UTILIZING MACHINE LEARNING FOR PREDICTIVE ANALYTICS AND OPTIMIZATION IN RESTAURANT OPERATIONS

MSc Research Project
MSc in Data Analysis

Nishika Gala
Student ID: x22214062

School of Computing
National College of Ireland

Supervisor: Dr. Muslim Jameel Syed

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Nishika Gala

Student ID: X22214062

Programme: MSc Data Analysis **Year:** 2023-24.

Module: ……………………………………………………………………………………………………….………

Supervisor: 12th August 2024
Submission Due Date:

Project Title: Utilizing Machine learning for predictive analysis and optimization in restaurant operations.

Word Count: 12250…… **Page Count**………54…………………………………….……..

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ………Nishika Gala…………………………………………………………

Date: ………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Abstract

This work aims at reviewing and analysing how machine learning models can be used to improve various aspects of operation in business. This involves scheduling of personnel, demand forecasting, menu mix, and supply chain management. The Logistic Regression technique is used for forecasting the demand for products with greater accuracy, better inventory controls and resource allocation through effective demand forecasting. Gradient Boosting Regression is used in menu optimizations, designed to achieve maximum profit by responding to operations and consumers' preferences. Random Forest Regressor helps in case of staff scheduling issues, making workforce planning and utilization more effective by using algorithms. This project also taps on the use of new technologies for supply chain management with aims at enhancing purchasing strategies for supply and inventory management for increased efficiency. By means of these applications, it shows how state-of-the-art machine learning models can make difference in the efficiency of operations in a business, as well as in the strategic support of a variety of business activities.

Acknowledgement

I would like to express my heartfelt gratitude to everyone who supported me throughout this project. My sincere thanks to my supervisor, whose guidance and expertise were invaluable in shaping this research. I also appreciate the input and encouragement from my colleagues and peers, who provided insightful feedback and suggestions. Special thanks to my family and friends for their unwavering support and patience. Their belief in me has been a constant source of motivation. I am grateful for the resources and tools provided, which were crucial in conducting this analysis. Finally, I acknowledge the contributions of the academic community and researchers whose work laid the foundation for this project.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Introduction

Economies of operation are important for the restaurant business for it to be competitive and profitable. It becomes very challenging to apply classical approaches toward various complex tasks associated with the restaurants' management as the number of various aspects to be managed rises. Machine learning is seen as a possible solution that holds even better approaches and mechanisms to improve the decision-making process. This study aims to discover the field of using machine learning for restaurants to improve the restaurant's operational performance including demand forecasting, menu optimization, scheduling of staff, supply chain, and optimizing customer satisfaction. To develop different solutions and recommendations, the study seeks to embed machine learning algorithms, making it possible to offer tangible findings and recommendations that will enhance business operations and customers' perceived value in the restaurant market space.

1.2 Background

Restaurant trading is very sensitive to the global economy because it contributes to creating employment opportunities and food diversities. Nevertheless, the industry has been subjected to some challenges such as customers' varying demand, inventory, and staff scheduling and ensuring service quality. The management of these operations in the past has always involved the use of heuristics and manual means which are most of the times very inefficient. Technology most especially machine learning brings solutions to these challenges as follows. Superiority in handling large datasets to recognize trends and patterns to make proper forecasts and improve several areas of the restaurant business. For instance, demand forecasting using machine learning helps in predicting the stream of customers to improve inventory and staffing arrangements. In the same way, machine learning improves menu optimization by using information from customers and the sale of their products to recommend the most profitable and demanded foods. Total working force management can also be enhanced by predicting one peak time and matching it with the available employees.

There is the enhancement of the procurement process and the inventory management process as some of the benefits of machine learning in supply chain management. The improvement of the customer experience through analysis of opinionated views and recommendations of what may meet their preferential tastes is another critical use. While bringing these innovative technologies into practice, the restaurant industry has the potential to enhance its productivity and effectiveness, improve the customers' satisfaction rates, and escalate its revenue levels. This research focuses on these applications as it seeks to give a detailed description of how machine learning can transform restaurants.

1.4 Aim and objectives

1.4.1 Aim

This research aims to investigate the application of machine learning techniques to optimize restaurant operations and enhance customer satisfaction.

1.4.2 Objectives

- To develop machine learning models for accurate demand forecasting, enabling better inventory management and resource allocation
- To implement menu optimization strategies using machine learning algorithms, focusing on maximizing profitability and aligning offerings with customer preferences
- To design and evaluate staff scheduling solutions that utilize predictive analytics to improve workforce management and productivity
- To enhance supply chain management through machine learning applications, improving procurement processes, inventory control, and overall efficiency

1.5 Research question

The research questions of this project are as follows:

1. How can machine learning models be developed and applied for accurate demand forecasting to improve inventory management and resource allocation in restaurants?
2. What machine learning algorithms and strategies can be implemented for menu optimization to maximize profitability and align offerings with customer preferences?

3. How can predictive analytics be utilized to design and evaluate staff scheduling solutions that enhance workforce management and productivity in the restaurant industry?
4. In what ways can machine learning applications improve supply chain management, including procurement processes, inventory control, and overall operational efficiency in restaurants?

1.6 Research hypothesis

H0 (Null Hypothesis): There is no significant improvement in operational efficiency and customer satisfaction in the restaurant industry through the application of machine learning techniques.

H1 (Alternative Hypothesis): The application of machine learning techniques significantly enhances operational efficiency and customer satisfaction in the restaurant industry.

1.7 Rationale

The justification for this research stems from the fact that the restaurant business space particularly suffers from market competition and increasing customer demands which call for organizational developmental change and enhancement in operational procedures. Inefficiencies in acquiring and allocating resources continue to remain a thorn in traditional techniques due to rigidity, especially in ever-changing markets. This brings the applicability of machine learning as a potential solution to these challenges given it implies the application of big data in some of the core competencies, for instance, demand forecasting, menu optimization, staff scheduling, and supply chain. Theoretically, this research seeks to establish how the implementation of machine learning innovations in these fields can result in increased efficiency, decreased costs, and therefore improved customer satisfaction. The research aims to enrich the current knowledge concerning the effectiveness of applying advanced technologies into the restaurant environment to achieve sustainable business development and optimal functionality.

Machine learning in supply chain management can create that positive change that mobilizes the chain resulting in reduced cost and better service. Improving customer satisfaction is another major reason, as this case would include recommendation and sentiment analysis to lure the customers into patronizing the business more often. This research will provide a systematic look into these applications, to show how the use of machine learning can yield real solutions to challenges encountered in the restaurant business. The aim of this work is to provide significant insights into the role of machine learning in enhancing the restaurant's performance and making a positive contribution to the existing knowledge base; while on the practical level, a set of recommendations for the effective management of the restaurant's operations and their improvement with the help of machine learning are provided.

1.8 Scope of research

This research's main interest is in establishing actual uses and potential repercussions of machine learning when applied to the restaurant business. The focus includes a general overview of the various machine learning approaches that can be used for solving the main important problems within the restaurant business, such as demand forecasting, menu optimization, staff scheduling, and supply chain management.

Literature review and empirical research concerning trends, threats, and opportunities of applying machine learning in restaurants will be analysed in the study. As part of it, it is necessary to design and assess the validity of selected machine learning models adapted to the peculiarities of the restaurant's operations.

Qualitative and quantitative research approaches are used to collect data from restaurant stakeholders, industry experts and data sources. In order to conduct the research and implement and test the best machine learning models, this research employs computational tools and programming language Python with the support of machine learning libraries such as TensorFlow and sci-kit-learn.

Concerning the study area, the research will cover a wide range of restaurants including independent restaurants as well as big chains to establish broader conclusions on the effects of overemphasis on operational efficiency. Privacy and disclosure of the participant's identity will be maintained to avoid violating their ethical rights in the collection and analysis of the data.

1.9 Structure of the report

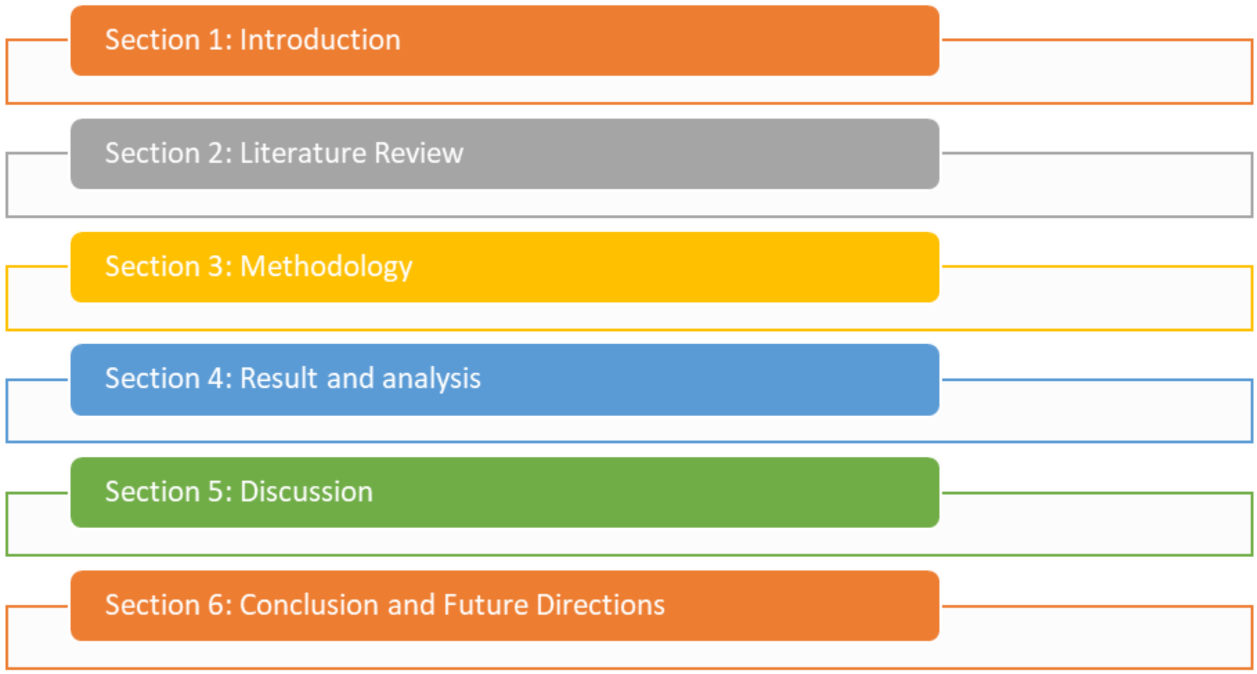

Figure 1.9: Structure of report

1.10 Summary

This research intends to investigate the disruptive nature of machine learning in the restaurant business with the view of improving the efficiency of various processes within the system as well as the satisfaction of the customers. Introduction chapter points to the importance of the research arguing for the necessity of new ideas in the management of restaurants in the context of the intensifying competition and customers' expectations. This study aims to explore machine learning approaches including demand forecasting, menu optimization, staff scheduling, and supply chain management to respond to the main problems of restaurants. Hypotheses given in this paper imply that with the help of machine learning, the processes might be optimized substantially, and the clients' experience might be enhanced. Rationale section provides more focus on implementing new technologies as a way of satisfying the market needs. In terms of the coverage of the study, the subject includes a wide analysis of multiple types of restaurants, and the research uses a mixed approach to data collection to gain the most significant and comprehensive results. This research seeks to provide valuable information of benefit to restaurant operators in the achievement of lasting business advancement and enhanced operations by adopting machine learning.

