Food Demand Prediction using Statistical and Machine Learning Models

MSc Research Project
Data Analytics

Sasikumar Jayapal
Student ID: x21153272

School of Computing
National College of Ireland

Supervisor: Cristina Hava Muntean

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Sasikumar Jayapal
Student ID:	x21153272
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Cristina Hava Muntean
Submission Due Date:	15/12/2022
Project Title:	Food Demand Prediction using Statistical and Machine Learning Models
Word Count:	8730
Page Count:	25

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:	Sasikumar Jayapal
Date:	29th January 2023

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):	

Food Demand Prediction using Statistical and Machine Learning Models

Sasikumar Jayapal
x21153272

Abstract

The demand for food is rising faster than the national economy due to population growth, climate change, and digitalization. The food-based industries like restaurants, canteens, catering services, fast food centers, etc., are the ones that handle perishable materials more often. One of the biggest challenges for the food-based industries to managing food orders for their customers. Sometimes, Inaccurate estimation of the food orders can lead to excessive or insufficient food can result in waste of both food and raw materials, as well as ineffective employee management and decreased business profit. This proposed research investigates the appropriate data mining, statistical, and machine learning models to predict the accurate food orders for a restaurant for the upcoming weeks. The machine learning models lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and gradient boosting regression models like Gradient Boosting, XGBoosting, LightGBM, CatBoost, and Facebook prophet can be applied to a large dataset with roughly 155 weeks' worth of food orders gathered from a restaurant in Italy. As a result, food-based industries like restaurants, canteens, and fast food centers can minimize their operation costs by reducing food and raw material waste and improving customer satisfaction by serving fresh and delicious food.

1 Introduction

Food-based businesses like restaurants, canteens, fast food centers, etc. are growing every day in today's fast-paced world because people's food preferences are constantly changing and they don't have enough time to prepare and eat their meals at home. Numerous factors, including population growth, climatic changes, and income convergence, are influencing the global food demand (Fukase and Martin; 2020). According to this research paper (van Dijk et al.; 2021), the global food demand is expected to increase from 35% to 56% by 2050, and if climate change is taken into account, it may rise from 30% to 60% by 2050. The Food and Agriculture Organization of the United Nations estimates that up to one-third of food produced globally, amounting to 1,3 billion tons of food per year, is lost and wasted (FAO; 2019). Food demand prediction is crucial for all food-based industries for both business operations and sustainable development. The sustainable development concerns are centered on food waste and food loss, while the business process includes manufacturing, logistics, supply chain process, inventory management, and customer satisfaction.

The food-based industries are the ones that often handle perishable materials which can spoil quickly. The ability to process the appropriate number of orders for their customers is a constant challenge for the food and beverage manufacturing industries. Inaccurate estimation of the food orders can lead to excessive or insufficient food can result in waste of both food and raw materials, as well as ineffective employee management and decreased business profit. Therefore, it has become a significant challenge for the food-based industries to forecast the appropriate volume of food orders for the upcoming weeks.

The goal of the research is to develop a demand forecasting system for predicting food demand in order to forecast food orders for upcoming weeks. This system will increase profit by decreasing food and raw material waste, providing customers with fresh and delicious food on time, and managing employees effectively. The machine learning models lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and gradian boosting regression models like Gradient Boosting, XGBoosting, LightGBM, CatBoost, and Facebook prophet can be used to build an accurate demand forecasting system on the large restaurant dataset. Regression, time series, or a combination of both can be used to approach this research problem. The following research questions can be addressed through this research.

1. How Can statistical time series and machine learning models be used to forecast food orders for the food-based industries for the upcoming weeks?

2. What degree of predictability can be achieved with the proposed demand prediction system for predicting food orders on a weekly basis for a restaurant using time series and machine learning models?

The first objective was to conduct a critical analysis of the field of food-based industries, food demand prediction, and its related works from 2010 to 2022. The following goal is to apply various statistical time series and machine learning to identify the most accurate model and evaluate the effectiveness of each. This research will be extremely beneficial to food-based industries, such as restaurants, canteens, fast-food outlets, etc. With the help of this research, business profits can be raised by cutting down on the waste of food and raw materials, managing inventories and employees effectively, and increasing customer satisfaction by delivering high-quality meals on time.

We assume that the number of orders placed at each store is independent of each other. The point of sale (POS) system has recorded the accumulated order quantity for the entire week for each store and meal, and it is accessible for research. Each row in the dataset is specific to the store and meal id. This research is limited to the weekly data and there are no external factors such as weekends, holidays, weather, and special events are not taken into account.

The paper is structured as shown below. The research environment and research objective are described in section 1. We examine the related literature on the food-based industries in Section 2. Section 3 describes the research methodology and its phases. The design specification for the study is presented in section 4. Section 5 includes the specific implementation steps for each forecasting model. Section 6 presents the evaluation's findings, case studies, results, and discussions. In Section 7 we present the conclusions and discuss several options for future research. Finally, the research ended up with a list of references.

2 Related Work

The state of the art for demand forecasting in the food industry over the previous twelve years, from 2010 to 2022, across three different domains has been discussed below in this section. The literatures that are grouped into the following three subsections below are critically reviewed and analyzed with regard to the dataset, methodology, techniques, and evaluation metrics to support the novelty of this research. Also, It helps to understand the background knowledge with respect to this study.

2.1 Review of demand forecasting using classical machine learning methods

In order to manage and operate a successful food-based business, demand forecasting is crucial. Predicting the number of customers who will enter a store is one method of demand forecasting. The research paper (Tanizaki et al.; 2019) was proposed to predict the number of customers for face-to-face service industries. To construct an accurate demand forecasting model, they use both internal point of sales (POS) data and external store-specific data, such as store location, weather, events, etc. The machine learning models Boosted Decision Tree Algorithm, Decision Forest Regression, and Bayesian Linear Regression methods are utilized. the stepwise method is used for statistical analysis, and Azure Machine Learning and SPSS tools were utilized for the implementation. The forecasting accuracy for the Bayesian, Decision, and stepwise analysis methods did not significantly differ, whereas the boosting regression method's forecasting accuracy was only moderately low. The Bayesian, Decision, and stepwise methods produce forecasts with an accuracy of up to 85%. The future work includes focusing on enhancing forecasting accuracy and investigating effective store management strategies, such as automated food material ordering and flexible work schedules for employees. There is no evidence that this research paper adheres to any particular methodology, and the dataset's feature variables and data volume are not clearly specified.

The future work of (Tanizaki et al.; 2019) has been carried out in the literature (Tanizaki et al.; 2020), the research illustrates how to manage the restaurants by having an accurate demand forecasting system. The researcher has used a random forest regression machine learning algorithm to predict customer order quantity and inventory order quantity of beers in restaurants at each store. They considered both internal point of sales (POS) data and external factors such as store location, weather, events, etc. The forecasting ratio for customer ordering quantity ranges between 22% to 68% and the accuracy is not significant while the forecasting ratio for inventory ordering quantity ranges between 45% to 71% and the accuracy is not commendable. Here, the model is trained with one year of data, this research can be further extended to increase the number of years of training data and examine the effect on the forecasting ratio and fitting ratio.

This (Donselaar et al.; 2016) study's goal is to analyze how promotional discounts affect the sale of perishable goods and to help retailers by developing a forecasting model to predict demand for perishable goods during promotions. For the investigation, data from a Dutch supermarket retail chain were gathered. There are 407 distinct perishable goods in 4 different product categories, including desserts, dairy drinks, cold cuts, and salads. This analysis doesn't offer any strong evidence for the presence of threshold and saturation levels for these categories. Instead, it uses regression models with five dummy variables for various pricing classes to capture potential threshold and saturation effects. The model comparison between linear and quadratic functions produced the worst forecasting accuracy when simulating saturation and threshold effects. When a product category uses routine and non-routine processes, and when a product has few or many observations, the forecasting accuracy is significantly increased. The researcher concludes that the study should be expanded in order to assess how generalizable the findings regarding saturation level and threshold are.

