output was decided and made is not explained nor traceable, especially when dealing with complex Deep Reinforcement Learning methods such (DRL) as Neural Networks [5]. This becomes a critical problem in a field such as finance where a DRL is still in the infancy for these "high-stakes tasks" [14]. A bug or error in reinforcement learning (RL) algorithms can cause substantial and devastating financial losses, highlighting an essential need for understanding how an RL agent makes a decision and breaking the black box. If ethics are addressed, a lack of transparency can lead to significant consequences regarding sensitive financial applications such as loan approval, credit scoring or insurance underwriting. For example, in the case of an RL agent rejecting a loan application, it is essential that the decision-making process is explainable in order to provide the applicant with enough relevant information behind the rejection. More attention is critical in regulated industries where fairness and accountability are a priority. Regulations such as the General Data Protection Regulation (GDPR) in the European Union or the Fair Lending Act in the United States require that automated decision-making processes be transparent and justifiable, particularly in cases where individuals' financial futures are at stake. The algorithm must explain the reasoning behind every decision in order not to undermine the consumer's trust and avoid violation of regulatory obligations. Such unjust rejection may

TScIT 42, January 31, 2024, Enschede, The Netherlands

result in the applicant filing a complaint and in the absence of a clear explanation of the decision, the financial institution could face legal consequences or regulatory penalties. To effectively address this challenge, Explainable AI (XAI), a subfield of AI, focuses on enhancing core principles like transparency and interpretability, thereby enabling human users to gain deeper insights into the machine's decision-making process [1]. XAI has established itself as an essential tool for finance scholars, industrial institutions, and strategists because of its accessible and logical methods, along with its clear computational processes, which have enabled numerous financial studies [2].

The purpose of this research is to explore and highlight the advantages of how explainable AI methods can be integrated into RL in low-predictability, dynamic, and complex environments to enhance transparency and understanding of these elaborated algorithms and methods. By analyzing the state of the art and also providing a practical example of the use of XRL methods, this paper aims to address the imperative problem between the potential efficiency and advantages that DRL methods can pose in finance and the need for trust and explainability of the processes and reasoning done by the RL agents in different financial applications.

The problem statement and objectives leave to the following research question:

> How do Explainable Reinforcement Learning (XRL) techniques impact the decision-making explainability of RL agents in the dynamic field of Finance and more specifically a low-predictability betting game?

This extended research question can be broken further into the following sub-questions:

(1) How do Explainable Reinforcement Learning (XRL) techniques impact the decision-making explainability of RL agents in the field of Finance?
(2) "How do Explainable Reinforcement Learning (XRL) techniques affect the explainability of RL agents' decision-making in a low-predictability betting game designed to simulate a simplified stock market?

This research aims at investigating advantages that XRL techniques can pose when analyzing dynamic and unpredictable realistic financial environments.

3 METHODOLOGY

This research will require multiple steps. First, literature research will be performed to obtain a general understanding of explainable reinforcement learning (XRL) in finance and existing methods. Considering that the existing literature presents complex financial systems (for example: stock market, portfolios) that refer to real-life scenarios with real hard-to-understand, analyze and explain data, the second part of the paper will focus on XRL on a case study of a simple betting game. The game provides an idealized scenario where agents must make decisions under uncertainty (trying to mimic a stock market environment). Consequently, LIME, an XRL model, will be implemented to showcase the use of explainable methods to understand and explain the decision-making process of the betting game's RL agent. The objective of this case study is to demonstrate and provide a practical application of XRL on an RL observer and to analyze the resulting outcomes.

The betting game, on which my experiments ar based on, is part of an ongoing project led by my supervisor, involving other students. The game implementation (game logic and generators) for the case study was done by Mette Weisfelt while the RL observer and LIME implementation was implemented by me.

4 LITERATURE REVIEW

Since Machine learning and reinforcement learning are experiencing expansion in many domains such as finance, expansion attributed to the high performance and great potential they provide, strong concerns arise about the opacity of this disruptive technology being voiced by stakeholders. As these algorithms become increasingly powerful and adaptable, they also become more opaque, rigid, and resemble black boxes, making their decision-making processes difficult to understand [18]. Finance is fundamentally concerned with the management, allocation, and investment of monetary resources, so when these systems are entrusted with any type of money management, the stakes become extremely high, and transparency and trust become vital. Eliminating the opaqueness becomes imperative, considering the gravity of monetary, ethical and safety risks when entrusting an RL agent with such sensitive operations. XAI aims at solving such problems by creating the means by which the decision process and decision points of an agent can be explained and made transparent to be more easily assessed and fixed in case of potential flaws. XAI methods can be categorized into two main types to aid in understanding and organizing XRL techniques (a subset of XAI): intrinsic and post hoc [12]. The intrinsic category refers to the construction of a model that is already compliant with the interpretability requirement. In the post-hoc category, an additional interpretable model or explanation technique is created to explain the existing non-interpretable model [11].

