AI-Powered Sales Forecasting for Event Catering: A Data-Driven Approach

Wouter Weijnen

University of Twente

P.O. Box 217, 7500AE Enschede

The Netherlands

Abstract: The Development of Artificial Intelligence (AI) has revolutionized many aspects of business. One such field is that of sales forecasting. Accurate sales forecasting provides businesses with many benefits, such as inventory and supply management optimization as well as increased operational efficiency, decreased purchasing cost and reduced obsolescence. A knowledge gap exists when considering AI sales forecasting in event catering. The field of event catering remains underrepresented in literature, and the possibilities for AI sales forecasting for event catering remain therefore relatively unknown. This study focuses on exploring the possibilities of AI sales forecasting for event catering by attempting to develop an AI sales forecasting model for the catering department of a football club. Doing so provides empirical insights into the performance, limitations and feasibility of AI forecasting for event catering. The event catering sector typically has infrequent days of operation, which results sparsely available data since there are limited opportunities for data collection. This results in several complications for the development and evaluation of AI sales forecasting models in the event catering industry. The findings and results of this paper indicate that these challenges are very relevant in the development of AI sales forecasting in this industry, but also show that AI sales forecasting is possible for event catering. These insights are relevant since it shows the capabilities and limitations of AI forecasting in this setting. Furthermore, the findings of this research demonstrate the practical impact of the prediction models. Thereby, this research confirms existing theories and adds to it in discovering the practical implications of existing theories. Furthermore, the exploration of the possibilities in AI sales forecasting for event catering contributes new insights and provides research for the underrepresented field of AI sales forecasting for event catering.

Graduation committee members:

1st Supervisor: Matthias De Visser

2nd Supervisor: Robin Effing

Keywords: Event catering, Artificial intelligence, Machine learning, Sales forecasting.

During the preparation of this work, the author used Chat-GPT in order to refine the bachelor thesis by checking grammar and spelling. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the work.

1 Introduction

1.1 Problem statement

Recent developments in the field of Artificial Intelligence (AI) and Machine learning have enabled ever more possibilities in various domains (Jordan & Mitchell, 2015). One such development is the increased accuracy and automation of demand and sales forecasting (Douaioui et al., 2024). Accurately predicting future sales can help companies to optimize inventory and supply management as well as increase operational efficiency, decrease purchasing cost and reduce obsolescence (Groene & Zakharov, 2024). Even though the field of AI is a relatively new development, its adoption has accelerated rapidly, and many sectors are integrating AI developments, this includes the field of sales and demand prediction.

However, a research gap exists when considering the application of AI developments in sales forecasting for event catering, even though this sector could benefit strongly from the new technology. A number of AI prediction models have been developed and have shown capable of accurately predicting sales. AI forecasting models can discover complex, non-linear patterns and can reduce forecasting errors (Douaioui et al., 2024). In today's ever changing and quickly developing business landscape, accurately and timely predicting demand and sales enables companies in various fields to gain critical knowledge, which has strong short- and long-term benefits (Doganis et al., 2006). For companies in the food industry, sales prediction is particularly important as food products typically have a short shelf life, which, when becoming obsolete, can cause a loss of income in both shortage and surplus situations (Tsoumakas, 2018). In terms of delivering appropriate service levels, accurate sale prediction is a prerequisite. Furthermore, accurate sales and demand prediction can strongly influence procurement efficiency and can help avoid food waste (Groene & Zakharov, 2024). Therefore, accurately predicting sales is highly important as it enables various benefits and allows for efficient operations.

Demand forecasting for event catering has proven especially challenging, due to event specific challenges such as a long setup, time-varying demand and complex customer food preference (Dolgui et al., 2021). These challenges are a result of unique factors in the event catering branch, factors that distinguish the event catering sector from regular food service industries. Unlike regular restaurants or catering operations, which operate on a daily basis, event catering occurs sporadically. This sparsity results in most standard forecasting techniques being impractical; it also leads to challenges on measuring forecast accuracy, model selection, and forecast aggregation (Turkmen et al., 2020). The event sector also requires tailored prediction models; as traditional prediction models have proven unsuited. Furthermore, events face complications in the AI modelling process due to the relatively limited availability of historical data, a result of the sporadic occurrence of events (Grolinger et al., 2015; Pang & Wang, 2024). The sparsely available data might result in overfitting difficulties maintaining a sufficiently large Events Per Variable ratio (Alwosheel et al., 2018).

This raises the question whether AI sales forecasting is possible within the event catering sector, given that it has far less historical data available and tends to be more complex due to its sporadic occurrence. AI powered prediction models have been developed in many industries, especially in supply chain and retail. (Douaioui et al., 2024). However, for event catering in specific, no clear research has been identified. For regular catering services, like restaurants or canteens, a number of studies have already been conducted, such as a sales forecasting model for a student canteen (Rodrigues et al., 2023). Such concepts prove that sales forecasting models for catering are possible, which further highlights the lack of such models in event catering.

1.2 Research Objective

Due to the existing knowledge gap, the event catering sector is yet to take full advantage of the recently developed new technologies such as AI. If developed, an AI prediction model can strongly influence the operational performance of companies in the event catering business. Therefore, this study revolves around exploring the opportunities in developing a data driven AI prediction model, attempting to accurately predict the sales of a range of items in the event catering setting. In particular, the case of a regional football club will be taken into consideration, with the goal of developing a model to accurately predict the sales of a range of selling points in the stadium. The exact capabilities of AI in this field are unknown and therefore it is hard to predict what exactly the possibilities are. By exploring this topic further this should become clearer, with the final goal of developing a data driven demand forecasting model that, with the input of certain variables, can accurately predict the demand and sales of a future match.

1.3 Research question

Now that the topic and goals have been identified, a research question can be shaped. As this study will be aimed at the development of a data driven AI sales forecasting model, focused on the catering sales of a local football club for future matches, the research question should revolve around this topic and should clearly reflect on the set goal of exploring this topic and the development of such a forecasting model. Therefore, the following research question has been identified:

How accurately can AI sales forecasting models predict the catering sales for sport events, based on historical data and event specific variables?