The livestock and agricultural industries also heavily rely on food price forecasts to boost profits and lower operational risk factors. This study (Wibowo and Yasmina; 2021) employs cross-validation and damping factor techniques to forecast food prices using linear regression and ridge regression. Due to their widespread usage in the market, the research is primarily focused on the price prediction of rice and eggs. The linear regression and ridge regression are implemented on the both rice and eggs datasets. On the rice dataset, the RMSE rates for linear and ridge regression with cross-validation are 0.7 and 0.6, respectively, and the error rates for both models are further reduced to 0.4 and 0.4 when the models are trained using damping factors. On the egg dataset, the RMSE rates for the models with cross-validation are 0.5 and 0.5, and the rates for the models with damping factors are 0.03 and 0.03. Thus the damping factor improves both the linear and ridge regressor accuracy.

2.2 Review of demand forecasting using artificial neural network methods

In general, rising food demand has an impact on the nation's economy. By using three decision tree methods—Classification and Regression Tree (CART), Chi-Squared Automatic Interaction Detection (CHAID), and Microsoft Decision Trees—this study (Bozkir and Sezer; 2011) forecasts actual food consumption demand for a certain menu on a selected date. For the analysis, two years' worth of data were gathered from the food courts at Hacctepe University in Turkey. Building a data mining model for predicting food demand is one of the two major objectives of this study, and the other is to pinpoint the variables that influence consumer behavior among various customers kinds. In which the MSDT model makes the better forecast with an accuracy of 80% while the CHAID model performed with a superior accuracy of 83%. The research will be expanded in the future to incorporate models of artificial neural networks (ANN) and support vector machines (SVM) to create a decision support system based on a more effective algorithm. Furthermore, whereas the study was conducted with data from a single year, the authors will train the model with data spanning three years.

Food quality and food safety are also other important facts that we think of in the food industry. Food quality will lead to higher food consumption and higher profits. This research (Anly Antony and Kumar; 2021) offers a systematic data-driven approach to predicting food quality based on a number of features, including size, color, shape, texture, and defects in the food products used for quality analysis. This study incorporates a variety of image processing and machine learning approaches, and it includes four steps: image acquisition, pre-processing, segmentation, feature extraction, and classification. This research work will be incorporated to use deep learning techniques such as convolutional neural networks. A single category of fruits or vegetables can be sorted and graded using the approach that has been devised. Future work can be incorporated with counting, sorting, grading, and defect identification in multiple categories of fruits and vegetables. Table 1 provides information on the model accuracy for the food goods.

Table 1: Food Product Accuracy

Food Products	Techniques	Accuracy
Tomato	Image acquisition, Segmentation	83%
Mango	SVM classifier	87%
Guava	Nearest-neighbour classification	90%
Corn Seed	K-means clustering	85%
Oil Palm	Multi-attribute decision making theory	Error-2.4%
Meat	mass spectrometry method	81.5%
Milk	deep and ensemble decision tree	98.76%

This research article (Pallathadka et al.; 2022) provides a detailed review of various artificial intelligence and machine learning techniques being utilized in the food and agricultural sectors mainly on supply chain optimization, crop selection, logistics, and food delivery, etc., The authors have contrasted the findings, methodologies, and approaches from various fields and datasets. Additionally, the research work done by (Kakani et al.; 2020) provides a significant view of how computer vision and artificial intelligence are used in the food and agricultural industry. Several agricultural applications, including food processing, agriculture-based applications, farming, plant data analysis, smart irrigation, and next-generation farming, are handled by computer vision and artificial intelligence methodologies. The paper also emphasizes the fundamental idea of using sustainable 4 IR technologies to help humanity achieve the required food supply by 2050 in an environmentally friendly way.

The study conducted by (Çetinkaya and Erdal; 2019) emphasizes demand forecasting in food production planning in a cafeteria and examines the variables influencing meal demand. The dataset is being gathered from a cafeteria at Kirikkale University and is divided into student and staff datasets. The artificial neural network models were applied to both student and staff datasets to predict food orders placed at the cafeteria on a specific set of meal items. RMSE and MSE were used to evaluate the artificial neural network models. As a result, the average MSE for the student and staff datasets is 0.00864 and 0.007014, respectively, and the average RMSE is reported as 0.09295 and 0.083748, respectively.

2.3 Review of demand forecasting using statistical methods

Predicting future sales for the food-based industries is also another way of forecasting food demand. A review of (Posch et al.; 2022) illustrates the demand prediction system by predicting future sales of daily sold quantity for the restaurants and canteens. The point of sale data is collected from the respective restaurants or canteens and considering the data has strong seasonality and trend changes. They considered two datasets consisting of multiple time series collected at a restaurant and a larger canteen including 20 months of data. There are two Bayesian general additive models were used for the prediction. The performance of both models was evaluated and compared with each other. The sales forecasting sales with the negative binomial distribution has more accurate than the prediction with the normal distribution. Currently, they used the POS data. In the future, the authors plan to expand on this work by adding holidays, special events, and weather data to the dataset and forecasting sales on an hourly basis.

The establishment of food sale prices for food goods is a significant responsibility for food-based industries. The analysis and forecasting of food prices are crucial, and the examination of seasonality and trends seen in time series data is made possible by this forecast. The Italian food wholesale corporation uses a variety of approaches to predict sale prices, including Facebook Prophet, Convolutional Neural Networks, Long Short-Term Memory (LSTM), and Auto-Regressive Integrated Moving Average (ARIMA). The LSTM model has an RMSE of 0.6, compared to 0.07 for ARIMA (3,1,1). When CNN and LSTM are combined, the RMSE is 0.05. With an RMSE of 0.07, the prophet model performs poorly. The results show that the ARIMA and LSTM models performed similarly, whereas CNN and LSTM combinations achieve better overall accuracy, but tuning the model takes more effort. Facebook Prophet, on the other hand, is a quick and simple to use but much less accurate model (Menculini et al.; 2021).

In this article (Fattah et al.; 2018), demand forecasting for food products used in food production is carried out using historical demand data for Moroccan food manufacturing industries from 2010 to 2015. Autoregressive integrated moving average (ARIMA) models were used by this author by applying the Box-Jenkins method. In this study, several ARIMA models with various parameter settings were built, and the best model, ARIMA (1,0,1), was chosen based on performance indicators like SBC, AIC, standard error, and maximum likelihood. The results show that the chosen model provides better forecasting accuracy of 83% in the food manufacturing sector. Future work will involve creating additional models using quantitative and qualitative methods in order to create reliable forecasting models and improve forecast accuracy.

Predicting food prices in the food-based industries is crucial for the business as there are many factors that affect food prices, such as types of food, weather, seasons, etc. The research paper (Kim et al.; 2018) outlines the food price trend analysis between healthy and unhealthy foods in South Korea. This research employs the time series analysis for 20 years of data from 1995 to 2015 on selected food items such as grains, vegetables, meats, sweets, spices, fast foods, and non-alcoholic beverages. Hence the result shows that healthy food prices are relative increases over the period when compared with unhealthy food prices.

The proposed research work by Vithitsoontorn and Chongstitvatana (2022) demonstrates demand forecasting in dairy products using a multistep forecasting approach. For this analysis, the authors have used 8 dairy products. The dairy transaction sale data were collected from the ERP system of the Dairy Farming Promotion Organization of Thailand (DPO) between October 2016 and September 2021 (5 financial years), as reported by 5 plants. The Bank of Thailand is where the holiday data is gathered. Models from statistics and neural networks, like ARIMA and LSTM, were used in this study. The models are contrasted with weekly and monthly observations. The findings demonstrate that both statistical and neural network models are trustworthy and can be used for demand forecasting. The results are nicely interpreted, and the structure and methodology used are presented clearly.

2.4 Conclusion

The limitations of existing models in the previous works and the need for better models and the current works with respect to previous work limitations are listed in the following Table 2.