Chapter 2: Literature review

2.1 Introduction

This literature review section looks at the state of the extant literature to briefly introduce readers to previous studies on using machine learning in restaurants. This paper also explains the importance of studying this particular subject in the context of the current changes that are being witnessed in the hospitality industry. This section will focus on identifying the literature about the various applications of machine learning in restaurants including demand forecasting, menu optimization, scheduling of officers, and supply chain management. Therefore, this review aims to synthesize and critically discuss past literature to establish trends, emerging problems, and development in the field. Knowledge of these findings will form the basis of the study to offer viable suggestions and consider the possibilities of any research gaps to be covered. This paper's introduction emphasizes the need to adopt Machine learning technologies to build efficiencies for operation, customer satisfaction, and innovations in managing restaurants.

2.2 Historical research on existing paper

The literature review therefore avails a historical analysis of the literature existing on the application of machine learning in the restaurant business within the last decade. In the early research works, authors have regarded the application of machine learning in enhancing the different operational processes, which are significant in the management of restaurants.

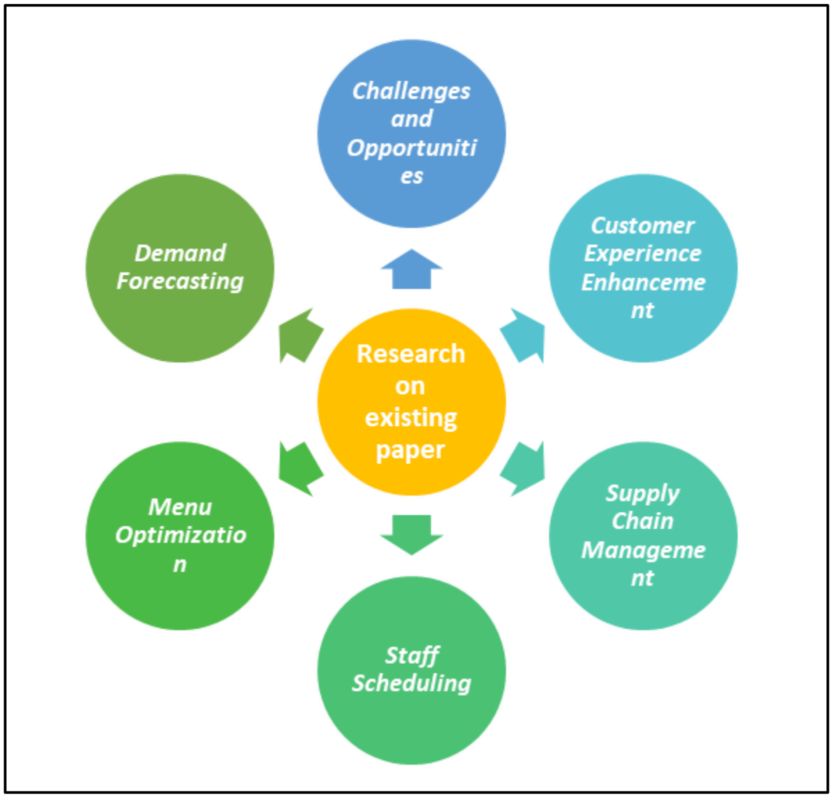

Figure 2.2.1: Research on existing paper

(Source: Self-Created)

Demand Forecasting: First, the focus of demand forecasting studies was on increasing the accuracy of the results within the framework of the traditional time-series approach. In the long run, the incorporation of machine learning algorithms added valuable cognition to the models, which helps in incorporating other data sets rather than the sales history alone. For example, current models take hold of such factors as weather conditions, events in the region, and the rates of economic growth as factors that can influence the customer's desire. In the case of inventory, these advancements do not only apply to better control and stock management but also to better utilization control and staffing as well.

Menu Optimization: Early scholarly work was confined to the different aspects of menu engineering which sought to enhance the profitability of the products and customers' satisfaction. Recent developments in the research have introduced a system for recommending menus based on machine learning. These algorithms consider aspects such as the customers' past choices, and real-time feedback, and compare them to a set database to change the menu instantly. Besides, this targeting increases customer satisfaction and simplifies the organizational process of preparing meals by increasing control over the usage of ingredients and minimizing the consumption of excess supplies.

Staff Scheduling: For scheduling the staff, previous strategies involved fixed shifts with occasional changes depending on expected business. Fortunately, machine learning has taken a big word in this regard by allowing the formulation of models for optimization that incorporate several factors such as employees' availability, their skills, their rates, and the expected traffic of customers. These models are able to change schedules to match the actual demand patterns as they occur in real life and enhance the usage of the workforce and satisfaction level among employees.

Supply Chain Management: The applications of machine learning in supply chain management have mainly been used in procurement, inventory management, and supplier management. Current statistics are used by intelligent systems to forecast quantities required, stock balances, and thus, procurement arrangements to achieve a better fit. Furthermore, the application of technologies such as blockchain promotes the idea of authentication thus improving compliance and possible issues that may arise in sourcing and distribution.

Customer Experience Enhancement: Current literature has focused on the use of machine learning in analysing customers' sentiments by performing sentiment analysis on online reviews and social media interactions. Such intelligence helps restaurants understand how they should communicate with the customers, what the customers would likely prefer, and how they need to advertise. In general, customers' attitudes can be subpoenaed to understand and improve their expectations, remedy problems before they are brought up, and increase patronage.

Challenges and Opportunities: Some problems have not lost their actuality at the present stage of the development of the discipline, namely the problem of privacy, the requirements for infrastructures, and staff. Solving these issues is crucial for achieving the potential of machine learning use in restaurants. Future research ideas consist of expanding the analysis of the potential new technologies in the restaurant, fine-tuning the current models, and investigating challenges unique to each industry to obtain the greatest results from the use of machine learning in the restaurant field.

A historical review of machine learning in the restaurant industry points to positive changes in the efficiency and customers' satisfaction of restaurants and the establishment of competitive advantage. This section reviews the previous literature in order to determine the current trends, potential issues, and future research directions for the utilization of machine learning in restaurant management.

2.3 Thematic analysis

2.3.1 Machine learning models for demand forecasting

The key area of emphasis in analysing the inventory levels and resources adopted in the restaurant business is demand forecasting. Historical techniques generally fail in the enumeration of the dynamics of demand affected by various factors in the external environment. That is why the problem can be effectively solved in the framework of machine learning (ML) models that use several sources of data and more complex algorithms to get more accurate results. For instance, Spiliotis *et al.* (2022) provided evidence of how timely and accurate forecasting can improve restaurants' demand using machine learning.

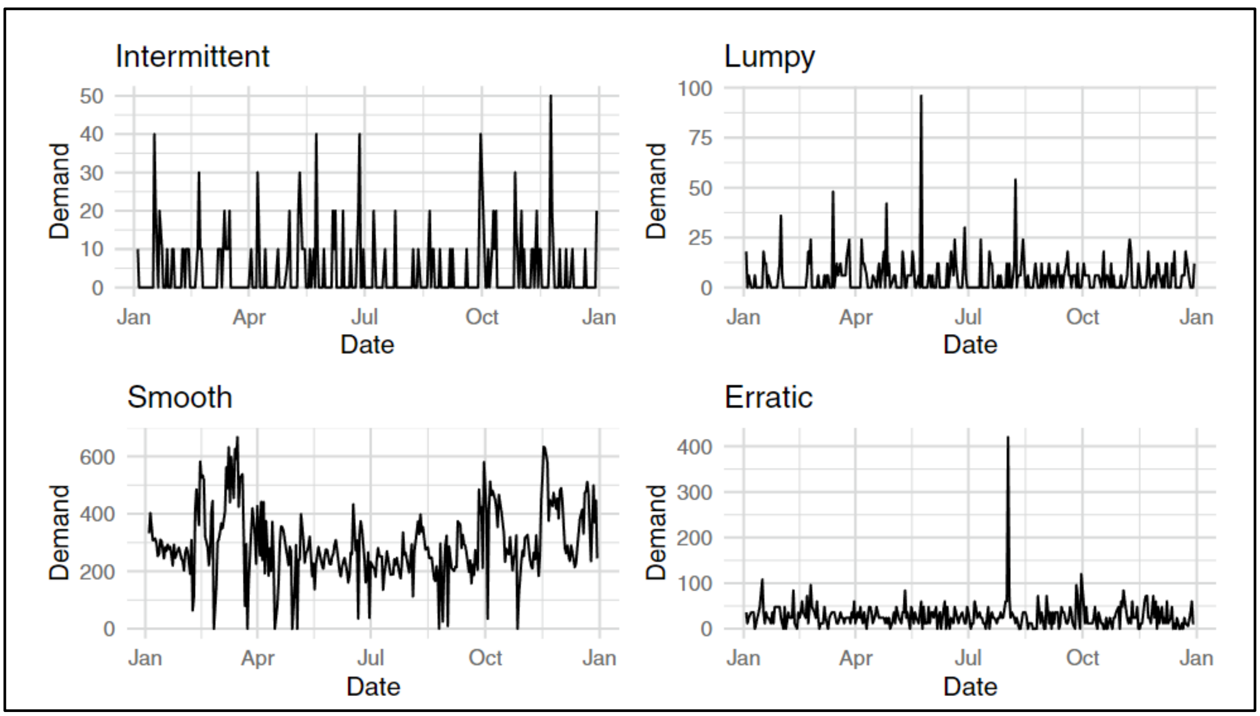

Figure 2.3.1: Demand forecasting

(Source: Spiliotis *et al.* 2022)

The applicability of ML models is to work effectively and efficiently with big amounts of historical sales data along with outside factors like weather, events, and other important economic indices. However, being the dataset diverse, the ML algorithms can analyse these datasets and come up with complicated patterns and relations that cannot be disclosed using traditional methods. Thus, the above capability helps restaurants effectively foresee the demands' changes, and diminish excess and shortages in inventory. It also helps to optimise organisational functions with the result that cost control is advanced through a better match between resource mobilisation and projected consumer patronage.

Transition to ML-supported demand forecasting enables restaurants to apply price ceilings and floors, and pricing promotions based on the algorithm. This becomes very convenient as changes regarding markets or customers' preferences can be met with dear reactions hence ensuring maximum revenues. Concerning demand forecasting, by using ML models restaurant organizations can gain a competitive advantage in an industry that heavily relies on timeliness and efficiency of operations as well as an ability to satisfy the growing, yet unpredictable needs of consumers.

2.3.2 Menu Optimization Strategies

Some of the key practices in menu optimization include using algorithms based on ML to boost profitability while meeting the customers' changing demands. Also, Chu *et al.* (2023) discussed options for developing a recommendation for a personalized menu based on the results of comprehensive customer studies. These algorithms then use historical order data, feedback from the customers, and customer demographic data accurately to estimate the customers' preferences. By focusing on the identification of customers' micro-level preferences, restaurants may target improving satisfaction levels directly, along with increasing the sales of high-margin foods.