Answering this research question should help fill the existing knowledge gap.

1.4 Sub questions

In order to answer the research question, several sub questions will be introduced. Answering these will allow for a definitive answer to the research question to be formulated. Each sub question will focus on a specific prerequisite for a final answer to the research question. The identified sub questions are: Why is sales forecasting relevant? What kind of sales prediction models exist? Which prediction models are suited for demand forecasting for event catering? Which input variables are suited for AI sales forecasting in event catering? How will data analysis be conducted and how can AI sales prediction models be evaluated? Which prediction model is most accurate for event catering sales forecasting? And finally, how do AI forecasting models perform compared to current methods? Formulating an answer to these questions will allow for the overall research question to be answered.

2. Theoretical Framework

2.1 Why is sales forecasting relevant?

The ability to predict future sales is widely sought after. Sales forecasting belongs to the important aspects of the business management process as accurately forecasting sales plays a major role in the allocation of resources, marketing, and finance. Furthermore, it influences the decision-making process which shapes the way companies provide products and services to consumers in the customer relationship management horizon (Hyndman & Athanasopoulos, 2019). A number of different methods have been developed, ranging in complexity, and accuracy.

2.2 What kind of sales prediction models exist?

Attempting to predict future sales is nothing new, (Yule, 1909) already included a version of the moving average. Since then, a large number of developments have been made in the field of demand forecasting, which results in a large number of forecasting techniques (Ahaggach et al., 2024). For the purpose of this study, these techniques will be divided into traditional models and AI and machine learning models.

2.2.1 Traditional methods

Many advancements have been made in the field of sales forecasting. Harrison (1965) suggests two main different types for forecasting; short term and seasonal forecasting. Brown's method of Adaptive Forecasting is the most satisfactory for the short term. To illustrate the advancements made in sales forecasting, Brown's model only includes one parameter; and multiple factor forecasting did not deliver superior results (Harrison, 1965). Other mentionable methods of sale forecasting are a (moving) weighted average and models such as Roodman's (1986) demand forecasting regression model. Multiple retail sales forecasting methods have been studied and developed in recent years to improve performance (Gustriansyah et al., 2022). Findings showed that classical sales forecasting methods still dominate in retail forecasting, innovation is needed to develop a new method that minimizes forecasting errors, including case study diversification. (Gustriansyah et al., 2022)

2.2.2 AI and machine learning applications

Next to traditional models of sales forecasting, new models relying on AI and machine learning are being developed. The application of AI and machine learning in sales forecasting mainly involves models that rely on historical data and identify patterns that influence future sales (Ganesan, 2024). Artificial intelligence (AI) is capable of figuring out complex relationships among data, behavioural patterns, classification schemes and can extract useful knowledge from these data (Türkbayrağí et al., 2021). A certain number of different techniques for AI application in sales forecasting exist, such as Regression trees and Artificial Neural Networks. These models operate more autonomously and can deal with a variety of criteria but are also much more complex than traditional models (Tulli, 2020).

2.2.3 AI and machine learning models compared to traditional models

Thanks to recent advances in the field of AI and machine learning, even more is possible. The majority of the traditional forecasting models methods only rely on one forecasting criterion, which is the number of sales taken from historical trends and data, in combination with seasonal effects to predict future sales (Fridley, 2018; Gustriansyah et al., 2019; Pavlyshenko, 2019). In these cases, the company needs to manually make adjustments to the statistical forecasts or rely on an expert (Gustriansyah et al., 2022). By making use of newly developed capabilities of AI many more criteria can be integrated into predictions, without the need for manual corrections. After a case study on sales prediction of smartphones (Biswas et al., 2023) concluded that with the development of Artificial Neural Network (ANN), accurate sales predictions significantly improved. The study showed a significant improvement over traditional statistical linear models such as multiple regressions (Biswas et al., 2023). A large variety of machine learning models have been tested against traditional forecasting models. Results vary but the overall conclusion is that traditional models benefit significantly from simplicity and, for situations where a multitude of factors and criteria are unnecessary, can outperform machine learning models (Cadavid et al., 2018). However, machine learning techniques that are properly implemented outperform most of the traditional forecasting methods (Cadavid et al., 2018).

2.2.4 Event and food related AI forecasting

The applications for AI sales forecasting for event catering remains underrepresented in research on AI applications. The need for accurate sales forecasting is especially valid in this area, given the typically limited shelf life of products. A compounding factor is that events usually run on a specific time and for a short period of time, meaning that supply has to be available at the correct time with little room for delays. Unsold products become obsolete very quickly as there is limited possibility for a later selling moment. Therefore, the application of AI is very promising for this sector, as it could strongly benefit from accurate sales forecasting. After researching the applications of AI powered predicting models on food and beverage sales, it was concluded that AI sales forecasting performed very well. Two major advantages were identified; an increase in prediction accuracy and the creation of a general feeling of relief among managers responsible for repetitive forecasting tasks (Groene & Zakharov, 2024).

2.3 Event catering specific challenges and characteristics

As mentioned previously, the event sector is characterised by infrequent days of operation (Goldblatt et al., 2005). This ultimately leads to a relatively limited availability of historical data to include in the development of the model. Further consequences arise with the selection of AI prediction models, as several models are very data intensive and might not work in combination with a small data set (Brigato & Iocchi, 2020). Further issues might result from imbalance of data and over or under fitting of the data. Possible solutions to this problem are to either increase the size of the data during the collection process or to select a prediction model that is especially suited to small data sets (Xu et al., 2023). Testing the eventual model might also be a complex and time-consuming phase due to the infrequent occurrence. This could result in a more complex integration process.