Table 2: Related Works and Its Limitations

Author and Area	Previous Works Limitations	Current Research Work
(Tanizaki et al.; 2019), Demand forecasting based on a number of customers.	There is no big difference in forecasting rate while using the Bayesian, boosted decision tree, and decision forest regression algorithms used by the authors. In actuality, the boosting algorithm's accuracy is a little low to others. Future works of this paper are to increase its accuracy and effective management including ordering food supplies and managing the workforce.	Since the accuracy of the boosting model is not significant, in comparison. The researcher employs boosting algorithms, including gradient boosting, XGBoost, Light GBM, and catboost algorithms, to increase the forecasting rate. Among these boosting models, the XGBoost model outperformed the previous research in this study.
(Tanizaki et al.; 2020), Forecasting customer order quantity of beers.	According to this study, the model's accuracy was not good for both customer and inventory ordering amounts that ranged from 22% to 68% and 45% to 71%, respectively. Future work on this paper is to expand the dataset size in order to assess the models' forecasting rate.	The models are trained using a sizable dataset containing 423k observations. All regression models have good average accuracy, which is over 80%. Hence the model's forecasting rate is good while predicting a number of food orders in a restaurant.
(Wibowo and Yasmina; 2021), Forecasts actual food consumption demand	This study uses linear and ridge regression models, but the author believes that using nonlinear forecasting models with a large data set would provide better results and computation	In order to assess the forecasting rates on the demand forecasting, many linear regression models, including lasso and ridge, multiple linear, and Bayesian ridge regression, as well as non-linear regression models, such as SVR, Decision Tree, Random Forest, and Boosting models, are utilized on a large dataset.
(Bozkir and Sezer; 2011), Forecasts actual food consumption demand	The authors intend to use ANN and SVM techniques in the future. While this study used data from a single year, the authors plan to build models with a dataset of three years in order to achieve accurate results.	The researcher tried to build SVR model, but the model takes more time about more than 24 hours to get trained. However, the hyperparameters are not improved the model's training time and the model was trained with 3 years of data including 423k observations.

(Posch et al.; 2022) illustrates the demand prediction system by predicting future sales of daily sold quantity for the restaurants and canteens	Since only point of sale data were used in this study, the researcher believes that enriching the data with information on the weather, special events, and holidays will improve forecasting accuracy. Planning for the workforce will be possible if this method is extended to predict on an hourly basis.	The Facebook prophet model was built to accommodate the details about the weather, special events, and holidays but the dataset contains weekly observations, so I am unable to provide additional information while building the model. It would be better suited if it's a daily observation.
(Menculini et al.; 2021), Predictions of food sale prices for food goods	To forecast the prices of foods being sold, this study used time series, neural network, and deep learning models. Facebook Prophet's accuracy is poor when compared to other models, according to comparisons.	The researcher attempted to build Facebook prophet model on a weekly observation dataset. However, the accuracy of the model is comparatively low.
(Fattah et al.; 2018), demand forecasting for food products in a food manufacturing industries	The ARIMA time series model is used in this study. The author considers creating additional models in the future in order to increase their precision and forecasting rate.	In order to assess the accuracy and forecasting rate, The researcher has built several statistical and machine learning models.

The model's training times are also captured to assess its performance with respect to computation which is not done in any of the previous works. This section 2 presents a number of previous works on the demand prediction on food-based products that have been published over a decade. These related works are divided into three categories and presented as such. by means of which an investigation based on food sales, prices, revenue, orders, and customers is conducted. The researcher has critically analyzed the prior works concerning the dataset, research methodology, techniques, feature scaling, hyperparameter tuning, evaluation metrics, and future works of each of the publications to ascertain the novelty of this research work.

According to the research made in the previous works section, there are numerous articles published for predicting food demand by food price, sales, and customers. , Hence the researcher considers that predicting food demand by means of food orders placed at a restaurant is worth investigating and implementing further including the future works of the reviewed literature works. With a large amount of data, this study introduces new work in terms of algorithms, such as multiple linear regression, lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and Gradient boosting regression models like Gradient Boosting, XGBoosting, LightGBM, CatBoost, and Facebook Prophet models.

3 Methodology

In order to extract useful information and patterns from a raw dataset, the researcher in this study used KDD (Knowledge Discovery from Databases) methodology. Fig.1 illustrates the stages of this method, which include goal setting, data selection, data cleaning and transformation, modeling, and evaluation.

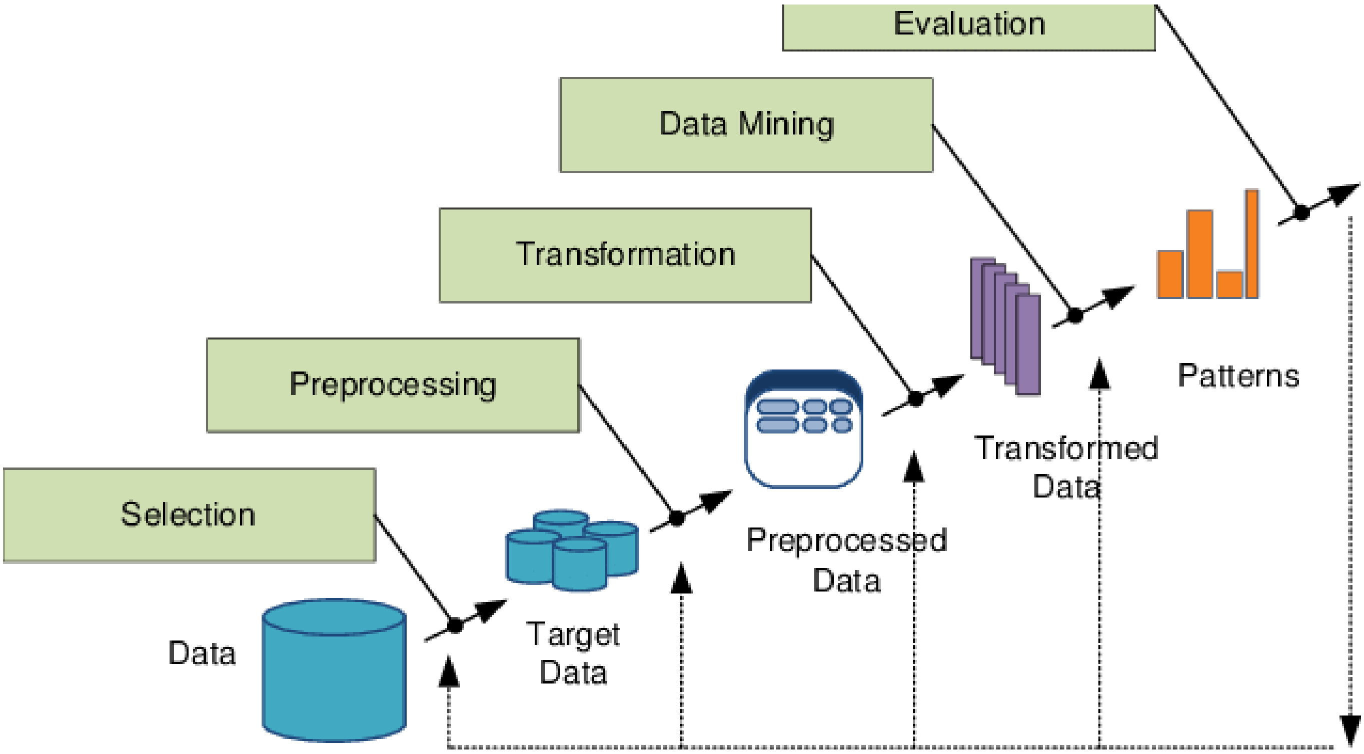

Figure 1: KDD Methodology (Costagliola et al.; 2009)

3.1 Data Selection

To address the research hypothesis, The dataset was chosen from the open platform Kaggle [1]. The information was gathered between 2017 and 2019 from a restaurant in Italy. Each row in the dataset corresponds to an aggregated food order made online over a week with respect to a specific meal. The actual dataset is a combination of three different datasets: the first dataset contains information about aggregated weekly food orders; the second, the details of the fulfillment centers; and the third, the information on each meal. The dataset has 423727 observations and 18 feature variables. The variable num_orders is the target variable and the other variables are considered predictor variables as listed in Table 4, which illustrates some of the most significant feature variables.