A mix of both intrinsic and post-hoc models can be found in existing XRL financial applications documentation. Existing work validates the benefits of using RL and XRL in different finance domains [3]. A DNQ (Deep neural Network) agent, which is commonly used in stock trading applications, uses a RL model that is too complicated to interpret and opaque. To this agent, a post-hoc explainability technique called SHAP was applied to help break down and understand the agent's reasoning. The method worked with two real stock datasets and explained the predicted outcomes and rewards for buying and/or selling stock.

In another field of finance, Misheva et al. (2021) [10] explored the use of explainable AI (XAI) methods: SHAP and LIME, in credit risk management to increase the transparency and explainability of machine learning models used in credit risk management. Since loans and credit approvals are high-risk decisions, entrusting RL models to do this critical operation would be impossible without proper and extensive reasoning and regulatory acceptance. Wrong decisions in this field could equate to significant loss of funds and reputation (lawsuits, ethical problems) of the firms or companies that would utilize these models. The study demonstrated that post-hoc explainability techniques could provide local and global interpretability for credit scoring models, making AI-based credit risk assessments more reliable and compliant with regulatory requirements. The research applied XRL techniques to ML models trained on peer-to-peer lending data from Lending Club, finding that SHAP effectively highlighted global feature importance. Furthermore, LIME provided detailed explanations for individual loan approval decisions. The study's findings reinforce the necessity and importance of explainable models in financial applications, as they facilitate trust and accountability in RL agents decision-making and make possible the integration of these machine learning models in these complex fields.

As seen in Lundberg and Lee (2017), it is much easier for humans to understand the factors that influence an agent decision with the help of SHAP values. Their framework brings into discussion the challenge of model interpretability in machine learning and AI by giving a compatible and explainable metric for understanding complex predictions across multiple domains. When making informed decisions, it is essential for financial experts to visualize SHAP values, for example plots of rewards, in order to show how past actions affect the process of decision-making. Moreover, this perspective helps identify key days and actions with the most significant impact on predictions [9]. The visualization of SHAP with the help of plots of rewards or possible influence of specific dates or periods, displayed how past actions contributed either negatively or positively to the decisions made by the agents. Key days and actions that contributed most to predictions can be explained and decisions to buy/sell can be broken down in order to help experts in making a more informed decision. Furthermore, this XRL model can also break down and find possible flaws or weak reasoning points of DNQ agents. Both intrinsic and post-hoc XRL models can be applied to portfolio management optimization and construction. Attention layers, Multi-head LSTM and Explainable Policy Network are intrinsic models that were integrated into the RL agent to be able to directly assess feature importance and weights. This made it possible to analyze more in-depth portfolio features and how these features influence and contribute to portfolio-related decisions. The post-hoc method used is Q-value analysis (with visual representation) resulting from attention layers to interpret which features were the most influential in specific decisions made post-training. As explained by Mao Guan and Xiao-Yang Liu (2021), when visualized, Q-value analysis can provide transparency in understanding which variables most significantly affected a model's decision, which makes it easier when interpreting factors that influence portfolio management decisions [6]. Since the RL model outperformed traditional techniques, the explanations offered by the XRL methods can offer reasoning and transparency in how and why this performance was possible.

The analysis of the articles indicates that these models can improve trust in RL methods and in the agent's reasoning as a whole by debugging. Furthermore, it also can bring real benefits to financial experts by highlighting properties or patterns that could not have been visible with other traditional financial methods. Mao Guan and Xiao-Yang Liu work highlights the ability to uncover hidden patterns in financial data. The study further emphasizes the extent to which XRL is capable of providing meaningful insights into financial systems, a process which would be significantly more complicated using only traditional methods of interpretation such as technical analysis and fundamental analysis, both being commonly used to interpret market trends and stock performance. While these previously mentioned methods have been valuable for decision-making, oftentimes they do not allow insight into complex, non-linear patterns in data [6].