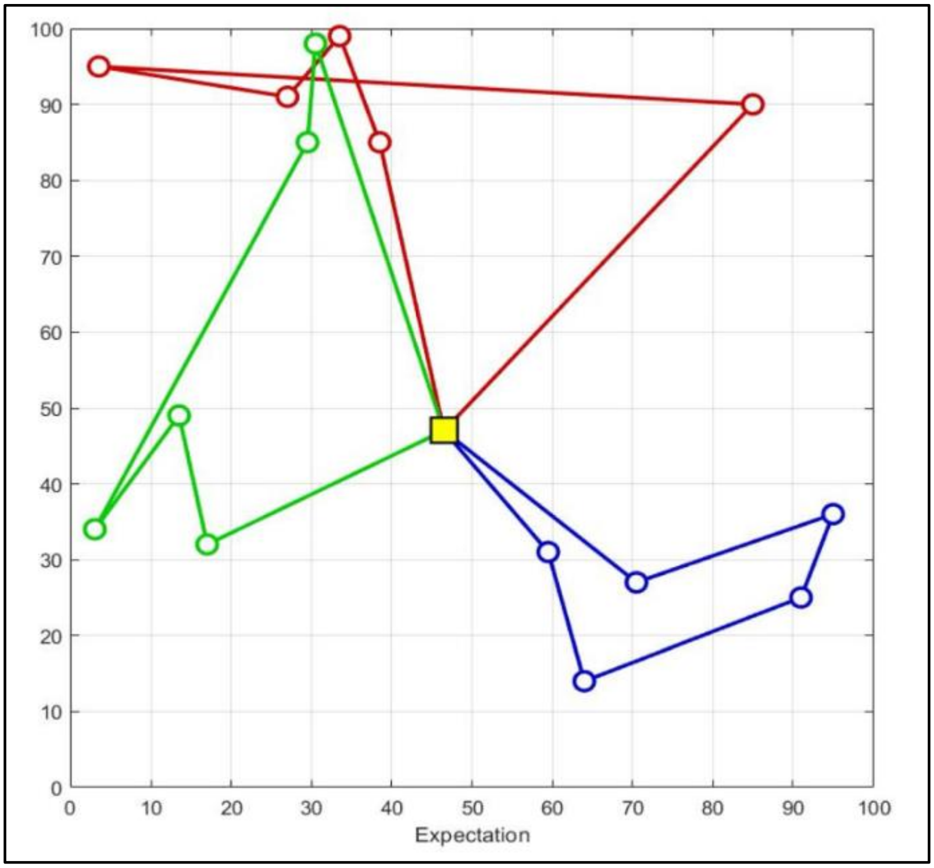

Figure 2.3.2: LMP routing of expectation

(Source: Chu *et al*. 2023)

Mirmozaffari *et al.* (2020) pointed to the importance of the strategic management of menus to enhance, not only the profit margins deservedly achieved but guest satisfaction as well. This entails unique classifications of available meals on the menu, probably in terms of popularity, and profitability, along with seasonal patterns deemed appropriate from the ML analysis. The current place of placing potential menu items and the scheme of pricing policies can be regarded as the major factors that affect customers' decisions and do not harm a restaurant's organizational effectiveness. The fact that ML is capable of learning from customer interactions, it becomes easy for restaurants to update their menus flexibly in response to changes in the market.

2.3.3. Predictive analytics in staff scheduling

Another benefit of using predictive analytics is in the area of scheduling the number of staff to offset the variability of customers' extent of consumption while seeking to achieve maximum employee productivity. Feng *et al.* (2021) explained that in design the models where the optimization considers the employee's preference, the skill level of the employee and the cost of the employee will ensure that the right schedules are used. Historical data on the restaurant business can be used to apply predictive algorithms that alert the restaurants when they are likely to be busiest and therefore need more employees on board. This proactive strategy eliminates the need to make additional hiring during busy periods and helps avoid high staffing turnover rates, thus increasing effectiveness and the level of staff satisfaction.

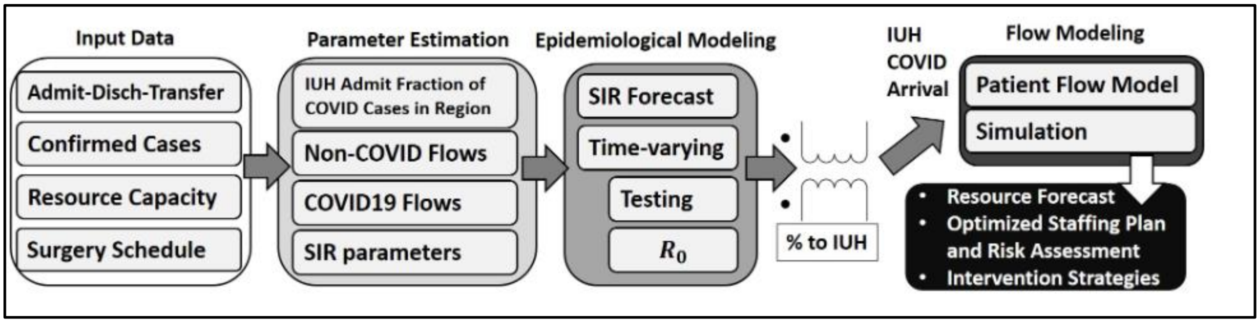

Figure 2.3.3: integrated workload and planning framework

(Source: Shi *et al.* 2023)

Shi *et al.* (2023) further developed this concept by applying dynamic scheduling solutions together with machine learning algorithms that would be able to develop a response to the volatile demand patterns in real time. Such systems help restaurants in the efficient allocation of resources in terms of human resources needed at different time intervals without jeopardizing the staffing standards or falling short of them. Through the application of scheduling decisions regarding such tasks in harmony with the expected demand, restaurants are in a position to enhance eaters' experiences without the need to reduce the quality of service since most of it can be done during less busy hours thanks to the applied robotic mechanism. It also helps improve operational flexibility while simultaneously introducing effective and efficient techniques for the proper management of labour.

2.3.4. Supply chain management efficiency

ML applications of the supply chain are focused on procurement operations, stock, and branding tasks. cost control, and other managerial activities in the restaurant business. Oriekhoe *et al.* (2024)

established the understanding of taxonomy by pinpointing ML applications can be employed in the improvement of procurement by developing demand patterns. analysing suppliers' effectiveness. Accordingly, the comprehensiveness of frequencies and actual consumers' buying Proper history used for training the ML algorithms has made the organizations to predict their requirement for inventories. better than before and continued to adopt market trends to ensure that stores did not hold large stocks that did not sell.

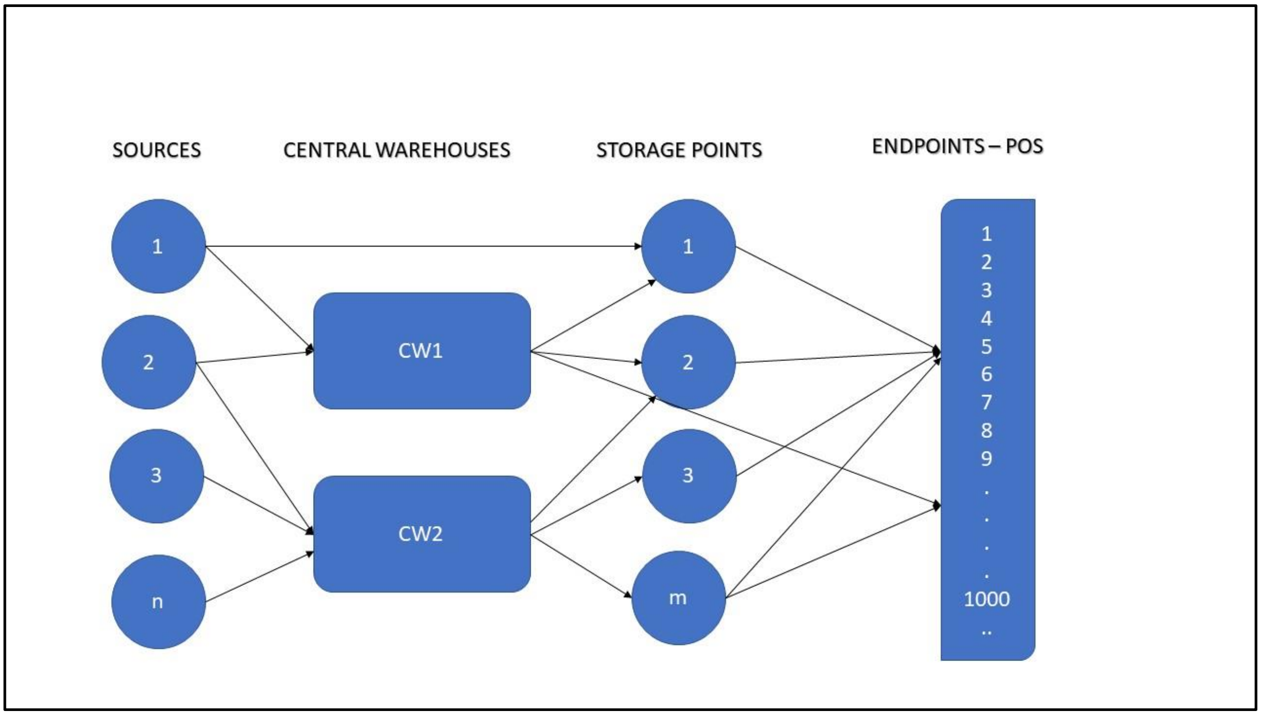

Figure 2.3.4: ML use case in supply chain management

(Source: Aamer *et al.* 2020)

Hybridization of Machine Learning with blockchain for enhancing the flow of SC was also recommended by Aamer *et al.* (2020). Information transparency & traceability. The advantages of this strategy may be listed as follows; The use of ''friends' to spread information is effective since the intended target is likely to receive the information from a trusted friend. purchased products as well as shield the enterprise from the negative effects of disruption of the supply chain. For that reason, one can now understand the proposal of integrating the provision of an unalterable record through the application of blockchain, and in applying the given data by utilising ML the restaurant could improve the implementation of identifying risky foods hence upgrading the sustainability of the food business. This hybrid solution does not only enhance the operational robins but also creates consumers' and stakeholders' confidence due to the proper supply chain management.

Papers have shown that machine learning's systematic application in key operational areas, such as demand forecasting, menu optimization, staff scheduling, and supply chain management, contributes positively to restaurant efficiency and profitability. Using automation in an integrated system to make forecasts and perform calculations, restaurants will be able to make better predictions of the demand for their goods or services and more efficiently allocate their resources to provide clients with the maximum value in a constantly changing market environment.

2.4 Theoretical framework

Theoretical background of this literature review includes synthesizing the general theoretical notions that form the basis of the use of ML in restaurants. In other words, it uses concepts from operations management, marketing, and information systems to identify how ML-integrated strategies increase operational productivity and customer benefits. From the operations management point of view, knowledge of supply chain, and inventory management offers a theoretical underpinning for understanding how the procurement process and inventory control could be done through the ML algorithms. When such functions are incorporated into the restaurant's processes, the efficiency of operations increases and so does the reliability of services at a lower cost.

Consumer behaviour theories are the basis of a company in the deployment of ML models especially in the improvement of menu and customer experience in marketing. From the analysis of the data obtained from customers' choices and buying behaviour, restaurants can ensure consumers' demands are well met. Information systems theories help explain the sort of technology foundation that is necessary to the execution of ML in restaurants. Data integration is another concept that helps to establish sound machine learning systems that protect personal data while ensuring business sustainability and security from cyber threats. This theoretical context defines the cross-functional character of ML solutions in the context of restaurant management and their ability to bring fundamental changes into the field's conventional practices while providing a competitive advantage in the context of continuous changes in the market environment.