3.2 Exploratory Data Analysis

Exploratory Data Analysis refers to a set of procedures for producing descriptive and graphical representations of the data which helps to understand, summarize, visualize, and become familiar with the important characteristics of the data. The trends in food

[1]https://www.kaggle.com/datasets/kannanaikkal/food-demand-forecasting

Table 4: **Feature Variables**

Feature Variables	Description
week	Week number in which the order is being placed
center_id	Unique center id for each fulfillment center
meal_id	Unique meal id for each meal
checkout_price	Final price including discount, taxes and delivery charges
base_price	Base price of the meal
num_orders	(Target) Number of food orders
center_type	Center type (TYPE A, TYPE B or TYPE C)
region_code	Unique code for region
city_code	Unique code for each city
category	Type of meal (beverages/snacks/soups....)
cuisine	Meal cuisine (Indian/Italian/...)
emailer_for_promotion	Emailer sent for promotion of meal
homepage_featured	Meal featured at homepage

orders from 2017 to 2019 are shown in figure 2. A high volume of food orders was placed at the restaurant in Jan 2017, and after a slow decline until March, food orders increased once more as a result of the Italian Easter Festival in April. Italy experiences hot weather in July, and festivals and other special events are in full swing. As a result, food orders are high during this month. Food orders are significantly higher in October due to special occasions and the end of the summer season. The trends for food orders in 2018 and 2019 are similar to those from 2017. However, there is no explanation for the sharp decline in food orders in October 2019.

The figure 3 illustrates the percentage of orders for each food category in the restaurant. where beverage orders are significant, accounting for 35% of total orders. Then, a rice bowl and a sandwich made up 17% and 15% of the total. In Italy, briyani orders for Indian food have decreased.

There are four types of cuisine that are particularly well-liked in Italy as shown in figure 4. Italian food orders make up about 37% of all orders, while orders for Thai, Indian, and continental cuisine are placed at about 27%, 21%, and 24%, respectively. Figure 5 shows that restaurants in the Apulia region of Italy received more orders overall by about 51% as a result of the dense population.

3.3 Data Cleaning and Preprocessing

Sometimes the performance of the model can be significantly diminished by the noisy data. So, at this point, preprocessing and data cleaning of the dataset are crucial. The research involves the following cleaning and preprocessing steps. The dataset does not contain any missing values or values with nulls or NA. The figure 6 indicates the presence of outliers in the target variable Num_Orders. The outliers were handled by removing them by applying the filter on the num_orders greater than 15000 from the dataset. In this section, there are some new feature variables including discount amount, discount percentage, discount y/n, and compare week price are derived from existing feature variables that make significant changes in the model performance.