Considering that the existing literature presents complex financial systems (stock market, portfolios) that refer to real-life scenarios with real hard-to-understand, analyze and explain data, the next section of the paper will focus on a case study of an idealized, simple betting game. It will very lightly mimic the stock market and the impact of different XRL techniques on a RL agent will be analyzed. The game and RL agent are both implemented in Python.

5 CASE STUDY

The case study analyzes the decision-making and reasoning of an RL agent developed for a Python-based betting game.

5.1 Components

The betting game consists of a generator and an observer. The generator produces sequences of numbers based on different distributions:

- **Linear distribution:** Generates linearly increasing sequences.
- **Normal distribution:** Produces values based on a Gaussian distribution.
- **Cellular automata distribution:** Creates binary sequences.
- **Stock data:** Uses historical stock data (e.g., Apple stock) as input.

The observer, implemented as an RL agent, predicts whether the next value will be higher or lower, earning rewards for correct predictions and penalties for incorrect ones.

DNQ RL agent: the agent was implemented using Deep Q-learning. The algorithm used combines Q-learning with a deep neural network to approximate the Q-function (the action-value function). The deep neural network is a complex network that mimics neurons of a brain and goes to thought computations that cannot be understood just by analysing or running the code [16]. They are referred to as "black boxes" due to elevated difficulty in the interpretation of the inner states of the models [16] which hinders the possibility of understanding the reasoning behind certain, if not any, decision the agent makes. .

The deep neural network takes input features that are mathematically relevant for predicting the trend of the distributions [15]. These features include:

- Rolling mean,
- EMA (Exponential Moving Average),
- Current value,
- Trend slope,
- Momentum,
- cons_pos_value (number of consecutive positive rewards).

The action states are:

- 1: Bet higher,
- 0: Bet lower.

LIME: Ribeiro et al. [13] addresses the problem of explaining opaque machine learning models by proposing LIME . This model-agnostic post-hoc XRL technique can be used locally, or globally to understand how the model behaves by giving user-friendly visual explanations. It builds simple surrogate (linear) models around these complex models in the local vicinity of each prediction in order to be able to offer a n interpretable representation that is locally faithful to the classifier. Since analysing single predictions would not be sufficient to explain how a model works, multiple LIME weights will be aggregated. By aggregating the LIME weights, the global behaviour of the Reinforcement Learning agent will be studied to get a general understanding of how important each feature is in the general decision process.

The vizualisation of the LIME results will be displayed using different charts:

- **Beeswarm plot:** combines weights and feature values, offering a comprehensive visualization of both feature contributions and values.
- **Global feature importance:** the absolute values of feature weights are averaged over all data points to detect most important features.
- **Trend plot of each feature:** feature weights are plotted alongside the corresponding feature values to detect patterns or relationships.

5.2 Result and analysis

Ten episodes of training were conducted multiple times, each consisting of 2,000 movements. After training the RL model with different generators and evaluating it using LIME methods the results provided several insights.

All the global feature importance plots values did not change much from first to last episode even if the RL agent was not performing optimally to get a substantial cumulative reward (not even in the easy case of the linear generator where the observer just has to learn to bet 1-higher every time). The features chosen, even if they represent mathematically correct and technical indicators, fail to deliver more than average results (or negative results) and seem to be used erratically and inconsistently by the DNQ agent. This is accentuated especially in the sock data observer where the values have a higher degree of randomness. This situation can indicate an incorrect trade-off between exploration and exploitation or hyperparameter values in the model but also problems with feature selection.

The inefficient feature selection strategy is also emphasized by the prevalence of the current value feature as one of the most influential in decision making across all generators, even if it represents just a numerical value and does not indicate anything about the generator direction. This can put noise in the decision and negatively influence the observer.

The importance of other features is also discrepant across the different generators, again questioning their validity. Nonetheless, even if the features utilized do not yield optimal results and generate inconsistencies in decision-making, they make a positive contribution to the decision-making process and lead to positive cumulative rewards (from 100-300 positive cumulative results in the presented training context) for most generators. To validate this, the agent trained just with current value as state feature demonstrated completely inefficient learning with many negative cumulative rewards across all generators.

Since different generators present different distributions, different features become important in the decision making; therefore, a deeper analysis is needed to assess what features can be changed to provide better results and accommodate all types of generators. Consequently, the inconsistencies in the learning strategies adopted by the observer in different training sessions (as shown by the beeswarm and feature trend plots) highlight learning issues that are subsequently demonstrated by inefficient performance when using a random distribution generator with the real stock data. This concludes that the observers and its selected features have a positive impact on a betting game with generators that do not exhibit much randomness but it is not appropriate yet for a betting game with random unpredictable data.