2.4 Which prediction models are suited for demand forecasting for event catering?

2.4.1 Selected models

A variety of prediction models have been discussed so far, it is important to look at which models might be suitable for demand forecasting in event catering in order to compare the prediction models' performance, several different kinds of models should be developed as this will allow for the direct comparison of different models. Furthermore, the current method of predicting should be analysed to gain a clear understanding of the capabilities of AI in this field. Comparing results across different forecasting techniques also facilitates a more meaningful interpretation of a model's accuracy, as the performance score becomes more informative when compared against those of other models. Keeping this in mind, the following models will be developed:

2.4.1.1 Three game moving average

A three 'month' (in this case game) moving average prediction model is a very simplistic model that is a very commonly used model. Such a model does not take into account any additional input variables besides the historical sales data of the last three data points. This model tends to be easy to develop and easy to implement, explaining its popularity (Huriati et al., 2022). The simplicity and popularity of this prediction model is the main reason why this model was selected as it will serve as a suitable point for comparison.

2.4.1.2 Multiple linear regression

Another common data analysis model, which can also be used to generate predictions is a multiple linear regression model. This model does make use of input variables but does not include artificial intelligence. The regression model attempts to plot a linear line based on the dependent variables, based on the independent variables. For a new prediction the model finds the best fit on this line based on the independent variables of the new prediction (Uyanık & Güler, 2013). This model could be appropriate for the football club predictions as the dataset is relatively small and might not be sufficiently large for AI powered prediction models (Abeysinghe et al., 2003). Furthermore, given the commonality of regression models, it once again serves as a suitable comparison point.

2.4.1.3 Random forest regression

Looking at AI and machine learning models that can be used to generate sales forecasts, the random forest regression model can predict numerical outcomes and makes use of several combined decision trees. Many such trees are generated and trained on a historical dataset without human interaction beyond set up. For a new prediction, every such tree predicts the outcome. A final prediction is then constructed from the average of all the trees' predictions (Malhotra & Karanicolas, 2020). Given that the relatively small size of the dataset other typical AI models such as neural networks are often unreliable or not compatible. The random forest regression is typically less data hungry but still makes use of the ability to use AI and non-linear relationships (Roßbach, 2018).

2.4.1.4 Gradient boosting regression trees

Another potentially suitable AI model is Gradient boosting. Such a model works somewhat similar to the Random forest regression in terms that it also uses decision trees to come to a prediction. However instead of having many independent predicting trees, the gradient boosting models use sequential trees where every new tree learns from the previous one. Gradient boosting often results in more accurate predictions but is usually slower than Random forest regressions (Shi, 2023). Another disadvantage is that gradient boosting models tend to be more vulnerable to overfitting; the process of training over training on practice data which results in poor performance on new, unseen data. Given that the data set is relatively small and that a prediction model constantly uses new data for independent variables to generate predictions, this might be something to watch out for (Cui et al., 2023).

2.5 Which input variables are suited for AI sales forecasting in event catering?

When looking at the development of an AI powered prediction model it is important to consider which variables and influential factors are involved. One of the most important factors in the success of a machine learning model are the import factors used (Domingos 2012). Input variables can be divided in two categories, internal and external. Internal input variables are acquired within an organization and usually describe data sets produced by the organization. External input variables describe related data sets that do not find their origin within the organization (Kuncheva & Žliobaitė, 2009). The most common type of internal features are variables such as product sales figures for past time units. Such variables are the main mechanism by which the relationship between past and current values of a series can be captured by learning algorithms (Tsoumakas, 2018). Other variables that are proven to influence food sales are supply, demand, price, e-commerce, delivery services, and advertising (Liu et al., 2022). Furthermore, a study focussing on predicting attendance rates at football games included the following input variables; Performance of home team, Day of the game, Performance of the away team, Distance and uncertainty of outcome (Şahin & Erol, 2018). Not all of these will be applicable for (catering) sales forecasting as these variables are aimed at predicting attendance rates, however the day of the game is a notable exception as this is likely to influence purchasing behaviour. Finally, the weather has a large impact on consumer behaviour, specifically during football matches (Elliott, 2022).

Finally, the risk of overfitting, applying too many variables on a data set, should be attributed to. As datasets for event catering tend to be relatively small, the risk of overfitting the data is relatively more likely Ying (2019). To prevent this from occurring the amount and types of variables should be taken into account, as these should remain respectable in regard to the size of the dataset Babyak (2004).

3. Methodology

3.1 Research design

This study will conduct quantitative, exploratory research; aimed at the development of a data driven AI sales forecasting model for catering sales of a regional football club during match days. A combination of historical internal and external data will be used in the development of this model. The goal of the development of this model is more accurately and more easily forecast sales of future matches. This process entails data collection, data normalization, developing an unsupervised machine learning model to process the data, and will involve performance evaluation. The development of this model will enable answering the research question.

3.2 Data Collection

As described previously this study will combine internal data and external data sources. Both data types are required to design the model which is necessary to answer the research question. Internal data sources will be focussed on club data whilst external data will be focussed on external factors such as the weather or calendar dates.

3.2.1 Internal data

The primary dataset for the research has been obtained from the catering department from a regional football club. This dataset includes sales data from 19 matches for 20 selling points. A personal agreement with the club was reached, to not break confidentiality the football club will remain anonymous and their raw data will preferably not be published and will otherwise be anonymized. The internal data set includes Sales numbers per match and per selling point as well as totals. These include historic sales data for 5 different food products that are sold across the stadium. Furthermore, the attendance numbers have been acquired as well as match dates and times. The acquired dataset should be very representative for the research as the data was directly acquired from the club at play and the model will be shaped around the acquired data.

3.2.1 External data

Secondary data will be acquired from external sources. In order to achieve data to match all identified input variables external data will have to be collected for the team's performance, the calendar dates and crucially the weather. The representativeness of the external data will have to be considered. Since the collection process will be independent from the club, in order to make sure the external data is representative the following will have to be taken into consideration; the data has to match the exact date and time periods, local data sources should be used as much as possible, during data normalization the structure of internal and external data should match, all data points that are taken into consideration should also have matching external data and finally the validity and trustworthy of data sources should be taken into consideration.