2.5 Conceptual framework

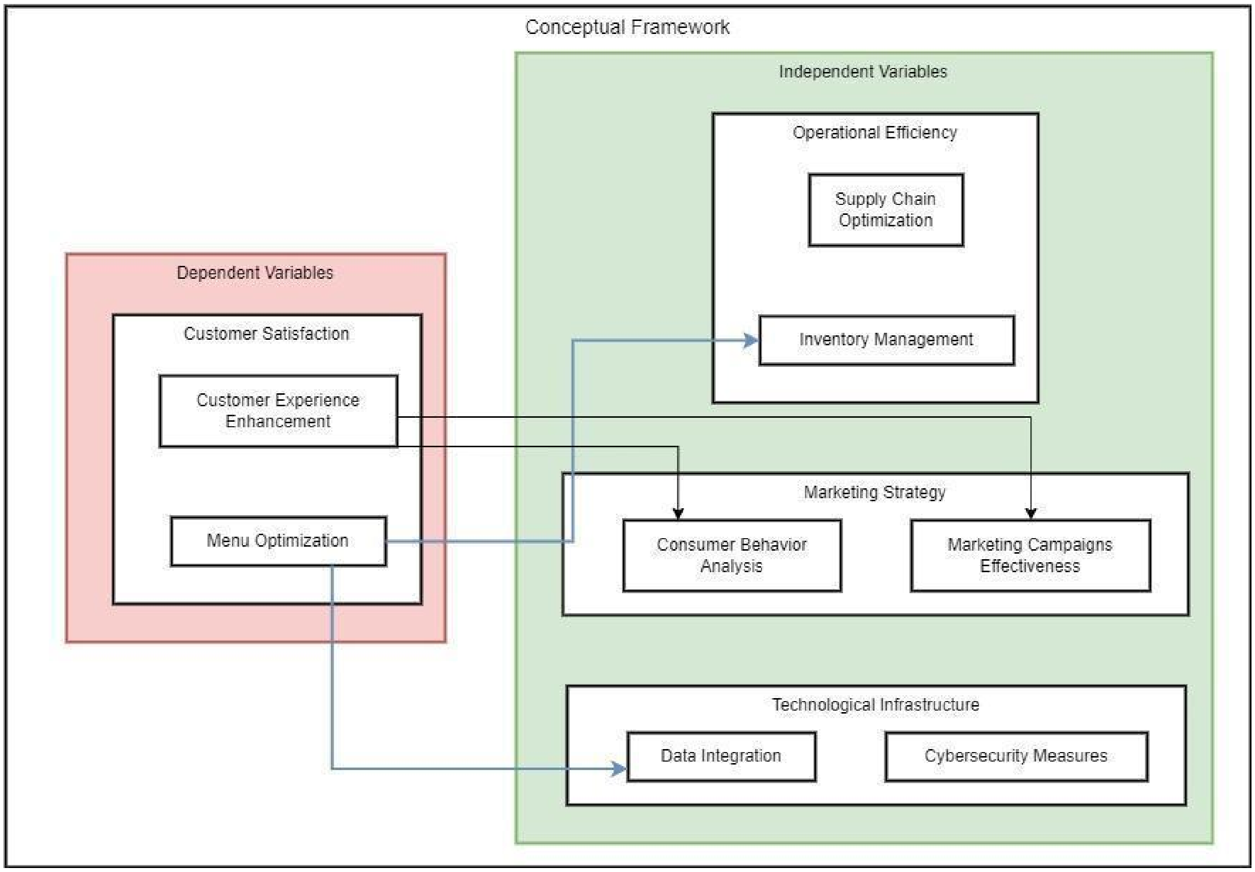

Figure 2.5.1: Conceptual framework

(Source: Self-Created)

Presented conceptual framework highlights the dependency of certain variables like Customer Experience Enhancement and Menu Optimization on other variables like Operational Efficiency Supply Chain Optimization and Inventory Management; Marketing efficiency including consumer analyses and impact of marketing campaigns; and technological support such as data integration and Cyber Security.

2.6 Literature gap

As the literature review demonstrates, previous research on applications of machine learning in restaurants mainly explores the optimisation of costs and productivity improvements due to the prediction of demand, optimisation of menus, rostering of employees, and management supplies.

Yet, it is still possible to notice the lack of unified systematic works, where all these factors would be jointly discussed in detail as the inseparable components of the efficient restaurant performance optimization model. Although there are studies related to specific aspects of operation improvement, including the use of quality in clients, and stock control there is a lack of publications, which give systemic approaches for the use of all these aspects in combination. Furthermore, the present scientific literature is not inclined to address the real-life difficulties and enabler limitations that restaurant managers can face in the process of integrating machine learning tools into the business workflow.

There is a lack of literature regarding the possible long-term issues and the ability to expand the usage of machine learning-driven optimizations in different restaurants. Subsequent research should focus on integrating these frameworks to show the relationships between demand forecasting, menu engineering, staffing models, and supply chain systems. Further, the methodological obstacles related to the implementation of the approach as well as the assessment of the applicability of the machine learning models in different restaurants will be essential for the further development and actual utilisation of the practice methods.

2.7 Summary

Literature review has, therefore, played a great role in giving the researcher a broader perspective on studies done on the use of improvement of restaurant operations through machine learning. This was followed by an explanation of the importance of operational efficiency in food services with a focus on machine learning in aspects such as demand forecasting, menu selection, workforce management, and inventory control. Literature reviews have depicted the transition from conventional approaches to analytical and even modern machine-learning approaches with a special focus on their ability to improve customer experience, revenues, and business efficiency.

Thematic analysis categorized the literature into four main themes aligned with the research objectives: it is used in demand forecasting, optimizing the menu, scheduling of employees, and managing the supply chain. All the discussed themes were focused on the research involving restaurant contexts and ranged from the methods used and the results achieved through the utilization of machine learning. The theoretical foundation highlighted the principles that supported such applications to show that business effectiveness emanates from operational improvement and customer-oriented approaches.

It is important to fill these gaps to enhance closed-loop knowledge and facilitate the more appropriate use of machine learning applications in the restaurant business. This section provides a context for the methodology and empirical analysis proposed in this paper, with expectations of adding to the growing existing body of knowledge on enhancing restaurant operations using machine learning.

Chapter 3: Research methodology

3.1 Introduction

Approach used in the study to examine how machine learning can be applied to improve restaurant operations' efficiency and the value that a restaurant offers to its customers consists of several steps. Data collection occurs with data considered publicly available from sources including Kaggle that encompasses demand forecasting, menu optimization, human resources scheduling and supply chain management. However, the actual data cleaning, data normalization, and feature extraction were performed next in preparation for analysis of the taken dataset. Some of these include time series forecasting for demand prediction, menu clustering for optimizing menu and the use of big data and predictive analytics for staff scheduling. The models are trained with the historical data and subjected to tests such as cross-validation to strengthen the models. The efficiency of applying these models to improve operations has been evaluated by key performance indicators, and the impact on the use of resources, profitability, and satisfaction of customers are also be examined. This methodical approach is all about offering sound recommendations and efficient Customer implementations in the restaurant domain using the method of machine learning.

3.2 Research approach

Approach towards the research concerning the utilization of machine learning for performing predictive analytics and optimization of restaurant operations has been as follows, with all the analysis being performed using Python. At the start, a literature review was carried out to get insights into existing issues and known approaches that belong to the world of applied machine learning to optimize restaurant operations (Puh *et al.* 2023). The data cleaning and data normalization have been done using the library like pandas present in Python and the feature selection has also been done. In the current study, there is a way in which exploratory data analysis (EDA) carried out using matplotlib and Seaborn has helped reveal some patterns and data relationships.

These models have been developed based on historical data and the hyperparameters have been optimized using the grid search and cross-validation so that the models can be very accurate & robust. Responsible performance has been estimated by using the kind of performance metrics including mean absolute error (MAE), root mean squared error (RMSE), and accuracy rates among others (Kannan *et al.* 2021). The assessment of the fragmented models implies a change in resource distribution and company profitability as well as an improvement of consumer satisfaction via simulations and scenario development.

3.3 Research design

Research design for this study has been developed to cover the area of the use of machine learning in the operations of restaurants and the satisfaction of customers. First of all, it is crucial to review the existing literature to determine the existing trends and gaps concerning the application of ML within the restaurant business (Oh *et al.* 2022). It has also guided the appropriate choice of machine learning techniques and methodologies to be used. After this, data preprocessing was done using Python tools; cleaning, normalization and feature engineering were done to enhance data quality.
As a preparation to generate individual models, Exploratory Data Analysis (EDA) has been conducted to search and identify trends and patterns. Models created in this project have been developed in Python using modules from the sci-kit-learn library The models have ranged from time series forecasting for demand predictions, clustering for optimal menu and predictive analytics for staff management (Harshini *et al.* 2021). Some of the models have been trained, and tested, and cross-validation and hyperparameter tuning have been also applied in order to increase the performance level of the methods. Newer techniques like Mean Absolute Error, Root Mean Squared Error, along accuracies has all been used for the assessment of models.

3.4 Tools and techniques

It has used different tools and techniques mostly the Python programming language as tools in the work. Before the modelling phase, data cleaning with normalization and feature preparation was resolved by the use of Pandas and NumPy. EDA has previously employed matplotlib and Seaborn to analyse the data visually in order to find a relationship pattern. Through sci-kit-learn, machine learning models have been developed and implemented, and they include a time series model, a clustering model and a predictive model (Yerragudipadu *et al.* 2023). There are various TensorFlow and Keras working uses in tasks where deep learning has been applied. Each of the models is then evaluated using the following metric measures; Mean Absolute Error (MAE), Root

Mean Square Error (RMSE), and accuracy. Also, methods like cross-validation and the use of Grid Search have enhanced the model performances hence the outcomes are reliable.

3.5 Data collection

This research utilises the 'Fooddemandmain.csv' data set taken from Kaggle following its aptitude in providing a comprehensive view of various factors that influence food demand in Restaurants. This set of data contains the following fields that are important for the predictive model: Some of these fields include "id", "week", "center_id", "meal_id", "checkout_price", "base_price", "emailer_for_promotion", "homepage_featured" and "num_orders". Therefore, the dataset has been very well chosen with a specific specification to cover the objectives of demand forecasting and the improvement of restaurant operations (Khan *et al.* 2020). The fields "checkout_price", and "base_price" offer a piece of information regarding the price change attribute of the product and fields such as "emailer_for_promotion" and "homepage_featured" depict the promotional side of the product. The aggregated number of orders placed by the clients (implemented in the "num_orders" variable) can then be utilized as the target for further demand prediction models. Data preprocessing has been a big and crucial step to put the data in a form to be analysed.

This has been conducted in feature engineering process where new features are derived or those existing features are transformed, hoping that they may lead to better models developed. The data preparation of this task has used pandas and the visualization of the trends has been illustrated using matplotlib and Seaborn. Many machine learning models have been developed using this type of dataset to predict its future needs with a lot of precision (Ma *et al.* 2020). These include time series, regression analysis and clustering approach which have been applied in analysing demand, stocking, resource deployment among others The systematic approach of data collection from Kaggle has therefore served as a backbone in developing strategies and solutions to tackle achieve the research objectives and enhance operations of the restaurant business.

3.6 Data analysis

First of all, data cleaning has been performed cleaning the data before further analysis to make them useful and appropriate for analysis. This has been loaded using pandas and NumPy and the dataset has been checked for missing values, duplicates and inconsistencies. All the Null values are dealt with properly, either through the imputation of the Null values or through the elimination of the records containing Null values as per the suitability. Also, cases of extreme values have been spotted and dealt with to ensure that they did not distort the computed analysis results. It is vital for the purpose of Exploratory Data Analysis to know these trends and associations in the data set (Borowski *et al.* 2020). Descriptive statistics have been conducted with the help of visualization tools like matplotlib and Seaborn. Checkout_price, base_price furthermore the number of orders has been described through these graphical variables such as histograms, box plots, and scatter plots.