Plots examples to justify the results can be found in Appendix A.

Below the analysis of the LIME plots is provided for every generator.

(1) **Linear generator:** The rl agent is performing the best on the linear generator. The final scores (cumulative rewards) are growing linearly, but some confusion in the predicting is present even for a linear distribution (the last score presented and the highest one was 430 and there are 2000 movements). The global feature importance graph shows three main important features: rolling mean, EMA and current value. The other three do not contribute much to the decision (momentum only contributes to one). The beeswarm and individual feature trend plots shows that the LIME weight contribution stays approximately the same but the feature values have a tendency to swap LIME feature weights between them in different training sessions (opposite feature weights between training sessions). Momentum stays the same with LIME weight zero across training sessions.

(2) **Normal distribution generator:** The final scores (cumulative rewards) are positive but seem to remain stuck at the 100-200 values. The global feature importance graph shows three main important features: EMA, current value and momentum. Slope also contributes to the decision, although not significantly and the last two do not have a real impact. Here the beeswarm and individual feature trend plots maintain similar LIME weight across training sessions for a feature value aside from the EMA feature that swaps lime weights and values, and current value that behaves inconsistently.

(3) **Cellular automata generator:** The final scores (cumulative rewards) are positive but also remain stuck at the 100-200 values. The global feature importance shows that: EMA, current value, momentum, and consecutive positive rewards feature influence most decision-making. Rolling mean and trend slope have low absolute mean weights values. the beeswarm and individual feature trend plots maintain the same, constant distribution of values and weights across different training sessions.

(4) **Stock data generator:** The stock data generator poses a challenge for the RL agent. The final scores across training episodes are both positive and negative indicating that the agent is not learning effectively. The global feature importance of each feature changes with different training sessions,

making it difficult to tell which features are the most important for this generator. The beeswarm is also looking different in different training sessions and for the trend of the features it is observed that the trend for current value and EMA stay the same proportionally but have different weights. The trend plots of the other features have changed similar to the ones from the other generator plots.

Drawbacks and Future Work: Lime provides a faithful single prediction explanation of the importance of features, so when aggregating multiple feature weights from different predictions, some inconsistencies can arise (LIME operates locally, and global patterns might not be captured perfectly)[19]. In addition, LIME's most important feature can cause either wrong or correct predictions. Both features with high positive or high negative LIME weights have a large impact on the predictions. For future work, more XRL methods can be applied (SHAP, silency maps). The predictions can be studied locally to check the influence of the specific features in correct and incorrect predictions since now the focus was analyzing the impact of the features in decision making regarding the correctness. Furthermore, a more in depth debugging can be done to the RL agent by removing non-relevant or noise inducing features and replacing them with relevant ones.

6 CONCLUSION

This paper has investigated the integration of Explainable Reinforcement Learning techniques into Reinforcement Learning (RL) models to overcome the current limitations of transparency and confidence in the financial decision making process. The research highlights the weaknesses of traditional "black-box" RL models, lack of transparency, and the potential risks in high-stakes environments such as finance, and thus the need to apply XRL techniques for insight into model reasoning.

This case study applies a simplified betting game to show how XRL techniques, notably LIME, may be used to better understand the RL agent's decision-making process in a variety of environments from linear distributions all the way through to more complex stock data generators. Furthermore, the results reveal differences in feature importance and confirm the need for more developed feature selection and tuning of the RL agents. These observations point towards the importance of the role XRL can play in identifying and preventing model weaknesses, such that RL applications in finance are not only effective but also understandable.

Though the case study provides meaningful insights, limitations such as the local scope of LIME and RL agent's poor performance suggest the need for further future research. Integration of additional XRL methods, improved feature selection, and performing a more in-depth analysis of predictions are further needed to improve the RL agent performance and decision-making transparency.

In conclusion, this research emphasizes that Explainable Reinforcement Learning is not only a technical requirement but a requirement for establishing trust, enhancing decision-making process, and ensuring compliance in dynamic financial environments. As RL advances and evolves, the intersection with XRL offers a promising solution to bridge innovative machine learning techniques with the explainability required to enable responsible and ethical practice in finance.