3.3 Data analysis

3.3.1 How will data analysis be conducted and how can AI sales prediction models be evaluated?

In order to answer the research question, a case study will take place using the dataset provided by the local football club. The previously discussed four prediction models will be attempted to develop. All four will then be tested on a subset of the data. The subset will not be included in the development of the models. Each model is later tasked with generating a prediction for the specific match in the test subset. The models will then be judged by calculating their average accuracy over all test cases, Accuracy will be determined by using the Mean Absolute Percentage Error (MAPE) (Lewis, 1982). This expresses the average absolute deviation as a percentage of total sales and is one of the most commonly used measures of accuracy (Chen et al., 2003). As accuracy is the selected measurement, the inverse of the MAPE will be used to express the model's performance.

3.4 Bias

During the data collection and normalization processes bias should be limited as much as possible. In comparison to qualitative research, quantitative methods are usually less susceptible to bias as the validity and reliability of studies are assessed using statistical tests that estimate the size of error in samples and calculate the significance of findings (Smith & Noble, 2025). In order to limit the biases In a well-designed research protocol explicitly outlining data collection and analysis can assist in reducing bias (Smith & Noble, 2025).

4. Results

4.1 Selected Variables

In chapter 2.5 potentially suited input (independent) variables were briefly discussed. During the input variable selection process, there are a number of factors that should be considered. First and foremost, the availability of data for all data points is necessary for any variable to be included, as the model should have consistent data for each data point and each prediction. Furthermore, the EPV (Events Per Variable) should have a ratio of at least 10. For this study in specific, that means a total degrees of freedom of 42, across all included input variables. Finally, the input data should be easily accessible to a user, as the model should remain user friendly. An example for this is to not include total millimetres of rain or cloud percentages but binary selectors for rainy and cloudy weather (either rainy or not), as needing such specific weather forecasts to make a new prediction would be very user unfriendly.

Considering the considerations outlined in the theoretical framework in chapter 2.5, the availability of data for all data points and user friendliness the following independent variables were selected:

Name	Type	Degrees of Freedom
Attendance	Numerical	1
Temperature	Numerical	1
Date	Date/Time, DD-MM-YYYY	1
Time	Categorical, HH:MM, 8 set timeslots	7
Rainy?	Binary, T/F	1
Cloudy?	Binary, T/F	1
Day	Categorical, Tue,Wed,Thu, Sat,Sun	4
Opponent Category	(Categorical, A-C	2
Days until the next home match	Numerical	1
Selling Point	Categorical, 1-20	19

4.2 Model Evaluation

After the previously outlined models were developed each model now has to be evaluated. The prediction models will be judged by a cross-validation, the process of testing a prediction model on unseen data after which the model is evaluated (Berrar, 2018) method. The cross-validation model was selected as it is the one of the most widely used data methods used to estimate true prediction errors (Berrar, 2018). Every model was used to generate a prediction for three test cases that were not included in the datasets that were used to develop the models. A comparison can now be made between the model's sales forecast and the actual sales. (Lewis, 1982) serves as a means to interpret the MAPE and therefore gives meaning to the accuracy scores.This means the MAPE scores can be interpreted as the following, under 10% is highly accurate, between 11% and 20% can be considered as a good forecast, between 21% and 50% is reasonable and a higher MAPE score can be considered inaccurate. This specific method was chosen because it is easy to calculate for all models, is intuitive, and is expressible both in error margin (MAPE) and as an accuracy percentage and for its scalability.

4.2.1 Three game moving average

After generating the predictions for the test cases, the 3 Game Moving Average (3GMA) model's predictions result in an overall average accuracy of 59,7% across all selling points. As previously mentioned, this number can be interpreted as an average absolute deviation (MAPE) from the actual sale number of roughly 40%. Using

Lewis' (1982) interpretation that leads to the conclusion that the model scores on the low end of the reasonable range. A notable feature is that one of the categories' predictions is far worse than the others. If this product were removed from the prediction the average accuracy rises to 68,9%, resulting in a MAPE of roughly 30%. This specific product has an overall average accuracy of 49%, compared to 56%,55%,77% and 66% respectively.

4.2.2 Multiple Linear Regression

When evaluating the MLR (multiple linear regression) model's performance the average accuracy comes down to 52,7%, resulting in a MAPE of 47,3%. This results once again in a reasonable prediction, albeit on the very low end. Once again however the same specific product performs much worse. If excluded this results in 77,2% accuracy, or a MAPE of 22,8%, which would place it within the same category but on top end rather than the bottom. This problem, however, occurs consistently across all selling points and all test cases, resulting in an average product accuracy of 29%, compared to 79%, 81%, 75% and 69% respectively. Overall, the MLR model performs reasonably, but does significantly better when one specific product is excluded from the predictions. It is also notable that the MLR model performs poorly on the last test case with an overall accuracy score (excluding the poorly performing product) of 64,4%, compared to 81,5% and 85,8%.

4.2.3 Random Forest Regression

The Random Forest Regression (RFR) model achieves an average accuracy of 64,3%, which translates to a MAPE of 35,7%, placing it in the reasonable category. If the same specific product is excluded once again, an average accuracy of 74,9% is achieved. Notably the average accuracy of the specific product is much higher than in the other methods (58%). The model also performs consistently across all test cases. The RFR model therefore scores as reasonable on Lewis' (1982) scale but performs consistently across all test cases.

4.2.4 Gradient Boosting Regression Trees

The Gradient Boosting Regression Trees (GBRT) model's predictions result in an average accuracy of 61,9%, resulting in a MAPE of 38,1%. However, the model's accuracy improves drastically when the same product is excluded. Such a change results in 77,1% accuracy and a MAPE of 22,9%. Interestingly the bad prediction for this product only occurs for one test case (-9% accuracy) and is quite good for the remaining test cases (83% and 82%). The model also performs consistently across the different test cases, with an exception for the previously mentioned occurrence. The model also scores as reasonable on Lewis' (1982) scale.

4.3 Model Comparison

After evaluating each model individually, the models will be compared to one another. Especially the performance of the AI models (RFR and GBRT) compared to the more traditional models (3GMA and MLR). What becomes clear is that all models struggle with accurately predicting one of the products. However, the AI models do achieve more accurate and more consistent results for that specific product. Another occurrence is that the AI models perform more consistently across the test cases than the traditional models (when the unreliable product is excluded from predictions). When the models are compared by average accuracy all models outscore the 3GMA prediction, which does not integrate the independent variables). The MLR and GBRT models score the best, with MLR achieving the highest accuracy in a single match and with GBRT scoring the most consistently. GBRT also generates the best prediction for the product that scored considerably worse than the others. It is also notable that the GBRT model did in fact consistently outperform the RFR model, which indicates that the added performance did outweigh the risk of overfitting the data.