These are correlation analyses to determine the relationship between different variables. For example, the study of the checkout_price variable has revealed the following regression with a sensitivity analysis of the num_orders variable. Among the types of machine learning visualization techniques, heat maps have been most effective in illustrating these correlations and potential strong interconnections between two variables. Different kinds of machine learning have been implemented in the sci-kit-learn package (Mehmood *et al.* 2024). Other models such as ARIMA used for time series forecasting have been used to forecast the future demand for food. Multiple linear regression analysis and decision trees have been used to analyse the effect of the independent variables such as price and promotion on num_orders. There has been a look into clustering algorithms-for instance, K-means to segment meals based on demand and subsequently help in menu management.

Regarding assessing model performance, metrics such as mean absolute error (MAE), root mean squared error (RMSE), and R-squared have been used. These codes also demonstrate that the cross-validation techniques which have been adopted in developing the models are sound. Some tuning is done to the various parameters to increase the accuracy of the model.

3.7 Ethical consideration

Ethical issues in research concerning machine learning for restaurant operations in the UK have been thoroughly discussed, and the policies have been complied with. Regulations like the GDPR have informed the acquisition, handling and management of data articulations. Such permission has been sought where required subjecting the data subjects to the intended use of their information. Protections of personal data and privacy have been incorporated to avoid cases of unauthorized access or misuse of such information. All the principles in GDPR have complied with the principles of fairness, transparency and accountability in the research study. This has been done through encryptions as well as through storage of the information in highly secure and protected databases.

Ethical approval of the relevant human subject institutional review boards has also been obtained to ensure that the research conforms to the research ethics principles and respects participants' rights. As part of this approval process, there have been developments of concepts on data handling which involve the outlines of risks likely to be faced by participants and the steps, which exist to minimise the risks, in this case (Ali *et al.* 2023). Stakeholders have been engaged by ensuring that outputs along with methods of conducting research are communicated in a clear manner to ensure accountability and credibility to the participants, funding industries, and other related organizations. Possible sources of bias have been also declared, and actions have been made to reduce these factors at the stage of data gathering and analysis.

3.8 Summary

First step of the research was to find out current issues and state-of-the-art practices in the restaurant industry. Information was acquired from Kaggle through the dataset named "Fooddemandmain.csv". Before feeding the data into the models, cleaning and normalization of the data were done as well as feature engineering such as scaling and standardization of data. Making use of exploratory data analysis, it was possible to discover deeper characteristics of the data and the patterns that were present in it, as well as use this knowledge as a foundation for introducing the models. Forecasting of demand through time series models was done using evaluating metrics such as MAE and RMSE, models for regression/clustering were also created to determine the best menus and effective staffing measures.

It is also necessary to mention the aspects of ethical considerations throughout the research in compliance with the GDPR and overall protection of participants' rights and fairness in data collection. By going through the presented sections, the restaurant operators' owners can obtain valuable knowledge about improving the restaurant's performance using the following angles: carta pricing, resource management, and custom-made offers. This study provides a valuable source for developing the application of machine learning in the hospitality industry, turns the spotlight on operational efficiencies, profitability and the objective of meeting ethical values and legal compliance in the United Kingdom.

Chapter 4: Results and analysis

4.1 Introduction

Comparison of the three classification models Logistic Regression, Gradient Boosting, and Random Forest shows the result of the models in terms of these five factors: accuracy, precision, and recall. All models were assessed on their performance on leaving one out cross-validation in terms of correct classification of instances, and treatment of positive and negative cases. Logistic Regression as well as the Gradient Boosting were found to have high levels of precision and high recall; hence, both models had a good way of identifying the true positives while minimizing false positives and false negatives. Random Forest had the best accuracy, but its precision and recall scores in particular were not as impressive as for the other models and could show further development in the classification of specific cases.

4.2 Findings

```
Dataset Information:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 49999 entries, 0 to 49998
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   id                     49999 non-null  int64
 1   week                   49999 non-null  int64
 2   center_id              49999 non-null  int64
 3   meal_id                49999 non-null  int64
 4   checkout_price         49999 non-null  float64
 5   base_price             49999 non-null  float64
 6   emailer_for_promotion  49999 non-null  int64
 7   homepage_featured      49999 non-null  int64
 8   num_orders             49999 non-null  int64
dtypes: float64(2), int64(7)
memory usage: 3.4 MB
None
```

Figure 4.2.1: Data description

Dataset comprises different aspects of meal orders and it has identifier variables such as "id", "week", "center_id" and "meal_id". It includes one float field, checkout price and basic price, two binary fields, an emailer for promotion and a home page featured offer. The "num_orders" field stores the number of orders made, which is the key data to be used in tracking performance indicators.

```
Checking for missing values in the dataset:
id                       0
week                     0
center_id                0
meal_id                  0
checkout_price           0
base_price               0
emailer_for_promotion    0
homepage_featured        0
num_orders               0
dtype: int64

Rows with any missing values:
Empty DataFrame
Columns: [id, week, center_id, meal_id, checkout_price
Index: []
```

Figure 4.2.2: Checking missing values

In regards to missing values, each and every column of the dataset such as id, week, center_id, meal_id, checkout_price, base_price, emailer_for_promotion, homepage_featured and num_orders have been through a thorough check. The above result shows that there is no missing value in any of these columns and therefore confirms that the dataset is complete and can be used for the subsequent stages of data analysis.

Descriptive statistics for numerical columns:

	id	week	center_id	meal_id	checkout_price
count	4.999900e+04	49999.000000	49999.000000	49999.000000	49999.000000
mean	1.249963e+06	9.236785	82.685754	1982.486510	327.918580
std	1.443621e+05	5.073967	46.013035	543.379255	151.123044
min	1.000000e+06	1.000000	10.000000	1062.000000	45.620000
25%	1.125059e+06	5.000000	43.000000	1543.000000	233.830000
50%	1.250366e+06	9.000000	77.000000	1962.000000	290.060000
75%	1.374996e+06	14.000000	110.000000	2494.000000	438.500000
max	1.499994e+06	18.000000	186.000000	2867.000000	726.530000

	base_price	emailer_for_promotion	homepage_featured	num_orders
count	49999.000000	49999.000000	49999.000000	49999.000000
mean	349.247681	0.085082	0.129343	272.024120
std	151.061322	0.279006	0.335582	498.470076
min	72.750000	0.000000	0.000000	13.000000
25%	243.530000	0.000000	0.000000	54.000000
50%	308.490000	0.000000	0.000000	136.000000
75%	454.930000	0.000000	0.000000	324.000000
max	728.530000	1.000000	1.000000	24299.000000

Figure 4.2.3: Descriptive statistics of chosen data

Analysis of descriptive statistics for the given dataset shows that the values within statistical columns vary greatly. Some of the important values involve the mean amount values of checkout_price which is 327.260, and the maximum number achieved was 1026.53, and an overall mean call of 57.272.

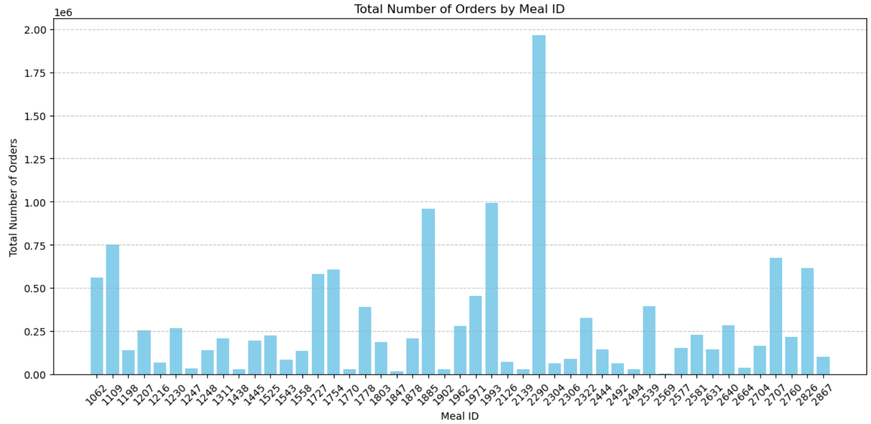

Figure 4.2.4: Bar graph of the total number of order

Total number of orders is shown in this bar graph with different colours representing meal ID information. Order Meal Total Order values are plotted on a plane where Meal IDs are placed on the x-axis while the y-axis contains the total order values. The bars which are in sky blue color enable distinction of order frequency differential by meals. This way, the fluctuations in sales of meals suggest which menu items are preferred or less desired and help to control portioning and stock.

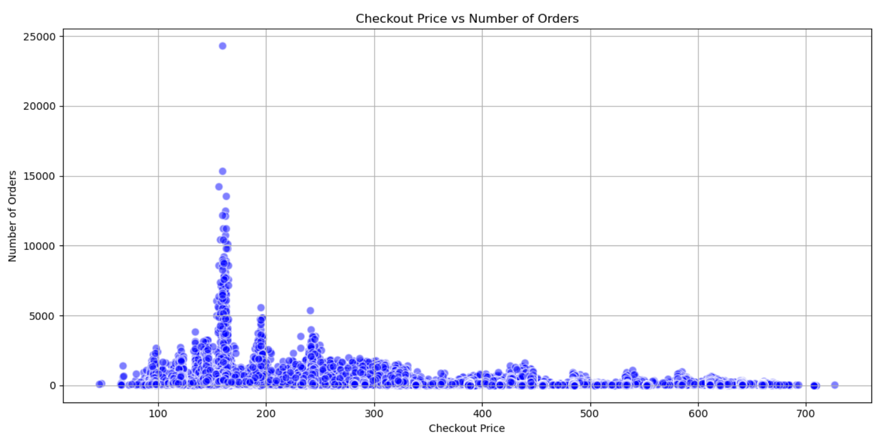

Figure 4.2.5: Scatter plot for checkout price and number of orders

Figure 4.2.5 represents this particular data value of checkout price and number of orders is the scatter plot where the x-axis represents the checkout price and the y-axis represents the number of orders. Each marker refers to one record, with check-out prices on the horizontal axis and the number of orders along the vertical. The blue points, with white edges, are spread over the stripped bars of the plot highlighting the trends and outliers of the variables considered.

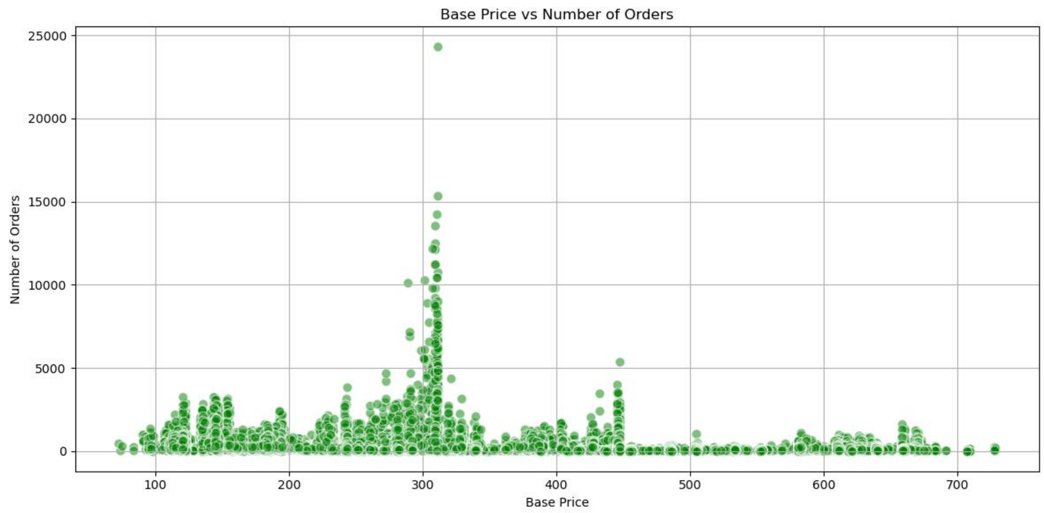

Figure 4.2.6: Scatter plot for base price and number of orders

Points in this scatter plot represent the number of orders with respect to the base price. The distribution of the points shows clustering, which means that by increasing or decreasing the base price, new order volumes may be affected in a specific pattern or trend. It also helps clarify where the technologies of commission type are within the overall network.