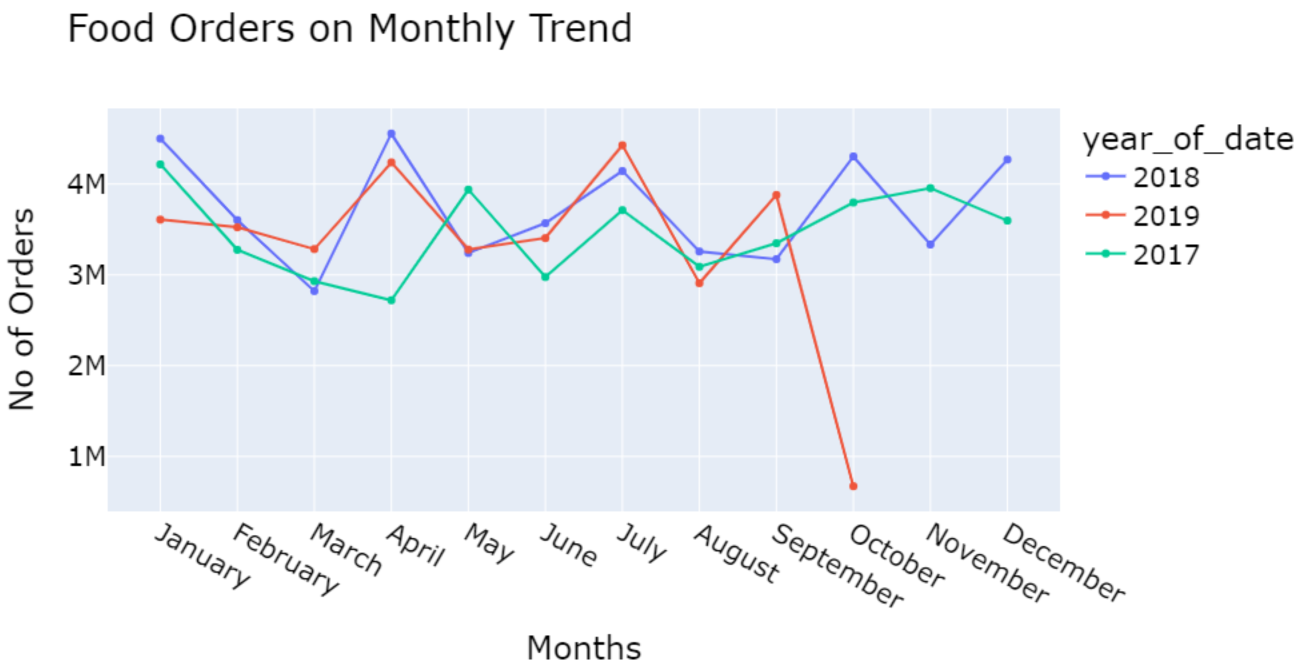

Figure 2: Food Orders Trend

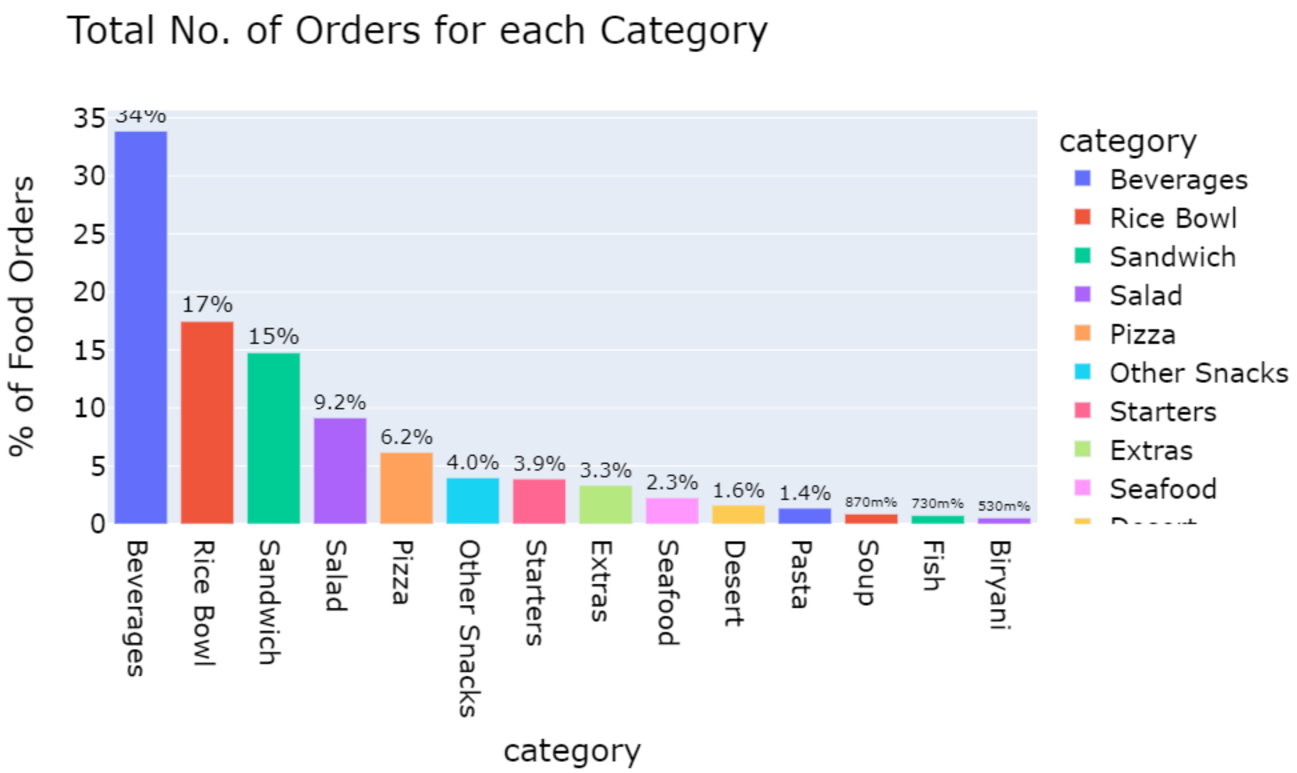

Figure 3: Food Orders by each Food Category

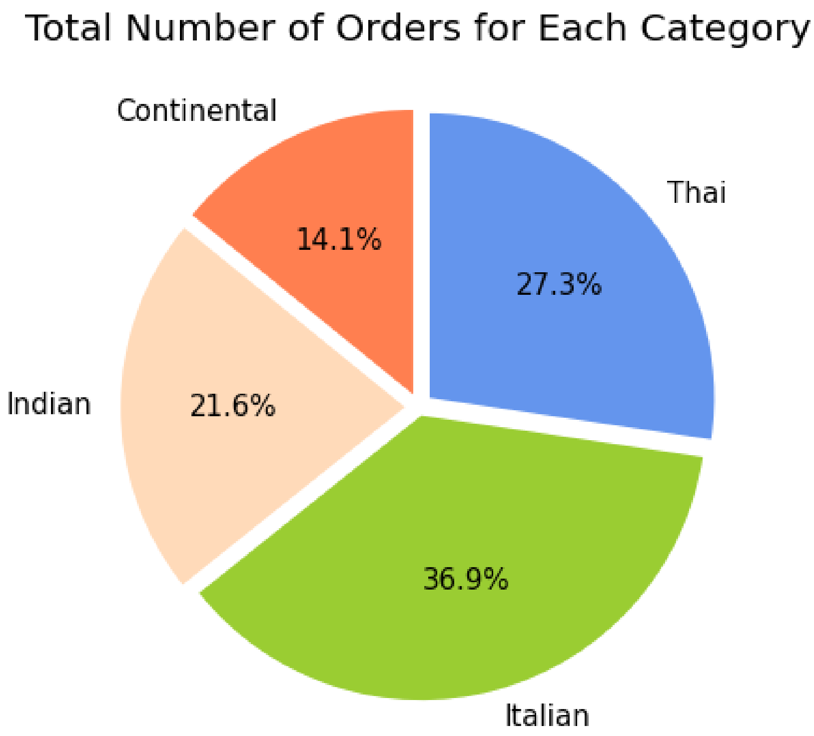

Figure 4: Food orders by Cuisines

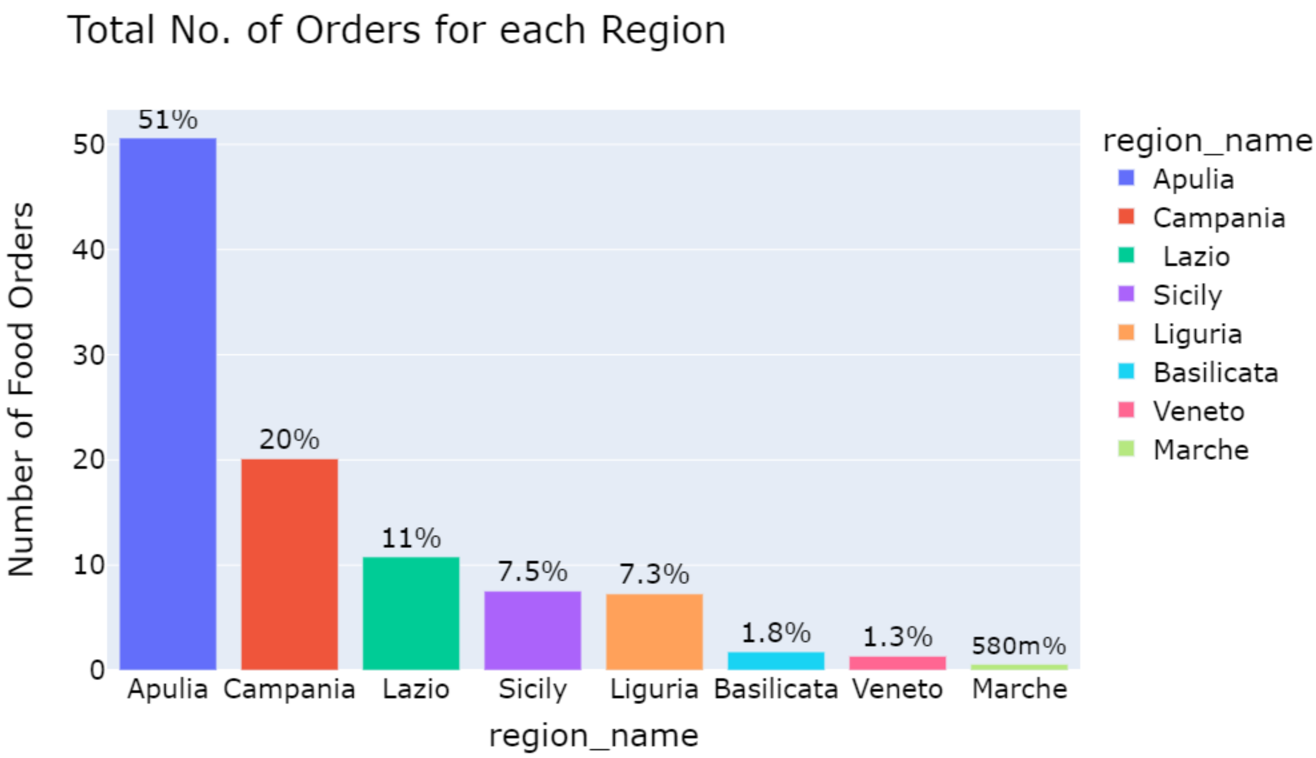

Figure 5: Food orders by Region

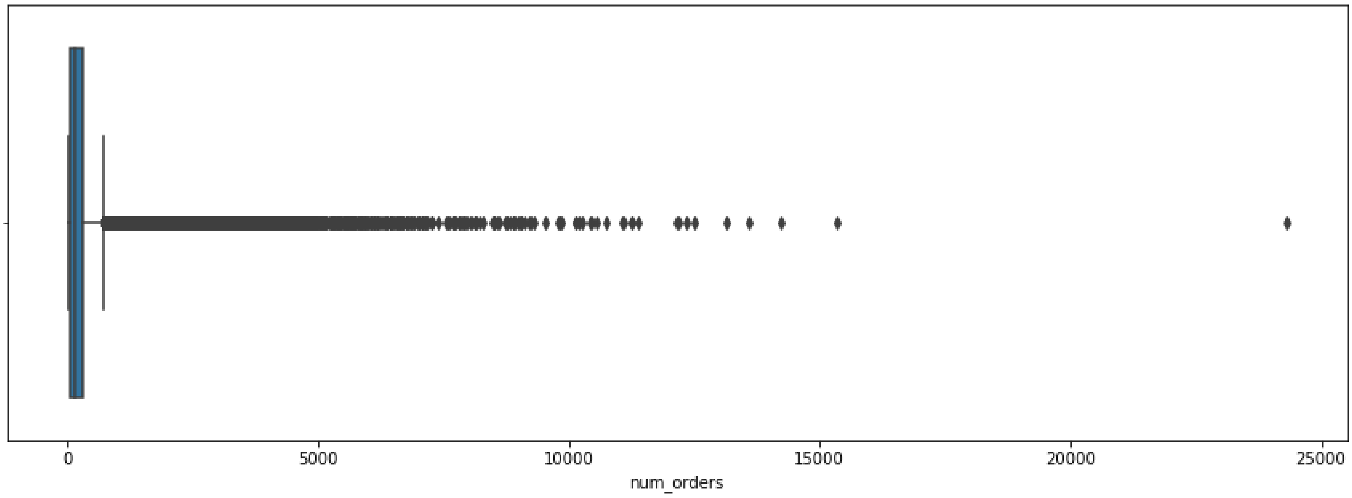

Figure 6: Outliers

3.4 Transformation

Using labelencoder, categorical variables such as center type, op area, cuisines, and category are converted into numerical variables at this stage. The log transformation is applied to the target variable to normalize the data because the target variable, num_order, is not normally distributed as shown in the figure 7. The scale-up of the feature variables is accomplished using the minmax scalar function. the highly correlated and less feature-scoring variables such as id, week, date, checkout price, month_of_date, year_of_date, city name, and region name are dropped from the original data frame. The splits for training and testing sets are based on 80% and 20%, and 70% and 30%.