4.3.1 Which prediction model is most accurate for event catering sales forecasting?

In general, the best two prediction models were the Multiple Linear Regression (MLR) model and the Gradient Boosting Regression Trees (GBRT) model, with the MLR model achieving the highest accuracy in single matches and the GBRT model performing more consistently. The average accuracy scores lie very close to one another. The Gradient Boosting Regression model, however, does achieve a much higher accuracy than the Multiple Linear Regression model when all products are included. In order to definitively prove which forecasting model is most suited for this case more test cases are needed as the accuracy of the MLR model is negatively affected by one of the test cases where it scores considerably worse than the other test cases. However, due to the limited availability of data there are not sufficient test cases to determine whether this specific match can be considered to be an outlier. Therefore, the Gradient Boosting Regression Trees model is selected as the most accurate model as it achieves equal accuracy whilst maintaining better consistency.

4.4 Current methods

In order to gain further insight into the performance of the AI model's performance the current prediction methods of the football club should be analysed. Currently the club uses two key variables to predict the sales of future matches, namely the day and the time. The club records matchday sales up to a year ago. For a new prediction, the match with the most total sales and the same day and time as the predicted match is selected. Sales or that match are then taken for the new prediction.

This method adds to simplicity but does not take other influential factors into account such as the attendance or the weather. When applied on the same test sets as the other models, this prediction method achieves an average accuracy of 75%, resulting in a MAPE of 25%. Compared on Lewis' (1982) scale, this results in the reasonable category.

4.5 How do AI forecasting models perform compared to current methods?

The two most accurate models, the Multiple Linear Regression model and the Gradient Boosted Regression Trees can now be directly compared with current methods, to evaluate the performance of the machine learning and AI powered models. What becomes clear is that even though the current method is relatively simplistic and only takes two variables into account, it does relatively well. The accuracy and MAPE scores of the current method are comparable to the MLR and GBRT models' scores (75% compared to 77% and 79%). The current method also has great difficulty predicting the product for which the other models were experiencing difficulties. What becomes clear is that the GBRT model does the best job of predicting this product, as it achieves 52% accuracy, compared to 29% (MLR) and 41% (current method). The consistency of the current method is comparable to that of the GBRT model but performs slightly less accurately across the predictions. Regarding the away section, the current method performs very poorly, as it only manages to achieve an average 38% accuracy, compared to 54% (MLR) and 62% (GBRT), however when the away section is excluded from the prediction score, Both the MLR and GBRT model outperform it, albeit marginally. When comparing the three models the following may be concluded, the Multiple Linear Regression achieves the highest accuracy for single predictions, but its inconsistency affects its overall accuracy, the Gradient Boosted Regression Trees achieve the best overall accuracy but are the most complex and achieve their accuracy through consistency. The current method is the most simplistic and easiest to implement, but although it performs consistently it achieves the least accuracy. Both the MLR and GBRT models also outperform the current method for the poorly predictable product and the away section.

4.6 Model Overview

	3GMA	MLR	RFR	GBRT	Current method
MAPE	40,3%	47,3%	35,7%	38,1%	40,2%
MAPE ex prod	31,1%	22,8%	25,1%	22,9%	39,1%
MAPE ex away	38,9%	46,8%	34,7%	38,1%	25,6%
MAPE ex prod, ex away	28,2%	21,3%	23,1%	21,9%	22,8%
Mape away section	67,6%	46,2%	53,6%	38,3%	62,0%

4.7 Notable Findings

During the model evaluation a number of notable occurrences could be observed. These observations include the poorly performing product, consistency of model performance across the test cases and the role of the away section in the stadium.

4.7.1 Product performance

As mentioned previously, predictions for one product in specific proved to be far less accurate than predictions for other models. What makes this especially interesting is that this occurred through most of the predictions, even though the individual models all use different predicting techniques. Surprisingly, even the 3 Game Moving Average model and the current method, which do not account for independent variables, perform significantly better on the other products. The reason for this is thought to partially lie in the way the MAPE is calculated. As mentioned previously there are multiple selling points. However, not all selling points have the same products, except for this specific product. Since the overall accuracy is calculated by taking the average of the different selling points' accuracy, this specific product has a higher influence on the final accuracy score. Even though this does explain why including or excluding this product has a relatively large influence on the final outcome, it does not explain why predictions for this product itself are worse across all test cases and models. It is even counterintuitive since for this specific product the most data is available. When the variance of this product is examined in the test data it turns out to be significantly less than for the other products (27% compared to 327%, 112%, 123% and 160% respectively). This could be a reason why the prediction models are struggling to accurately forecast sales for this product as the low variance results in an increased difficulty to recognise patterns. This however does not explain the poor performance of the Three Game Moving Average, as this model does not include pattern recognition. However, due to the low variance, a single case of a large deviation does have a higher influence. This does occur in the first test case, where an average accuracy score of -20% is achieved for that product.

4.7.2 Match Consistency

Another interesting attribute of the predictions is the consistency of the models across the different test cases. The Multiple Linear Regression model and the Three Game Moving Average model both encountered inconsistencies between matches with an accuracy range of 29% for the 3GMA and 21% for the MLR model, compared to 6% for the RFR and GBRT models. This shows that the AI model appears to be more reliable in terms of performance. The specific match where both the 3GMA and MLR model had inaccurate predictions was a play-off match and experienced an unusually low attendance, which might explain the poor performance of the non-AI models. The fact that the RFR and GBRT models perform significantly better than the other two models here might be explained by the fact that matches such as that one are typically rare. This might hint at an increased ability of the AI models to predict for uncommon matches compared to the more traditional models.