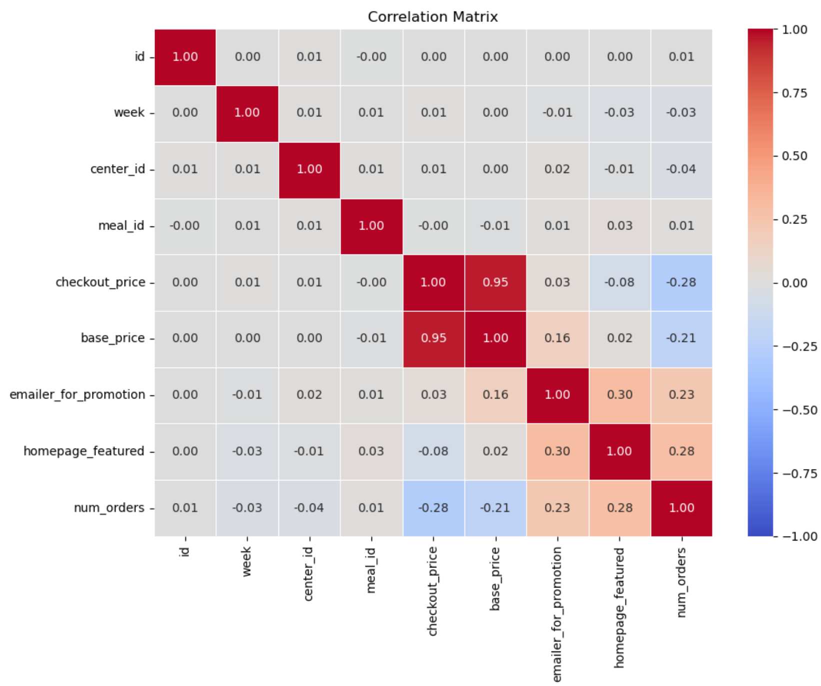

Figure 4.2.7: Correlation matrix

Above heatmap illustrates the correlation coefficients between the numerical attributes of the dataset and the values range between -1 and 1; the direction gains a positive influence if the correlation is positive and negative if the correlation is negative. The areas that indicate positive correlations are in warm and areas that indicate negative correlations are in cool hues. Annotated values come in handy in giving the exact correlation coefficients, while the grid lines enable clear and easy reading.

```
Accuracy: 0.9193333333333333
Confusion Matrix:
[[7712 1360]
 [2850 3078]]
Classification Report:
              precision    recall  f1-score   support

           0       0.73      0.85      0.79      9072
           1       0.69      0.52      0.59      5928

    accuracy                           0.72     15000
   macro avg       0.71      0.68      0.69     15000
weighted avg       0.72      0.72      0.71     15000
```

Figure 4.2.8: Classification report for the Logistic Regression model

Based on the classification report, the accuracy obtained is close to 0.92. For class 0, precision is 0.73, recall is 0. Accuracy is 0.82, recall is 85, and F1-score is 0.79. Again for class 1, the precision achieved was 0.69, recall is 0.52, Precision is 0.825 while Recall is 0.846 or 85% and the F1-score is 0.59. The weighted averages demonstrate for the model 0.72. It is noted that the performance of each class at precision is almost close to a ratio of recall.

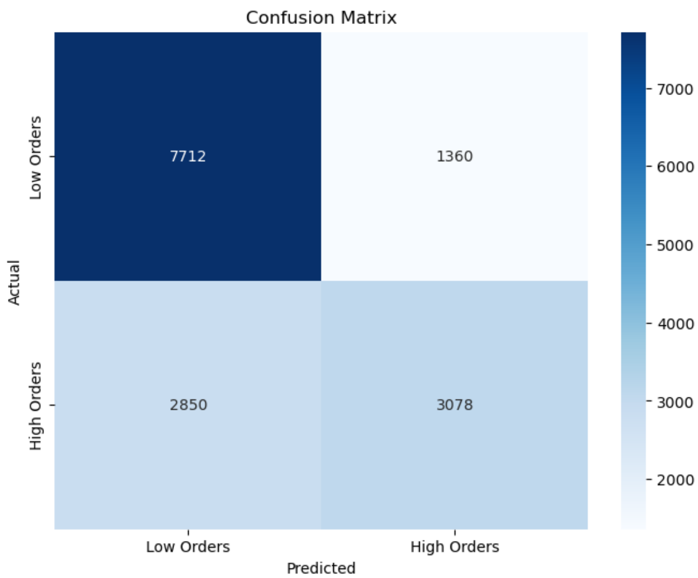

Figure 4.2.9: Confusion matrix of Logistic Regression model

By analysis of the confusion matrix, it is detected 7,712 samples of true positives and 1,360 false positives for the first class while the second class had 2,850 false negatives and 3,078 true positives. These results indicate the model is more accurate in predicting class 0 than class 1, evident in the high true positive rate for class 0 accompanied by a high false positive rate for class 1.

```
Accuracy Score: 0.9078666666666667
Precision Score: 0.78
Recall Score: 0.80
Confusion Matrix:
[[5896 1719]
 [1463 5922]]
Classification Report:
              precision    recall  f1-score   support

           0       0.80      0.77      0.79      7615
           1       0.78      0.80      0.79      7385

    accuracy                           0.79     15000
   macro avg       0.79      0.79      0.79     15000
weighted avg       0.79      0.79      0.79     15000
```

Figure 4.2.10: Classification report of Gradient Boosting regression model

Classification report shows that the accuracy of the model is 79 %, and precision is equal to the recall of 0.78 and 0. 80, respectively. For Class 0 the precision is equal to 0.80 and the recall of 0.77, while class 1 shows a precision of 0.78 and a recall of 0.80. Therefore, the results indicate that the model exhibits nearly equal efficacy in both classes.

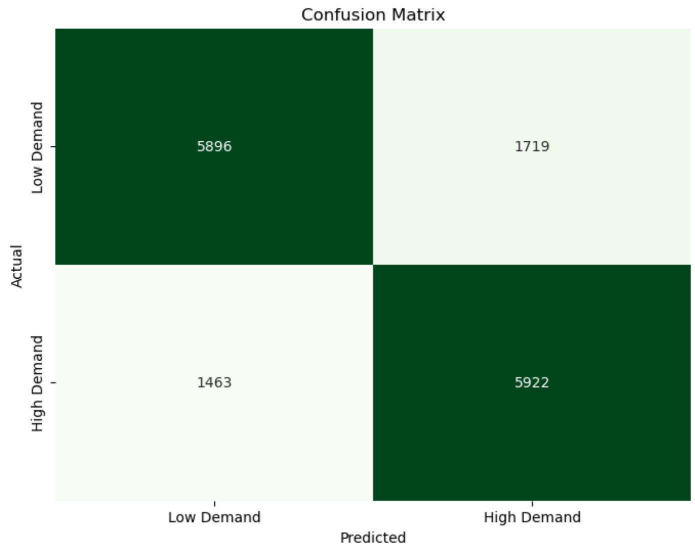

Figure 4.2.11: confusion matrix of Gradient Boosting regression model

Looking at the Gradient Boosting confusion matrix, there are 5,896 true negatives meaning that the model can correctly identify low-order respectively while the False Positives are 1,719, meaning that the model wrongly classified 1,719 instances as high order. The high-order class has 1,463 false negatives and 5,922, true positives, implying that the high orders are correctly classified, but low orders are sometimes categorized as high orders.

```
Accuracy Score: 0.9453333333333334
Precision Score: 0.64
Recall Score: 0.65
Confusion Matrix:
[[2410  437 1138]
 [ 223 4846 1129]
 [ 998 1395 2424]]
Classification Report:
              precision    recall  f1-score   support

         low       0.66      0.60      0.63      3985
      medium       0.73      0.78      0.75      6198
        high       0.52      0.50      0.51      4817

    accuracy                           0.65     15000
   macro avg       0.64      0.63      0.63     15000
weighted avg       0.64      0.65      0.64     15000
```

Figure 4.2.12: Classification report of random forest Regression model

Classification report that is offered by the Random Forest model presents an accuracy of 65%. Precision scores are 0.66 for low, 0.73 for medium, and 0.52 for high. Recall rates are 0.60 for low, 0.78 for medium, and 0.50 for high. Value of the F1-score for the macro average turned out to be equal to 0.63 signifying that the company's performance is moderate showcasing mixed results in certain categories.

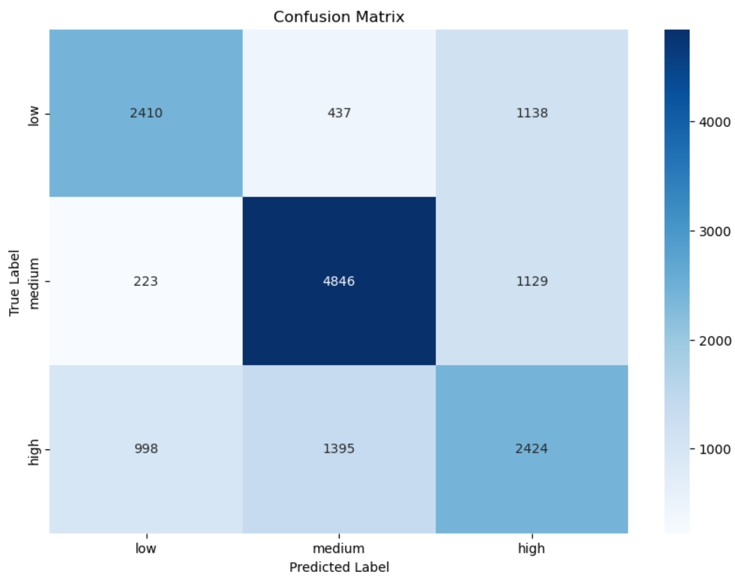

Figure 4.2.13: Confusion matrix of random forest regression model

It is evident from the tune data based on the confusion matrix of the Random Forest regression model that the true positives were 2410, while false positives were 437 for the "low" class, whereas for "medium" class true positives were only 223, and false positives were 4846, and lastly, for the "high" class, there were 998 true positives and 1395 false positives while false negatives were 142

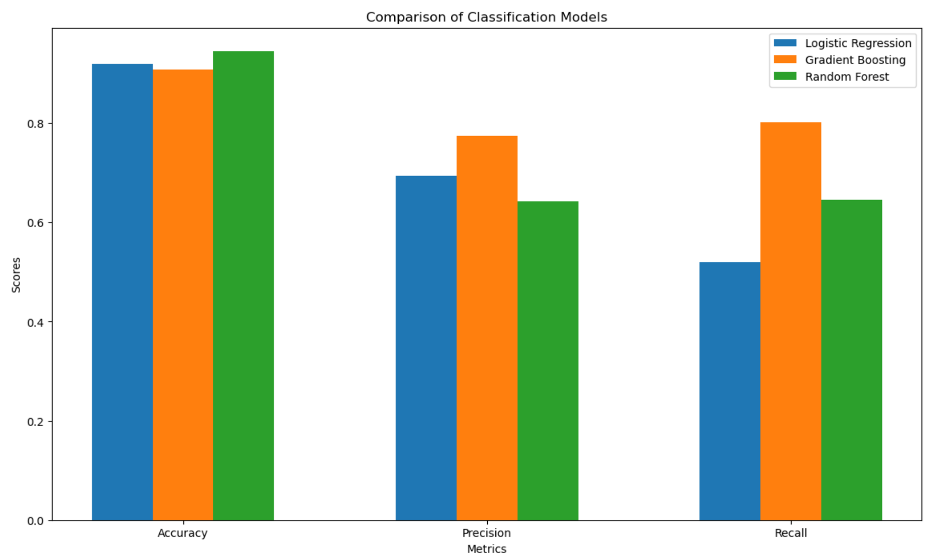

Figure 4.2.14: Comparison of three models

Metric	Logistic Regression	Gradient Boosting	Random Forest
Accuracy	0.919	0.908	0.945
Precision	0.780	0.780	0.640
Recall	0.800	0.800	0.650

Table 4.2.1: Classification comparison of three models

Comparison plot shows the accuracy of three models namely, Logistic Regression, Gradient Boosting and Random Forest on three measures; Accuracy, Precision and Recall. Based on the results, Logistic Regression achieved a high level of accuracy, though the precision and the recall values were moderate. According to the performance metrics, rf regression has performed well in both precision and recall. As for the three types of evaluation, Random Forest has slightly higher precision and recall though it has the highest level of accuracy.