6.1 Limitations and future research

Time constraints were the biggest limitations considering the broadness of the research. A more in-depth literature analysis could be done to get a wider understanding of the "state of the art" of XRL in finance with a separate focus on different fields of this broad domain. The case study presents a global analysis of the decision-making patterns of the RL agent, but there is a need for a local analysis to enhance understanding and more accurately assess which features influence which outcomes. Furthermore, applying more XRL models could be use for better analysis and comparison of results. Future research could dive into the optimization of the RL agent based on the feedback and results of the XRL techniques and explore their relationship.

A APPENDIX

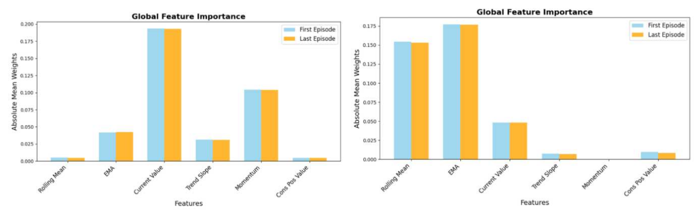

Fig. 1. (Global feature impotence graphs for normal and linear generators respectively: emphasizing current value as a influencial feature and the importance of different values across different generators)

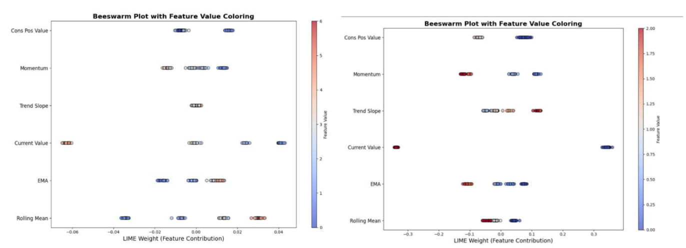

Fig. 2. (Beeswarm graphs of stock and cellular generators: emphasizing different LIME wight distributions of different feature values across generators)

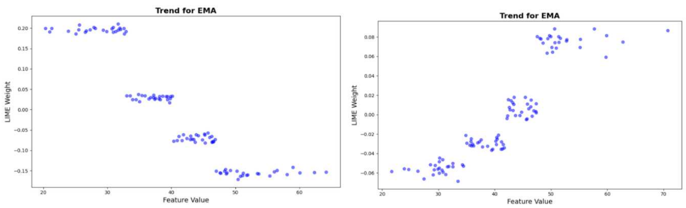

Fig. 3. (Trend graph feature for EMA in two separate training occasion with a normal generator: there is a noticed shift of LIME weights compared to different feature values emphasizing the change of learning strategy on the same generator)

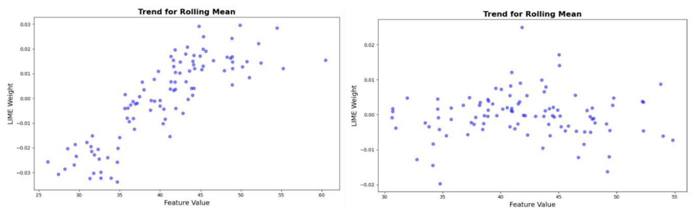

Fig. 4. (Trend graph feature for Rolling Mean in two separate training occasion with a normal generator: there is a noticed inconsistencies of LIME weights compared to different feature values emphasizing irregular learning patterns)

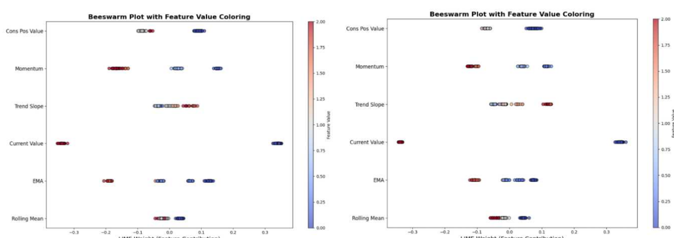

Fig. 5. (Beeswarm graphs for the cellular generator in two separate training sessions: there is a consistency between learning patterns emphasizing that the suboptimal consistent cumulative reward across episodes can be caused also by the exploration exploitation trade-off or hyperparameters values)

REFERENCES

[1] Amina Adadi and Mohammed Berrada. 2018. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* 6 (2018), 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

[2] J. O. Adelusi and F. Zhang. 2023. Explainable AI (XAI) in Financial Model Transparency: Bridging Accountability and Interpretability in Modern Finance. *ResearchGate* (2023). https://www.researchgate.net/publication/387133832