4.7.3 Away Section

Finally, an expected but notable occurrence is the relatively poor performance of the prediction models for the away section. One of the 20 selling points is located in the away section, which has different demographics every match. For example, some clubs fill the away section in its completion, other clubs bring just a few dozen away fans. For some rare matches the away section is even available to home supporters because the opponent has been disallowed from bringing away fans. Therefore, the attendance numbers for the away section, which were not available, are not included in the historical data but do in fact have a strong influence on sale numbers. Furthermore, during international matches against opponents from Muslim majority countries, sales for pig-based products are far less than against normal opponents. Given that such data is not included in the historical data and is not included in the dependent (input) variables the models are unable to analyse or predict this trend. Finally, the background of the away supporters is different for each match, which makes pattern recognition especially difficult. It is therefore not surprising that the away section consistently has the worst performing predictions, except for a few rare occurrences. This trend can be seen across all models and all test cases.

5. Discussion

5.1 Theoretical Implications

As outlined in the theoretical framework section, a research gap exists considering AI sales forecasting for event catering. Prior research such as (Turkmen et al., 2020) and (Pang & Wang, 2024) indicated a number of challenges and complications for demand forecasting in event catering, such as the limited availability of historical data points due to infrequent occurrences, the impracticality of most standard forecasting techniques and challenges on measuring forecast accuracy, model selection, and forecast aggregation. Very few studies have however, empirically tested the performance of AI prediction model for event catering under such limitations. The findings of this research align with the theorized limitations, as such challenges proved to be a challenging component in the development of the predicting models. The AI models that were chosen were highly influenced by the lack of sufficient historical data; the selected AI model (Random Forest Regression and Gradient Boosted Regression Trees) are among the least data hungry models available. This aligned with the theorized shortcomings of the acquired dataset. However, both AI prediction models managed to generate predictions with acceptable error margins (MAPE). This supports the theorized concept that less data intensive models such as Gradient Boosted Regression Trees and Random Forest Regression can still be implemented, despite the relatively small dataset. This partially confirms the existing theory of heavy data dependency but also challenges the assumptions commonly made about the unsuitability of AI prediction models in data sparse environments. As described in chapter 2.5, InTo prevent major issues with the prediction models, the Events Per Variable should be sufficiently large. To stay clear of problems an EPV of at least 10 should be implemented (Peduzzi et al., 1996). Given the nature of events, with limited access to historical data and many potential predictors, it was uncertain whether enough observations could be acquired to maintain an EPV of at least 10. The EPV can be determined by dividing the total number of observations by the total degrees of freedom for the predicting (independent) variables (Austin & Steyerberg, 2014). Mainly due to the relatively limited degrees of freedom for the predictors, an EPV of 11.1 was maintained, indicating that even though there are typically many predictors and limited observations associated with generating predictions for the event sector, a sufficiently large EPV can be maintained, resulting in appropriately performing AI prediction models.

Therefore, it may be concluded that although more complex and limiting, the event sector can make use of the new technology to use Artificial Intelligence for demand and sales forecasting. The theoretical implications for this research support concepts in existing literature for the challenges in AI sales forecasting in event catering but also shows that the concept of AI sales forecasting in event catering is possible and achieves satisfactory results. Therefore, this research not only confirms the existing theories but also challenges common assumptions and proves that AI sales forecasting is possible in data sparse fields.

5.2 Practical Implications

Given that AI demand forecasting was successfully experimented, other experiments can take place to attempt to duplicate the results. If proven to be successful to be developed and implemented, AI demand forecasting could strongly influence the event catering sector. In practice this could potentially lead to waste reduction, labour reduction, optimized inventory management and better financial performance (Rodrigues et al., 2023). Integrating automated and accurate demand forecasting can help businesses to shift from a reactive to a predictive operation, resulting in the business planning for demand instead of reacting to demand. The implication of accurate AI demand forecasting models has the potential to overcome shortcomings of traditional methods and uncover complex patterns. This has the potential to improve inventory management, recourse allocation, pricing strategies and supply chain management efficiency (Kumar & Nayak, 2024). Practical implications could however be complex as they are often complex and difficult to develop from prototype to product (Bosch et al., 2020). The practical implications for AI powered sales forecasting in the event catering industry are therefore very promising with a high potential for optimization but could be complex and difficult to develop and integrate. Currently the football club will further evaluate the models' performance by comparing predictions with actual sales for future matches. Eventually the models can be used to efficiently supply the different selling points, which will offer benefits in terms of supply management as well as create a calmer working environment as there will be less resupply movements.

5.3 Limitations

Even though this research demonstrates the possibility of the development for an AI powered sales forecasting model for event catering industries, several limitations were encountered. Due to the lack of historical data, other commonly used AI prediction models were unable to be tested or integrated. Artificial Neural Networks (ANN) are one such example. This prediction model is widely used for similar prediction models in other industries and is usually performing quite well. In this case however, this technique proved highly unsuited. A comparison between the performance of Artificial Neural Networks and Gradient Boosting Regression Trees would have been interesting and could have been used to evaluate the performance of different AI prediction models. This option, however, remains unexplored as a functional ANN prediction model could not be developed. Therefore, the best suited AI prediction model can only be determined from options available, rather than from all potentially suited models.

Another complicating factor resulting from the lack of data was the amount of predicting (independent) variables that could be included in the model. As described previously, an Events Per Variable ratio of at least 10 had to be maintained. Since the number of observations available was fixed, the total degree of freedom was reduced to at most 42. This directly influenced how many and which predictors could be included in the model, as too many variables or variables having a high degree of freedom could strongly influence the EPV ratio. Furthermore, given the predictions are generated from historical sales data, these numbers could be misleading. The sales data does not include supply; therefore, the recorded sales do not reflect the demand when a product is sold out. The sales numbers do not include uncaptured demand, which may lead to a biased demand forecast. A potential solution for this would be to include supply as a predictor, this would however strain the EPV ratio. Furthermore, this would require the business to have such numbers available for every data point included in the historical data set. As discussed previously the model is also unable to reflect on the changing demographics of the away section of the stadium. This represents another limitation in the model. Knowledge outside the historical data and provided secondary data is unavailable to the prediction, highlighting the need for human interaction with the prediction to correct for such instances.