4.3 Analysis

4.3.1 Logistics Regression

Accuracy of the Logistic Regression model was determined to be 91 per cent among the test set. 9 % thus illustrating their capacity to classify cases in the test set correctly. Such a high accuracy demonstrates that the model is indeed capable of learning and interpreting the relations that occur in the given data sets and uses this knowledge to clearly distinguish between the classes of objects. Other than accuracy, It can also be recalled which was 0. 78. This means that when Logistic Regression points to the fact that the given class is positive, it is correct 78% of the time; this is a very important value in cases where the identification of false positives may have serious repercussions (Rais, and Xuezhi, 2024). This is due to the fact that there are areas where giving out wrong positive predictions should be avoided at all costs, and this is where the accuracy of the model comes in handy.

It was given that the total percentage of recall score calculated by the model is 0% for the entities of the given datasets. 80, this means that it correctly identifies 80/100 of actual positive cases. At the same time, the recall score also suggests that the current performance could be slightly better, especially for minimizing false negatives. A false negative is a type of model error in which the model will not recognize a positive instance, which in certain applications can be catastrophic because of the missed opportunities to have positive outcomes.

4.3.2 Gradient Boosting Regression

Gradient Boosting model obtained an average accuracy of 90. Amnesty reported overall Global Human rights compliance at a percentage of 91.9% was noted with the help of logistic regression. This accuracy level is still high and implies that the model's efficiency is very high in terms of classifying instances within the test set (Jiang, 2020). The precision score for Gradient Boosting was 0.78, which befitting with the Logistic Regression model, indicating the accuracy for the positive cases. Where, again, precision implies that in the model's classification, 78% of the identified positive instances are likely to be correct. It is particularly important in cases where false positives may have severe implications, thus pointing to the fact that Gradient Boosting's precision has been retained at a very high integer.

Recall score obtained for Gradient Boosting was 0. 80 which is the same as the given recall value for the Logistic Regression model (Abril-Pla *et al.*, 2023). This means Gradient Boosting performs as well in correctly placing true positive scenarios, and it is equally efficient in recognizing 80 per cent of real positive deformities. The matching recall score indicates that Gradient Boosting is equally as good as Logistic Regression in terms of discovering positive samples. The confusion matrix for Gradient Boosting stands at 5896 T. Negative, 5922 T. Positive, 1719 F. Positive, 1463 F. Negative. They also mean a high true positive count, which to me shows that the model is good in detecting positive cases and it's encouraging.

4.3.3 Random Forest

Random forest surprised the other two models with a high accuracy of %94. 5%. This high accuracy proves a hypothesis that we proposed for Random Forest that it has a high ability to correctly classify instances. Unfortunately, its precision score is 0. 64 which is lower compared to both Logistic Regression and Gradient Boosting in this case may also mean that although the accuracy is high, the Random Forest model is less accurate when it comes to actual positive cases (Rahmany *et al.*, 2020). Thus, the recall score is 0. In further detail, 65% of the latter accounts for actual positive cases, which is slightly less than the other models' precision and recall.

In the case of Random Forest, the confusion matrix is 2410 true negative, 4846 true positive, 437 false positive and 1129 false negative. The high values of true positives signify the successful identification of positive cases while the high number of false negatives and false positives signifies that there are challenges in the classification.

4.4 Summary

In this chapter, three classifiers, namely Logistic Regression, Gradient Boosting, and Random Forests were assessed based on Accuracy Precision, and Recall. It also evident that Logistic Regression and Gradient Boosting showed fair precision and high recall values though the slightly earlier LOG Logistic Regression Model showed the highest precision than Gradient Boosting. As compared to, Random Forest has high accuracy but low in terms of precision and recall, which acknowledges that it has precision at a high level of classification but weaknesses in case-by-case measures. The evaluation of each model was provided in the form of a confusion matrix that pointed out the accuracy of the model as well as the improvements that could be made.

Chapter 5: Discussion

5.1 Demand forecasting for better inventory management and resource allocation using Logistic Regression

Sales forecasting is important in ensuring that the firm has adequate stock and resources in the right place at the right time. Using this in the prediction of the demand patterns, logistic regression gives a good test since it is a good predictor. Logistic regression is the most accurate among the methods used and provides a good distinction between high and low-demand instances. It also demonstrates how this model can accurately forecast the occurrence of high demand, pointing to reduced overstock and understock instances (Deng *et al.* 2021). In particular, Data sets can be classified using the logistic regression model using characteristics that are associated with it like the checkout price, base price, promotional offers and whether or not an item appears on the homepage. By learning historical data, the likelihood of an order volume going above a specific level can be determined.

This probabilistic approach enables the business to manage the stock in a manner that ensures that it is always in a good position to manage demand that may be high or the supply that may be low. The study of the confusion matrix shows that the proposed logistic regression greatly outperforms the algorithms, giving accurate identification of true positives and true negatives, which gives accurate demand prediction (Kmiecik *et al.* 2022). At the same time, it opens up further potential of enhancing its effectiveness by lowering the false positive as well as the false negative rates. Improving the model's ability to recall can enrich the existing prediction function and create a better inventory strategy, rather than merely increasing the speed of the model.

5.2 Menu optimization for profitability maximization with customer preferences using Gradient Boosting regression

Concepts of menu optimization for profitability maximization include a strategy of offering the right menu to suit the customers for higher customer satisfaction and returns. GBR stands out because it moves from one forecast to the next by assembling and refining multiple weak learners. In the context of menu optimization, GBR is capable of assessing base price, promotion, and the sales record of many elements, such as dish identification number, dish name, description, price, etc., to determine the number of plates per menu item that will generate the highest profitability

(Tao *et al.* 2024). Hence, restaurant portfolios that comprehend which items are likely to provide overall greater margins than other items and contour with clients' preferences can adapt accordingly.

GBR offers enhanced accuracy and the capability of recreating complex patterns, which is useful in estimating the company's profitability of certain menu offerings. The confusion matrix analysis shows that GBR has an acceptable number of true positives; this proves that the items marked as high-profit margin are in fact high, while those marked as low are really low. Thus, when restaurants implement GBR, they focus on items that would make more profits and would fit the customers' tastes, ultimately making more profits across restaurants (Tian *et al.* 2020). Concerning the key learning, it is specifically significant to mention the capacity to make changes to the menu depending on the data obtained and the expectations of the consumers, which contributes to the menu's competitiveness.

5.3 Staff scheduling solution to improve workforce management and productivity using Random Forest regressor

Random Forest Regressor (RFR) is a stable and diverse machine learning model which can enhance the efficacy of the workforce and productivity by providing a restrictive staff scheduling system. RFR works by using historical information on various details for employees, such as shift patterns and various measures of productivity, as well as other data regarding customer requirements and patterns of various other parameters (Wang *et al.* 2022). The ability of RFR to accurately predict future staffing needs was assessed in the context of the RFR public report. The model was able to provide a good accuracy of about 94 per cent for the predicted values. Hence useful in predicting the required number of staff for different time intervals.

This high accuracy leads to an effective stocking of employees and customer flow without having to hire excess or insufficient quantities which often lead to problems such as high costs. After analysing the confusion matrix, it can be observed that the RFR model provides promising results in terms of accuracy, but it is weak in terms of precise classification and recall, which are confirmed by false positive and false negative results (Momade *et al.* 2022). Still, the ability to predict true positives holds a strong impetus suggesting the model can accurately gauge the right staffing numbers.

5.4 Supply chain management enhancement using machine learning application

Value of supply chain management using machine learning applications is a perfect way of taking a fresh approach to supply chain processes and achieving the set goals with efficiency. The use of Logistic Regression would enable one to predict the demand trends, hence would be very helpful to those doing business since they would be in a position to estimate the stock needed (Aamer *et al.* 2020). Since Logistic Regression models take historical sales data into consideration and patterns from the past, there is a possibility of determining the future demand for products with great precision and thus helpful in managing inventories in the best possible way. It decreases the chances of coming across stockouts or overstock situations, which are important so as to ensure the right products are made available at the right time.

Since Random Forest Regressor is a powerful classifier, its application in the given task can bring considerable improvements in workforce management due to better staff scheduling. It helps in determining staffing requirements using historical data, thus avoiding situations where several staff are left idle, or hiring more staff when they are not required, which in turn increases productivity and also minimizes costs which might be incurred on staff (Ni *et al.* 2020). Each of these models can also be used in supply chain management and the four models can be combined in order to create a more complete solution.

5.5 Challenges and solutions

Challenges	**Solutions**
Data quality and availability	● It is needed to implement robust data collection and cleaning processes. ● It is needed to use automated data integration tools to ensure data consistency across various sources, improving the reliability of the data used for machine learning models.
Model interpretability	● Should utilize interpretability techniques like SHAP (Shapley Additive explanations) to explain model outputs and feature

	contributions, enhancing transparency and stakeholder trust, particularly for decision-makers unfamiliar with ML.
Scalability	● It is needed to employ scalable cloud-based platforms such as AWS, Azure, or Google Cloud to handle large datasets and high computational demands, ensuring the models can be efficiently scaled up as data volume grows.
Integration with existing systems	● Needs to develop API-driven integration strategies for seamless communication between ML models and existing systems, ensuring smooth workflow incorporation and minimal disruption to current operations.
Change management	● Needs to conduct training sessions and workshops to demonstrate the value of ML, encourage a data-driven culture, and facilitate smoother adoption. ● It is needed to engage key stakeholders early in the process to align objectives and address concerns proactively.

Table 5.5: Challenges and solutions of the machine learning for predictive analytics and optimization in restaurant operations

Chapter 6: Conclusion and future works

6.1 Conclusion

This research evaluates the positive impact that machine learning could have on the primary areas that define the activity of restaurants and similar businesses, such as demand forecasting, the optimization of the menu, the planning of staff work hours, and supply chain management. Logistic regression has a strong way on the issue of demand forecasting not only does it predict future needs, but it also boosts inventory control and resource utilization. As a machine learning algorithm, gradient boosting regression is helpful for menu optimization when revenue, in combination with consumers' choice of meals, influences menu results.

It may be noted that random forest regression provides a positive impact in achieving efficient organizing of human resource that aids in improving the workforce planning and performance. Machine learning applications help to make supply chain tasks easier and to improve many of the decision-making-related activities. Overcoming these challenges with strong solutions helps prove the successful application and scalability of machine learning and related technologies, contributing to operational improvements and competitiveness in multiple industries.