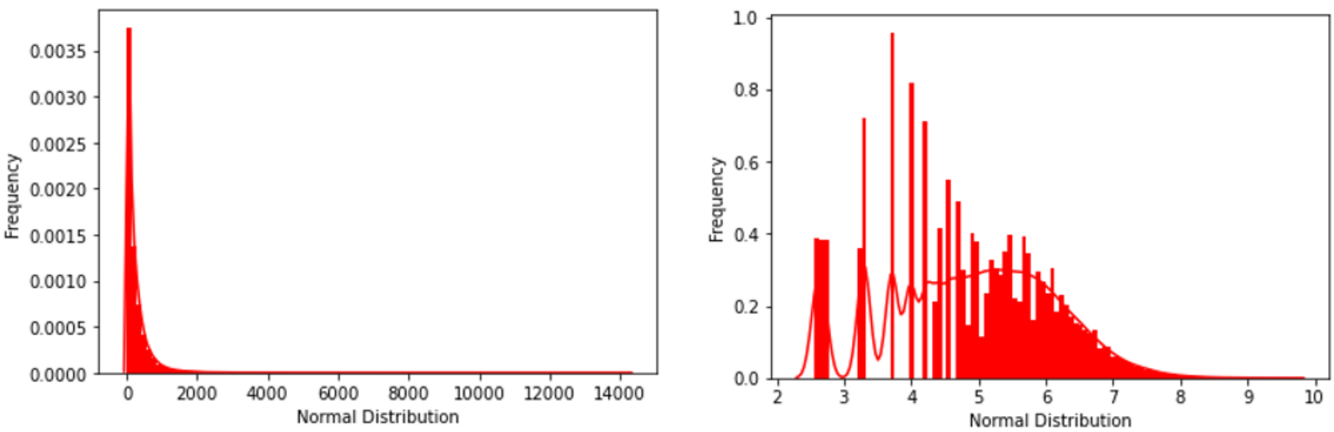

Figure 7: Distribution of Target Variable

3.5 Model Building

In order to predict the food orders for the future weeks, the following machine learning, statistical time series, and neural network models are built with better parameters identified in the hyperparameter tuning section.

- **Multiple Linear Regression**: A statistical technique known as multiple regressors is multiple linear regression. This method forecasts the outcome of the target variable using a number of feature variables. This model seeks to identify the linear relationship between the independent and dependent variables. The model that fits the data best is one that has a low residual error rate between the actual and predicted observation. This model can be assessed using R^2 and RMSE rate.

- **Support Vector Regressor (SVR)**: SVR is One of the commonly used regression algorithms, it is a component of the support vector machine and can be used to solve regression problems. The SVR algorithm will create an insensitive tube with a single best-fitting line or hyperplane. The SVR gives us the flexibility to allow some error margin while drawing the best filling line, i.e., an error inside the insensitive tube can be acceptable and the error outside of the tube can be taken into account for the evaluation.

- **Decision Tree Regression**: It is a supervised, tree-based machine learning technique that is also used to predict target variables in a nonlinear manner and is appropriate for continuous target variables. We can make decisions in a tree structure by using a decision tree. By choosing a feature variable that has a better split based on entropy and information gain, the root node is chosen. The splitting process is continued until all leaf nodes are pure and entropy is reduced.

- **Random Forest Regression**: The random forest algorithm builds multiple randomly selected decision trees from the data using the ensembling and boosting techniques, combining them with a decision tree framework, and then averaging the results to produce a new result that frequently produces accurate predictions and classifications.

- **Gradient Boosting Regression**: On the decision tree, the gradient boosting algorithm applies ensemble and boosting techniques. The algorithm that combines weak decisions and models is continually tuned using the residuals of earlier weak models to produce better predictions. To prevent high variance and bias, the model uses shrinkage regularization. The random forest algorithm cannot compare to this model.

- **XGBoost Regression**: The XG Boost ensemble machine learning algorithm is based on decision trees and employs the gradient boosting method to solve regression, classification, ranking, and user-defined prediction problems. This model can be better fit into small to medium size structural/tabular data.

- **LightGBM Regression**: This distributed gradient boosting algorithm employs a tree-based learning strategy. The model is quickly and more effectively trained. In addition to producing better accuracy, this model also supports parallel, distributed, and GPU learning.

- **CatBoost Regression**: The CatBoost regressor is an extension of the XGBoost and LightGBM regressors and utilizes decision trees in addition to boosting methods. With the help of this model, weak models are sequentially combined to produce powerful predictive models. The catboost model differs from previous models in that it builds oblivious trees, which enforce the requirement that all nodes be at the same level, and tests the same predictor under the same conditions. In order to produce the optimal solution and avoid overfitting, this model employs regularization techniques.

- **Lasso and Ridge Regression**: The ridge and lasso regressions are powerful regularization techniques generally used for creating demand forecasting models. The main objective of this model is to avoid high bias and variance. The ridge regression performs L2 regularization which adds the sum of squares of the coefficient with the linear regression. The lasso performs L1 regularization which adds the sum of the absolute value of the coefficient with the linear regression.

- **Bayesian Ridge Regression**: Regularization hyperparameters are also used in the Bayesian ridge regression. This algorithm is helpful when the dataset has insufficient data and the data is unevenly distributed. The model was created using the probability distribution. The goal of the model is to determine the posterior distribution or probability that an event will occur, in addition to predicting the best parameter.

- **Facebook Prophet**: Facebook Prophet is a straightforward statistical time series model that is effective at predicting multiple seasonality time series data, accurate and quick, and capable of handling outliers and other data issues on its own. This model takes into account extraneous elements like weekends and holidays. As shown in the equation 1, the Facebook prophet is a sum of three different elements.

$$Y(t) = g(t) + s(t) + h(t) + E_t \tag{1}$$

  where g(t) is the growth function and it can be linear, logistic, or flat, s(t) is seasonality, h(t) is holidays and E_t is an error term.

3.6 Model Evaluation

The most popular regression evaluation metrics are described below in order to assess the model's prediction ability.

- **Mean Absolute Error (MAE)**: The Mean Absolute Error is the mean value of the sum of the absolute difference between actual and predicted observations. This is referred in the equation 2

$$MAE = 1/n \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{2}$$

- **Mean Squared Error (MSE)**: It is the mean of the squared differences between the actual and the predicted observations as shown in the equation 3.

$$MSE = 1/n \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{3}$$

- **Root Mean Squared Error (RMSE)**: The equation 4 defines the mean root-squared difference between the actual and the predicted observations.

$$RMSE = \sqrt{1/n \sum_{j=1}^{n} (\hat{y}_j - y_j)^2} \tag{4}$$

where n - the number of observations, y_i - actual value for the each observations, and $\hat{y}_i$ - predicted value for the each observations.

- **R-Squared**: In a regression model, R-Squared (also known as R^2 or the coefficient of determination) is a statistical measure that quantifies how much of the variance in the dependent variable can be accounted for by the independent variable. The equation 5 defines the R-Squared.

$$R^2 = \frac{SS_{regression}}{SS_{total}} \tag{5}$$

Whereas $SS_{regression}$ Sum of Square due to regression and SS_{total} is the total sum of the square.