A final limitation of the model is the insufficiency of test cases, once again resulting from the sparsely available datasets. As described in the previous section, predictions for matches that could potentially be considered outliers could not be defined as such due to insufficient other test cases for which they would be compared against. A commonly used ratio to determine the amount of test cases relative to the historical data is 20% test cases and 80% historical data (Joseph, 2022). This however would result in a conflict with the EPV ratio of at least 10, therefore 12,5% were test cases. This did allow for a functional model but did hurt the ability to determine outliers and gain a more meaningful insight into the model's performance.

5.4 Future Research

Future research into this specific topic has many applications. First and foremost, the duplication of results could give insight into sector wide possibilities, in comparison to business specific.

With the availability of more data different AI prediction models could be tested and evaluated. This would allow for optimization of the model as the best suited predicting technique could be found. Furthermore, the models could be tested more extensively, which would allow for better and more accurate model evaluation. If available the difference between demand and sales could be explored, allowing for more precise predictions and a reduction of uncaptured demand.

Another interesting topic for future research is to explore the integration process within the event catering industry, focussing on challenges related to the shift from prototype to product as well as employee adaptation to emerging technologies.

Finally, the potential effects of AI sales forecasting on different aspects of the event catering industry, such as supply chain management, waste reduction management, inventory management and purchasing and supply management would be areas for future research. Evaluating whether the theoretical benefits of accurate AI sales forecasting actually materialize would extend the implications of the new technology.

5.5 Conclusion

Even though the developed model has considerable limitations, mainly due to the lack of sufficient historical data, AI powered sales forecasting for the event catering sector is possible. Further research is however required to determine the most optimal predicting technique. The main limitation for data driven AI sales forecasting for event catering is the limited availability of historical data, stemming from the infrequent occurrence of events. That having said, accurate sales forecasting using AI has very high potential, especially for the event catering sector and is therefore a promising topic for future research and experimentation.

How accurately can AI sales forecasting models predict the catering sales for sport events, based on historical data and event specific variables?

To formulate a concluding answer to the research question: AI sales forecasting of matchday catering sales of a football club is possible and, albeit with notable limitations, can achieve between 60 and 80% accuracy, depending on configuration and included products.

References

[1] Abeysinghe, T., Balasooriya, U., & Tsui, A. (2003). Small-Sample Forecasting Regression or Arima Models? *Journal Of Quantitative Economics*, *1*(1), 103–113. https://doi.org/10.1007/bf03404652

[2] Ahaggach, H., Abrouk, L., & Lebon, E. (2024). Systematic Mapping Study of Sales Forecasting: Methods, Trends, and Future Directions. *Forecasting*, *6*(3), 502–532. https://doi.org/10.3390/forecast6030028

[3] Alwosheel, A., Van Cranenburgh, S., & Chorus, C. G. (2018). Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal Of Choice Modelling*, *28*, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

[4] Austin, P. C., & Steyerberg, E. W. (2014). Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Statistical Methods in Medical Research*, *26*(2), 796–808. https://doi.org/10.1177/0962280214558972

[5] Babyak, M. A. (2004). What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine*, *66*(3), 411–421. https://doi.org/10.1097/01.psy.0000127692.23278.a9

[6] Berrar, D. (2018). Cross-Validation. In *Elsevier eBooks* (pp. 542–545). https://doi.org/10.1016/b978-0-12-809633-8.20349-x

[7] Biswas, B., Sanyal, M. K., & Mukherjee, T. (2023). AI-Based Sales Forecasting Model for Digital Marketing. *International Journal Of E-Business Research*, *19*(1), 1–14. https://doi.org/10.4018/ijebr.317888

[8] Bosch, J., Crnkovic, I., & Olsson, H. H. (2020). Engineering AI Systems: a research agenda. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2001.07522

[9] Brigato, L., & Iocchi, L. (2020). A Close Look at Deep Learning with Small Data. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2003.12843

[10] Cadavid, J. P. U., Lamouri, S., & Grabot, B. (2018, 8 juli). *Trends in Machine Learning Applied to Demand & Sales Forecasting: A Review*. https://hal.science/hal-01881362v1

[11] Chen, R. J. C., Bloomfield, P., & Fu, J. S. (2003). An Evaluation of Alternative Forecasting Methods to Recreation Visitation. Journal Of Leisure Research, 35(4), 441–454. https://doi.org/10.1080/00222216.2003.11950005

[12] Cui, S., Sudjianto, A., Zhang, A., & Li, R. (2023). Enhancing Robustness of Gradient-Boosted Decision Trees through One-Hot Encoding and Regularization. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.13761

[13] Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. *Journal Of Food Engineering*, *75*(2), 196–204. https://doi.org/10.1016/j.jfoodeng.2005.03.056

[14] Dolgui, A., Bernard, A., Lemoine, D., Von Cieminski, G., & Romero, D. (2021). Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems. In IFIP advances in information and communication technology. https://doi.org/10.1007/978-3-030-85906-0

[15] Domingos, P. (2012). A few useful things to know about machine learning. *Communications Of The ACM*, *55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

[16] Douaioui, K., Oucheikh, R., Benmoussa, O., & Mabrouki, C. (2024). Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review. *Applied System Innovation*, *7*(5), 93. https://doi.org/10.3390/asi7050093

[17] Elliott, H. (2022, 12 augustus). *Football and fine weather result in winning drinks sales - CGA*. CGA. https://cgastrategy.com/football-and-fine-weather-result-in-winning-drinks-sales

[18] Fridley, L. (2018). *Improving Online Demand Forecast using Novel Features in Website Data: A Case Study at Zara* (Massachusetts Institute of Technology, MIT Sloan School of Management, Department of Civil Engineering, Leaders for Global Operations Program at the Massachusetts Institute of Technology, G. Perakis, P. Jaillet, G. Perakis, & P. Jaillet). https://dspace.mit.edu/bitstream/handle/1721.1/117976/1051238258-MIT.pdf