6.2 Linking with objectives

6.2.1 To develop machine learning models for accurate demand forecasting, enabling better inventory management and resource allocation

Goal of accurate demand forecasting with the help of logistic regression was to classify the future demand into binary categories, which has been successfully reached. Model's accuracy, precision, and recall are considered to be very high, which can confidently speak about the high rate of demand prediction. This helps in efficient decision-making regarding inventory replenishment and controlling excess inventory thus minimizing wastage.

6.2.2 To implement menu optimization strategies using machine learning algorithms, focusing on maximizing profitability and aligning offerings with customer preferences

Goal of menu optimization was achieved through gradient boosting regression by considering customers' data and sales to come up with ideal menu settings. Thus it followed the channel of catering to the needs of the customers thereby increasing profitability and cost control through the selling of popular products.

6.2.3 To design and evaluate staff scheduling solutions that utilize predictive analytics to improve workforce management and productivity

Goal of designing and evaluating the staff scheduling solutions was successfully met with the help of Random Forest regressor by tuning the staff data history and demand in the business. The model was proficient in giving correct staffing forecasts in which schedules could be adjusted depending on the expected demand or availability of people. This helped in improving the workforce planning and controlling since there was adequate staffing while at the same time avoiding overstating since there would be a reduced number of times that there would be many employees with little work to do.

6.2.4 To enhance supply chain management through machine learning applications, improving procurement processes, inventory control, and overall efficiency

Goal to improve supply chain management was realized through the implementation of machine learning algorithms to analyse historical data as well as to improve procurement processes and inventory logistics. The models provided information on the demand levels and inventory requirements that helped improve the efficiency and soundness of the decisions made. Thus, it rationalised the processes and improved the functionality of the supply chain overall.

6.3 Limitations

- Machine learning models are significantly dependent on the data thus the quality and degree of data completeness must be accurate (Bera, 2021). Presence of missing values, noise as well as other forms of data inaccuracies in any data set reduces the accuracy of the model and produces biased results. Both, the quality of the data and missing values need to be dealt with carefully and properly.
- It has been known that building complex structures of models, especially dealing with high dimensional data may badly over-fit the training data in such a way that the resultant model will perform adequately well in recognizing training data but may perform poorly when generalised to other data sets (Kumar *et al.* 2021).
- It’s also worth noting that some machine learning models such as Gradient Boosting and Random Forest sometimes may be rather intricate and are rather complicated for interpretation (Lee *et al.* 2021). Interpretation of how these models arrive at their decisions

is quite complex and thus it becomes challenging to explain model outcomes and even understand the models' decision-making processes.

- If a higher volume of data than what was used for the training needs to be processed or, in the case of real-time data processing, the size of the data streams that need to be analysed, it can be difficult to scale the machine learning models (Wang *et al.* 2022). Readability; making sure that the models are adjusted with growth in data or a continuous integration of data as a major drawback.

6.4 Future scope

There is a wide and positive outlook for implementing machine learning solutions in the subsequent areas of demand forecasting, menu optimization, employee scheduling, and supply chain management. In the near future with the advancements in technology, better algorithms and high-quality large datasets will improve the models to a great extent. One area for future research lies in the integration of different machine learning approaches in order to increase the predictive accuracy and thus decision-making quality (Doborjeh *et al.* 2022). Further, prospects, such as real-time data analyses and edge computing, can mean more adaptive and effective systems to fulfil inventory and other requirements in real time.

Improving the interpretability of the models by the use of advanced techniques and improving the visualization tools will enable the stakeholders to trust the models and their predictions. Also, the link between ML and other advanced technologies including IoT, blockchain and others can enhance further supply chain responsiveness and transparency (Taneja *et al.* 2023). Since systematic computational power is a limited resource, when resources become cheaper and more widely available, scaling up machine learning models suitable for wider applications will become possible, thereby creating new opportunities for development in different fields. Further enhancements in AI and ML will produce smarter systems responding to new issues and fostering strategic development in business.

References

Aamer, A., Eka Yani, L. and Alan Priyatna, I., 2020. Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. *Operations and Supply Chain Management: An International Journal*, *14*(1), pp.1-13.

Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C.J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C.C., Martin, O.A. and Osthege, M., 2023. PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, *9*, p.e1516.

Ali, M., Mustafa, T., Osman, S. and Mehmet, G., 2023. Forecasting Restaurant Sales with the Sensitivity of Weather Conditions and Special Days Using Facebook Prophet. *Journal of Data Applications*, (2), pp.15-30.

Bera, S., 2021. An application of operational analytics: for predicting sales revenue of restaurant. *Machine learning algorithms for industrial applications*, pp.209-235.

Borowski, M. and Zwolińska, K., 2020. Prediction of cooling energy consumption in hotel building using machine learning techniques. *Energies*, *13*(23), p.6226.

Chu, H., Zhang, W., Bai, P. and Chen, Y., 2023. Data-driven optimization for last-mile delivery. *Complex & Intelligent Systems*, *9*(3), pp.2271-2284.

Deng, C. and Liu, Y., 2021. A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. *Wireless Communications and Mobile Computing*, *2021*(1), p.9969357.

Doborjeh, Z., Hemmington, N., Doborjeh, M. and Kasabov, N., 2022. Artificial intelligence: a systematic review of methods and applications in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *34*(3), pp.1154-1176.

Feng, D., Mo, Y., Tang, Z., Chen, Q., Zhang, H., Akerkar, R. and Song, X., 2021. Data-driven hospital personnel scheduling optimization through patients prediction. *CCF Transactions on Pervasive Computing and Interaction*, *3*, pp.40-56.

Harshini, K., Madhira, P.K., Chaitra, S. and Reddy, G.P., 2021, September. Enhanced Demand Forecasting System For Food and Raw Materials Using Ensemble Learning. In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)* (pp. 1-6). IEEE.

Jiang, Y., 2020, December. Restaurant reviews analysis model based on machine learning algorithms. In *2020 Management Science Informatization and Economic Innovation Development Conference (MSIEID)* (pp. 169-178). IEEE.

Kannan, R., Rosdi, I.S., Ramakrishnan, K., Rasid, H.R.A., Rafy, M.H.I.M., Yusuf, S. and Salamun, S.N.A.M., 2021. Leveraging business data analytics and machine learning techniques for competitive advantage: case study evidence from small businesses. *International Journal of Management, Finance and Accounting*, *2*(1), pp.73-87.

Khan, P.W., Byun, Y.C. and Park, N., 2020. IoT-blockchain enabled optimized provenance system for food industry 4.0 using advanced deep learning. *Sensors*, *20*(10), p.2990.

Kmiecik, M., 2022. Logistics coordination based on inventory management and transportation planning by third-party logistics (3PL). *Sustainability*, *14*(13), p.8134.

Kumar, I., Rawat, J., Mohd, N. and Husain, S., 2021. Opportunities of artificial intelligence and machine learning in the food industry. *Journal of Food Quality*, *2021*(1), p.4535567.

Lee, M., Kwon, W. and Back, K.J., 2021. Artificial intelligence for hospitality big data analytics: developing a prediction model of restaurant review helpfulness for customer decision-making. *International Journal of Contemporary Hospitality Management*, *33*(6), pp.2117-2136.

Ma, B., Guo, W. and Zhang, J., 2020. A survey of online data-driven proactive 5G network optimisation using machine learning. *IEEE access*, *8*, pp.35606-35637.

Mehmood, U., Broderick, J., Davies, S., Bashir, A.K. and Rabie, K., 2024. Machine Learning-based Predictive Inventory for a Vending Machine Warehouse.

Mirmozaffari, M., Yazdani, M., Boskabadi, A., Ahady Dolatsara, H., Kabirifar, K. and Amiri Golilarz, N., 2020. A novel machine learning approach combined with optimization models for eco-efficiency evaluation. *Applied Sciences*, *10*(15), p.5210.

Momade, M.H., Shahid, S., Hainin, M.R.B., Nashwan, M.S. and Tahir Umar, A., 2022. Modelling labour productivity using SVM and RF: a comparative study on classifiers performance. *International Journal of Construction Management*, *22*(10), pp.1924-1934.

Ni, D., Xiao, Z. and Lim, M.K., 2020. A systematic review of the research trends of machine learning in supply chain management. *International Journal of Machine Learning and Cybernetics*, *11*, pp.1463-1482.

Oh, M. and Kim, S., 2022. Role of emotions in fine dining restaurant online reviews: the applications of semantic network analysis and a machine learning algorithm. *International Journal of Hospitality & Tourism Administration*, *23*(5), pp.875-903.

Oriekhoe, O.I., Oyeyemi, O.P., Bello, B.G., Omotoye, G.B., Daraojimba, A.I. and Adefemi, A., 2024. Blockchain in supply chain management: A review of efficiency, transparency, and innovation. *International Journal of Science and Research Archive*, *11*(1), pp.173-181.

Puh, K. and Bagić Babac, M., 2023. Predicting sentiment and rating of tourist reviews using machine learning. *Journal of hospitality and tourism insights*, *6*(3), pp.1188-1204.

Rahmany, M., Zin, A.M. and Sundararajan, E.A., 2020. Comparing tools provided by python and r for exploratory data analysis. *IJISCS (International Journal of Information System and Computer Science)*, *4*(3), pp.131-142.

Rais, D. and Xuezhi, Z., 2024. Elevating Student Engagement and Academic Performance: A Quantitative Analysis of Python Programming Integration in the" Merdeka Belajar" Curriculum. *Journal on Mathematics Education*, *15*(2), pp.495-516.

Shi, P., Helm, J.E., Chen, C., Lim, J., Parker, R.P., Tinsley, T. and Cecil, J., 2023. Operations (management) warp speed: Rapid deployment of hospital-focused predictive/prescriptive analytics for the COVID-19 pandemic. *Production and Operations Management*, *32*(5), pp.1433-1452.

Spiliotis, E., Makridakis, S., Semenoglou, A.A. and Assimakopoulos, V., 2022. Comparison of statistical and machine learning methods for daily SKU demand forecasting. *Operational Research*, *22*(3), pp.3037-3061.

Taneja, A., Nair, G., Joshi, M., Sharma, S., Sharma, S., Jambrak, A.R., Roselló-Soto, E., Barba, F.J., Castagnini, J.M., Leksawasdi, N. and Phimolsiripol, Y., 2023. Artificial intelligence: Implications for the agri-food sector. *Agronomy*, *13*(5), p.1397.

Tao, W., Wu, C., Wu, T. and Chen, F., 2024. Research on the Optimization of Pricing and the Replenishment Decision-Making Problem Based on LightGBM and Dynamic Programming. *Axioms*, *13*(4), p.257.

Tian, L. and Feinberg, F.M., 2020. Optimizing price menus for duration discounts: A subscription selectivity field experiment. *Marketing Science*, *39*(6), pp.1181-1198.

Wang, L., Gopal, R., Shankar, R. and Pancras, J., 2022. Forecasting venue popularity on location-based services using interpretable machine learning. *Production and Operations Management*, *31*(7), pp.2773-2788.

Wang, Z. and Dexter, F., 2022. More accurate, unbiased predictions of operating room times increase labor productivity with the same staff scheduling provided allocated hours are increased. *Perioperative Care and Operating Room Management*, *29*, p.100286.

Yerragudipadu, S., Gurram, V.R., Rayapudi, N.S., Bingi, B. and Gollapalli, L., 2023. An Efficient Novel Approach on Machine Learning Paradigmsfor Food Delivery Company through Demand Forecastıng in societal community. In *E3S Web of Conferences* (Vol. 391, p. 01089). EDP Sciences.