4 Design Specification

The research process flow has been adhered to in this section, as shown in figure 8. The steps in the process flow start with the research hypothesis, then include dataset acquisition, dataset analysis, data cleaning and preparation, feature selection, model building, model evaluation, and business decision-making based on the results. The research hypothesis section outlines the business goals and objectives, and the dataset acquisition stage includes steps for selecting the datasets and preparing the data. Exploratory data analysis techniques can be used to gain a deeper understanding of the data. At the data cleaning and preprocessing stage, various techniques are used to clean and prepare data. The feature selection section finds the best features using feature scoring, grid search, and correlation matrics. The implementation and evaluation of the machine learning models, including multiple linear regression, lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and gradient boosting regression models like Gradian Boosting, XG-Boosting, LightGBM, CatBoost and Facebook prophet. finally, the business can make a decision based on the evaluation results.

4.1 Software and Hardware Configuration

This section outlines the software and hardware setups necessary for this research work. The following hardware configurations will be used to conduct the proposed research: an Intel Core i7 processor, 8GB of RAM, and a 64-bit operating system.

In the implementation section, the software configurations for the research work including Jupyter Notebook 6.4.5, Google Colab, and R studio 2022.07.01+554 were used. Microsoft Office 2018 and the online LaTex editor overleaf were employed for the documentation.

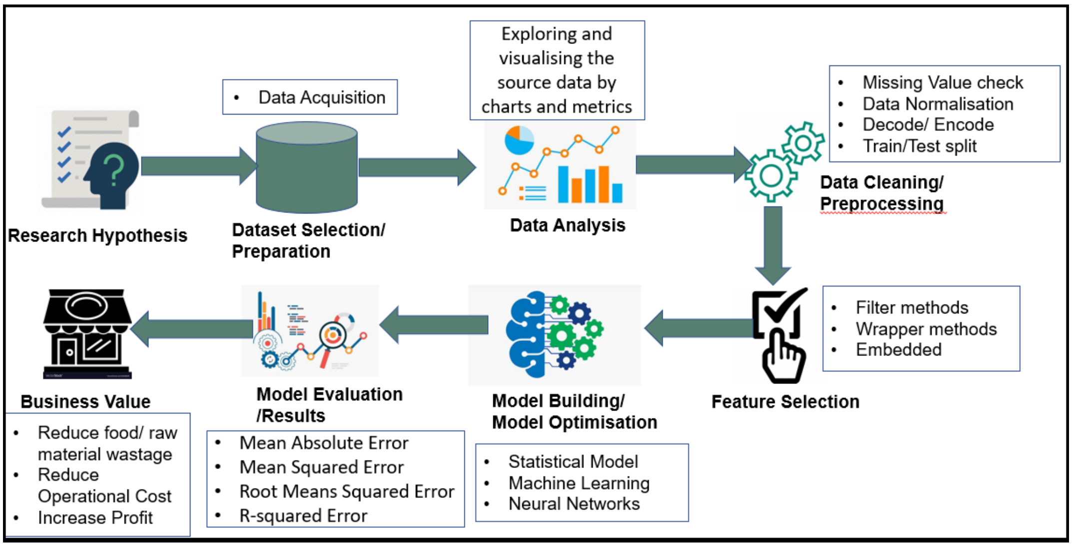

Figure 8: Process Flow

5 Implementation

In this section, the machine learning model implementation steps and strategies are described below with respect to food demand prediction. The following machine learning models are trained and tested using dataset splits of 70% and 30% and 80% and 20%.

5.1 Multiple Linear Regression

A number of independent variables and one dependent variable are used to construct the base model for the research, which is known as multiple linear regression. The backward elimination technique is used to eliminate independent variables when the statistical significance(P) of the variable is greater than 0.05 (default significant score). Each of the multiple linear regression assumptions, including homoscedasticity, independence of observation, normality, and linearity, was verified, and all of the requirements were met. A significant impact is being felt on the target variable prediction by the independent features region code, city code, meal id, emailer for promotion, and homepage features.

5.2 Lasso and Ridge Regression

Both the Lasso and the Ridge Regression use the packages LassoCV and RidgeCV to train and fit the source data, respectively. Unlike ridge, which introduces a penalty factor to the coefficients and takes the square root of the coefficients, lasso causes the coefficients to shrink towards a mean of zero. On the basis of alpha, the coefficients will shrink. While the alpha in Ridge is not set and behaves like linear regression, the alpha in Lasso is set in the range of 1 to 10, but there is no change in the results, hence we set alpha as a minimum of 1.

5.3 Bayesian Ridge Regression

The Bayesian Ridge Regression model is created using sklearn.linear_model.BeyesianRidge module. Since the dataset is not normally distributed, this Bayesian regression model is created. The response variable Y is taken to have a Gaussian distribution in order to derive the fully probabilistic model. One type of Bayesian regression that estimates a probabilistic model of the regression problem is Bayesian ridge regression. The model is fitted with both training and testing sets with 70% and 30% and 80% and 20% split.

5.4 Support Vector Regressor

The support vector Regressor (SVR) allows for some latitude in determining the error margin. The linear kernel was used to construct the SVR model. The slack parameter measures the distance between points outside the tube, while the hyperparameters C and ϵ (epsilon) refer to the regularization parameter and error margin (tube width), respectively. To make the model fit, the hyperparameters C and ϵ (epsilon) are changed because as C increases, the tolerance for a point outside ϵ also increases. In order to find the best parameters and fit them to the model, the grid-search technique is employed. The goal of the algorithm is to maximize performance while minimizing error by placing more points into the tube and reducing slack.

5.5 Decision Tree Regression

In decision tree regression, X_train is fitted into the model, which sorts all the independent variables, chooses a specific point on each variable, divides them into two data sets, and computes the MSE. Based on low MSE calculations, the split was done more effectively. When the minimum MSE is reached, the selection and splitting process repeats until a tree-like structure is formed.

5.6 Random Forest Regression

The random forest regression employs bootstrapping and ensemble learning techniques. The model is being trained using the Sklearn module with default parameters such as n_estimators=50 ,max_depth=True ,max_features=True. After that, the best optimal parameter, such as max_depth: 10, 'max_features': 'auto', 'n_estimators': 100, are chosen using grid search. Finally, the model's performance is evaluated after the model is trained with improved optimal hyperparameters.

5.7 Gradient Boosting Regression

The Gradient Boosting Regression model constructs a base tree F_0 with a single root node, calculates residual error r_0 of the same, and makes prediction P_0, Similarly builds a decision tree F_1 from the residuals of F_0 and make prediction P_1, here the model is being trained from residuals of previous trees to make better predictions. The hyperparameters n_estimators=70, learning_rate=0.1,max_depth=4, random_state=0, loss='ls' are being utilized while training the model.

5.8 XGBoost Regression

The XGBoost Regression also works based on boosting and ensemble techniques and it is built with a combination of the base leaner and objective function. Loss function and regularization term are among the objective functions. This model uses the objective loss

'Reg: Linear'. Therefore, the model will combine the base learner's predictions before eliminating the inaccurate ones and adding up the accurate ones. The hyperparameters: max_depth = 9,learning_rate=0.5,objective= 'reg:linear', eval_metric= 'rmse', seed= 4 are used to train and fit the model.

5.9 LightGBM Regression

The LightGBM Regression works based on boosting techniques on decision trees. The library lightGBM is used to train and fit the model. The GOSS and EFB are the 2 techniques used to define the characteristics of LightGBM. The hyperparameters are boosting_type='gbdt',objective='regression', max_depth=9,learning_rate = 0.5,num_leaves =60,feature_fraction = 0.8,min_data_in_leaf= 100, bagging_fraction= 0.3,metric= 'rmse', random_state=100,seed=4.

5.10 CatBoost Regression

The CatBoost library is being utilized to train and fit the regression model and SciKit Grid Search function is being employed in this model to find out the best optimal hyperparameters to train the data. Here the RMSE metric is used to evaluate the loss function. L2 regularization parameters are used. The hyperparameters used are 'depth': 8, 'iterations': 200, 'learning_rate': 0.1, 'l2_leaf_reg': 0.5.