[19] Ganesan, P. (2024). AI-Powered Sales Forecasting: Transforming Accuracy and Efficiency in Predictive Analytics. *Journal Of Artificial Intelligence Machine Learning And Data Science*, *2*(1), 1213–1216. https://doi.org/10.51219/jaimld/premkumar-ganesan/280

[20] Goldblatt, 2005, Getz, 2008, Bowdin, Jago, 1998, & Shaw, 1998. (2005). Event Management HM-402. In *Uttarakhand Open University*. https://www.uou.ac.in/sites/default/files/slm/HM-402.pdf

[21] Groene, N., & Zakharov, S. (2024). Introduction of AI-based sales forecasting: How to drive digital transformation in food and beverage outlets. *Discover Artificial Intelligence*, *4*(1). https://doi.org/10.1007/s44163-023-00097-x

[22] Grolinger, K., L'Heureux, A., Capretz, M. A., & Seewald, L. (2015). Energy Forecasting for Event Venues: Big Data and Prediction Accuracy. *Energy And Buildings*, *112*, 222–233. https://doi.org/10.1016/j.enbuild.2015.12.010

[23] Gustriansyah, R., Ermatita, E., & Rini, D. P. (2022). An approach for sales forecasting. *Expert Systems With Applications*, *207*, 118043. https://doi.org/10.1016/j.eswa.2022.118043

[24] Gustriansyah, R., Suhandi, N., Antony, F., & Sanmorino, A. (2019). Single exponential smoothing method to predict sales multiple products. *Journal Of Physics Conference Series*, *1175*, 012036. https://doi.org/10.1088/1742-6596/1175/1/012036

[25] Guo, Z., Wong, W., & Li, M. (2013). A multivariate intelligent decision-making model for retail sales forecasting. *Decision Support Systems*, *55*(1), 247–255. https://doi.org/10.1016/j.dss.2013.01.026

[26] Harrison, P. J. (1965). Short-Term sales forecasting. *Journal Of The Royal Statistical Society Series C (Applied Statistics)*, *14*(2/3), 102. https://doi.org/10.2307/2985338

[27] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts. https://otexts.com/fpp2/

[28] Huriati, P., Erianda, A., Alanda, A., Meidelfi, D., Rasyidah, Defni, & Suryani, A. I. (2022). Implementation of The Moving Average Method for Forecasting Inventory in CV. Tre Jaya Perkasa. *International Journal Of Advanced Science Computing And Engineering*, *4*(2), 67–75. https://doi.org/10.62527/ijasce.4.2.77

[29] Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis And Data Mining The ASA Data Science Journal*, *15*(4), 531–538. https://doi.org/10.1002/sam.11583

[30] Kumar, S., & Nayak, A. (2024). Predictive Analytics for Demand Forecasting: A deep Learning-based Decision Support System. *International Journal Of Engineering And Computer Science*, *13*(07), 26291–26299. https://doi.org/10.18535/ijecs/v13i07.4853

[31] Kuncheva, L. I., & Žliobaitė, I. (2009). On the window size for classification in changing environments. *Intelligent Data Analysis*, *13*(6), 861–872. https://doi.org/10.3233/ida-2009-0397

[32] Lewis, C. D. (1982). *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*. Butterworth-Heinemann. https://archive.org/details/industrialbusine0000lewi/page/n3/mode/2up

[33] Malhotra, S., & Karanicolas, J. (2020). A Numerical Transform of Random Forest Regressors corrects Systematically-Biased Predictions. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2003.07445

[34] Pang, Y., & Wang, F. (2024). Forecasting Stadium Attendance Using Machine Learning Models: A Case of the National Football League. *Studia Sportiva*, *18*(2). https://doi.org/10.5817/sts2024-2-12

[35] Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal Of Clinical Epidemiology*, *49*(12), 1373–1379. https://doi.org/10.1016/s0895-4356(96)00236-3

[36] Rodrigues, M., Miguéis, V., Freitas, S., & Machado, T. (2023). Machine learning models for short-term demand forecasting in food catering services: A solution to reduce food waste. *Journal Of Cleaner Production*, *435*, 140265. https://doi.org/10.1016/j.jclepro.2023.140265

[37] Roßbach, P. (2018). *Neural Networks vs. Random Forests – Does it always have to be Deep Learning?* https://blog.frankfurt-school.de/wp-content/uploads/2018/10/Neural-Networks-vs-Random-Forests.pdf

[38] Schreyer, D., & Ansari, P. (2021). Stadium Attendance Demand Research: A Scoping review. *Journal Of Sports Economics*, *23*(6), 749–788. https://doi.org/10.1177/15270025211000404

[39] Shi, Y. (2023). Sales Prediction Optimization via Gradient Boosting Decision Tree. *Highlights in Science Engineering And Technology*, *44*, 121–128.

https://doi.org/10.54097/hset.v44i.7274

[40] Tulli, S. K. C. (2020). Comparative Analysis of Traditional and AI-based Demand Forecasting Models. *INTERNATIONAL JOURNAL OF EMERGING TRENDS IN SCIENCE AND TECHNOLOGY*, 6933–6956. https://doi.org/10.18535/ijetst/v7i6.02

[41] Turkmen, A. C., Januschowski, T., Wang, Y., & Cemgil, A. T. (2020). Intermittent Demand Forecasting with Renewal Processes. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2010.01550

[42] Uyanık, G. K., & Güler, N. (2013). A Study on Multiple Linear Regression Analysis. *Procedia - Social And Behavioral Sciences*, *106*, 234–240. https://doi.org/10.1016/j.sbspro.2013.12.027

[43] Xu, P., Ji, X., Li, M., & Lu, W. (2023). Small data machine learning in materials science. *Npj Computational Materials*, *9*(1). https://doi.org/10.1038/s41524-023-01000-z

[44] Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal Of Physics Conference Series*, *1168*, 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

[45] Yule, G. U. (1909). Method of Correlation to Social and Economic Statistics. In *Journal Of The Royal Statistical Society* (Vol. 72, Nummer 4, pp. 721–723). https://academic.oup.com/jrsssa/article/72/4/721/7091264