[3] P. Sequeira D. G. Cortés and B. Rosman. 2024. Portfolio Construction Using Explainable Reinforcement Learning. *Expert Systems* 41, 1 (2024). https://onlinelibrary.wiley.com/doi/10.1111/exsy.13667

[4] Scott M. McKinney et al. 2020. International Evaluation of an AI System for Breast Cancer Screening. *Nature Medicine* 26 (2020), 908–914. https://www.researchgate.net/publication/338332372_International_evaluation_of_an_AI_system_for_breast_cancer_screening

[5] Wojciech Samek et al. 2019. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. *IT Professional* 21, 3 (2019), 16–20. https://doi.org/10.1109/MITP.2019.291718

[6] Min Guan and X. Y. Liu. 2021. Explainable Deep Reinforcement Learning for Portfolio Management: An Empirical Approach. *arXiv* (2021). arxiv.org/abs/2111.03995

[7] N. Gârleanu and L. H. Pedersen. 2016. Dynamic Portfolio Choice with Frictions. *Mathematical Finance* 26, 1 (2016), 3–41. https://onlinelibrary.wiley.com/doi/full/10.1111/mafi.12382

[8] H. Holger and B. Schölkopf. 2012. Kernel Methods in Machine Learning. In *Machine Learning and Knowledge Discovery in Databases*, A. Gionis M. Sebag J. L. Balcázar, F. Bonchi (Ed.). Springer, 145–168. https://link.springer.com/content/pdf/10.1007/978-3-642-27645-3.pdf

[9] S. M. Lundberg and S. Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*. https://arxiv.org/abs/1705.07874

[10] Branka Hadji Misheva, Ali Hirsa, Joerg Osterrieder, Onkar Kulkarni, and Stephen Fung Lin. 2021. Explainable AI in Credit Risk Management. *arXiv preprint* arXiv:2103.00949 (2021). https://arxiv.org/abs/2103.00949

[11] L. Sonenberg P. Madumal, T. Miller and F. Vetere. 2023. Explainable Reinforcement Learning through a Causal Lens: A Survey. *Commun. ACM* 66, 3 (2023), 52–61. https://dl.acm.org/doi/full/10.1145/3616864

[12] Elias Puiutta and Eric Schneider. 2020. Explainable Reinforcement Learning: A Survey. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*. 4813–4819. https://doi.org/10.24963/ijcai.2020/666

[13] R. Ribeiro, S. Singh, and C. Guestrin. 2016. Why Should I Trust You? Explaining the Predictions of Any Classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. 1135–1144. https://dl.acm.org/doi/abs/10.1145/2939672.2939778

[14] M. Vishal S. Kumar and V. Ravi. 2022. Explainable Reinforcement Learning on Financial Stock Trading using SHAP. *arXiv* (2022). https://arxiv.org/abs/2208.08790

[15] D. J. Siami-Namini and N. Tavakoli. 2021. Stock Prediction Based on Technical Indicators Using Deep Learning Model. ResearchGate. https://www.researchgate.net/publication/354507095_Stock_Prediction_Based_on_Technical_Indicators_Using_Deep_Learning_Model

[16] M. K. Singh, P. K. Gupta, and R. Singh. 2023. Feature Engineering for Predictive Modeling in Finance. In *Artificial Intelligence and Machine Learning for Business Applications: A Data-Driven Approach*. Springer, 125–142. https://link.springer.com/chapter/10.1007/978-981-19-7784-8_10

[17] Richard S. Sutton and Andrew G. Barto. 1998. *Reinforcement Learning: An Introduction* (1st ed.). MIT Press. www.incompleteideas.net

[18] Richard S. Sutton and Andrew G. Barto. 2018. *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press. http://incompleteideas.net/book/bookdraft2018jan1.pdf

[19] R. Tibshirani and J. H. Friedman. 2024. Explainable Machine Learning: Methods and Challenges. *Annual Review of Statistics and Its Application* 11 (2024), 1–26. https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-112723-034423

[20] Waymo. 2020. *Safety Report*. Technical Report. https://downloads.ctfassets.net/sv23gofxcuiz/4gZ7ZUxd4SRj1D1W6z3rpR/2ea16814cdb42f9e8eb34cae4f30b35d/2021-03-waymo-safety-report.pdf

USE OF AI TOOLS

During the research process, AI tools were utilized for code completion and grammar refinement. All AI-generated outputs were thoroughly reviewed, evaluated, and edited by the author. The author assumes full responsibility for the content and conclusions presented in this work.