5.11 Facebook prophet

Prophet library is similar to the Sklearn library. The weekly aggregated food orders for the beverage category are predicted as one of the most ordered food items among others. The Facebook prophet is a statistical time series model, the dataset is selected with a date variable and target variable num_order. The first step is to build the Facebook Prophet model, the second is to fit the training data to the model, and the third and final step is to predict the upcoming food orders for the following 10 weeks.

5.12 HyperParameter Tuning

Hyperparameter tuning was not performed on all the models; rather, it was performed only in the appropriate places. Some of the models are trained with default parameters. I could only notice a significant difference in random forest regression models' error rate with the tuned and default hyperparameters, hence the same was captured in the results section. However, I did not notice a significant difference in the other model's performance between tuned and default hyperparameters. The detailed descriptions of models and their hyperparameters, how I have chosen are listed below.

- For the linear regression models, The multiple linear regression model is trained without any hyperparameter because I consider it a base model. The lasso regression model was tuned with hyperparameters such as n_alphas=1, max_iter=3000, random_state=0. However, I manipulated n_alphas between 1 to 10, there is no change in the result, hence it kept it at a minimum of 1. The parameter max_iter is randomly changed between 100 to 3000 and the model error rate is keep on increasing. There are no parameters are used for ridge and BayesianRidge regression models because these models work similarly to linear regression.

- For the decision tree model, I have chosen default parameters and it resulted in an overfitting model hence I moved on to the Random forest regression model. The random forest regression model is trained with hyperparameters which are selected using GridSearchCV.
- for the boosting models, the Gradient Boosting Regression model is selected with default parameters for the regression problems and the default parameters n_estimators and max_depth are randomly changed. The default hyperparameter for XGBoost and Light GBM Regression is chosen and then the parameters are randomly modified to achieve better results. This experiment was run multiple times in order to achieve better results. Catboost regression is trained with hyperparameters which are selected through grid search.

6 Evaluation

To support the research questions and goal of predicting food demand, the comprehensive analysis of each study/experiment and its findings are described and critically analyzed with one another in this section. RMSE, MAE, R^2, and model train time are used to assess the efficiency of machine learning models. The split between the train and test portions of the model is 70% and 30%, and then 80% and 20%, respectively. The evaluation results for both the 70% and 30% split and the 80% and 20% split are described in Table 5 and Table 6.

Table 5: Evaluation Results for 70% and 30% split

Machine Learning Models	**RMSE**	**MAE**	**R^2**	**Training Time(sec)**
Multiple Linear Regression	0.64	0.71	0.73	3.16
Lasso Regression	1.04	0.85	0.27	11.10
Ridge Regression	0.64	0.49	0.73	5.40
Bayesian Ridge Regression	0.64	0.49	0.73	3.68
Decision Tree Regression	**0.18**	**0.06**	**0.98**	**7.69**
Random Forest Regression	**0.25**	**0.17**	**0.96**	**246**
Random Forest Regression with Tuned parameter	0.88	0.71	0.48	79
Gradient Boosting Regression	0.65	0.51	0.72	111
XGBoost Regression	**0.45**	**0.34**	**0.87**	**88**
LightGBM Regression	0.49	0.38	0.84	2.88
CatBoost Regression	0.55	0.42	0.80	438

6.1 Multiple Linear Regression

Multiple linear regression is used to determine the linear relationship between the predictor and response variables. The findings indicate that a number of independent features, including region code, city code, meal id, emailer for promotion, and homepage features, are significantly influencing the response variable num_orders will be placed.

Table 6: Evaluation Results for 80% and 20% Split

Machine Learning Models	**RMSE**	**MAE**	**R^2**	**Training Time(sec)**
Multiple Linear Regression	0.64	0.71	0.73	3.57
Lasso Regression	1.04	0.85	0.27	15.66
Ridge Regression	0.64	0.49	0.73	6.43
Bayesian Ridge Regression	0.64	0.49	0.73	4.51
Decision Tree Regression	**0.18**	**0.06**	**0.98**	**8.48**
Random Forest Regression	**0.26**	**0.18**	**0.96**	**274**
Random Forest Regression with Tuned parameter	0.88	0.71	0.48	101
Gradient Boosting Regression	0.65	0.51	0.72	132
XGBoost Regression	**0.45**	**0.34**	**0.86**	**100.14**
LightGBM Regression	0.49	0.37	0.84	3.30
CatBoost Regression	0.55	0.42	0.80	483

Following backward elimination techniques, the model's RMSE for both splits are reported as being 0.64 and 0.64, respectively, However, there is a slight variation in the model training time of both splits as shown in Table 5 and 6.

6.2 Lasso and Ridge Regression

Overfitting and underfitting are the two main issues that arise during model training. As a result, regularization is implemented to eliminate overfitting and underfitting using lasso and ridge regression. The RMSE of the lasso and ridge regression models are 1.04 and 0.64 respectively. The lasso with alpha 1 does not fit better with training data, while alpha with 0 in the ridge performs similarly to linear regression and the outcome is also confirmed to be the same.

6.3 Bayesian Ridge Regression

The Bayesian Ridge Regression is fit with both training and testing sets and the RMSE of the model is accounted as 0.64. There are no significant differences in both Bayesian Ridge and linear regression.

6.4 Decision Tree Regression

The Decision Tree Regression model is fitted with the training set. According to Table 5 and 6, The Decision Tree Regression model is reported as having less error(RMSE) rate is 0.18. This model outperformed all the existing models. However, it seems like the model is being overfitted with the training set.

6.5 Random Forest Regression

The decision tree model's shortcomings of overfitting and subpar performance are overcome by the Random Forest Regression model. The Random Forest Regression model has a better fit to the data, with an RMSE of 0.26. Additionally, The model's performance is downgraded to 0.88 after the model is tuned with hyperparameters. Thus, the outcome confirms that the Random Forest Regression model with default parameter outperformed. Since the training and testing portion is changed between 70% and 30% and 80% and 20%, there is a slight variation in the model's training time of 246 and 274 seconds respectively.

6.6 Gradient Boosting Regression models

The random forest is not a better option when the data is having multiple trends, The Gradient Boosting Regression model performed with an RMSE value of 0.65 on the test set. The Extreme Gradient Boosting implementation offers better model performance and is computationally efficient. The RMSE of the model is 0.45 for both data splits. XGBoost regression model outperforms all the models except random forest. The LightGBM Regression model performs better with training data and has an RMSE of 0.49. Its training time is also relatively shorter. The CatBoost Regressor Model's RMSE is 0.55 and its training time is 485 seconds which is the highest among all of them. In light of this, the XGBoost and LightGBM outperform all other boosting models and provide a better fit for the training and test data.

6.7 Facebook Prophet

The Facebook prophet model is built for predicting food orders in the beverage category. The RMSE, MAE, and R^2 for this model are 30060.78, 18559.81, and 0.48 respectively. The number of records used to train the model is limited. Thus, the performance of the model is not significant and this model can be run separately for all the food categories.

6.8 Discussion

This research aims to provide a detailed view of the methodology, implementation, and evaluation strategies for an accurate demand forecasting system for a restaurant to predict the number of food orders for the upcoming week. For which, there are several machine learning and statistical models are implemented and evaluated in the above section. The dataset is split between train and test sets, with 70% and 30% and 80% and 20% respectively. The simple multiple linear regression is the base model for this research and its RMSE is 0.64 which fits better than the research work done by (Wibowo and Yasmina; 2021) with respect to RMSE score. Additionally, lasso, ridge, and Bayesian ridge regression are used to evaluate the effectiveness of the previous research (Wibowo and Yasmina; 2021) paper, but no significant changes were found in this investigation. The SVR machine learning model is implemented with a respectable error margin, but it does not fit the train and test sets better. Additionally, training this model takes longer about 12 hours which is not recommended, and it's less efficient. When developing decision tree regression, overfitting is discovered. To address these issues, random forest regression was developed, taking into account its bagging nature. This model outperformed in terms of RMSE and accuracy all linear regression models as well as the random forest regression in the research article (Tanizaki et al.; 2020). However, the random forest regression with hyperparameters downgraded the model's performance. The random forest model is not a better option when the data exhibits multiple trends, hence the gradient-boosting

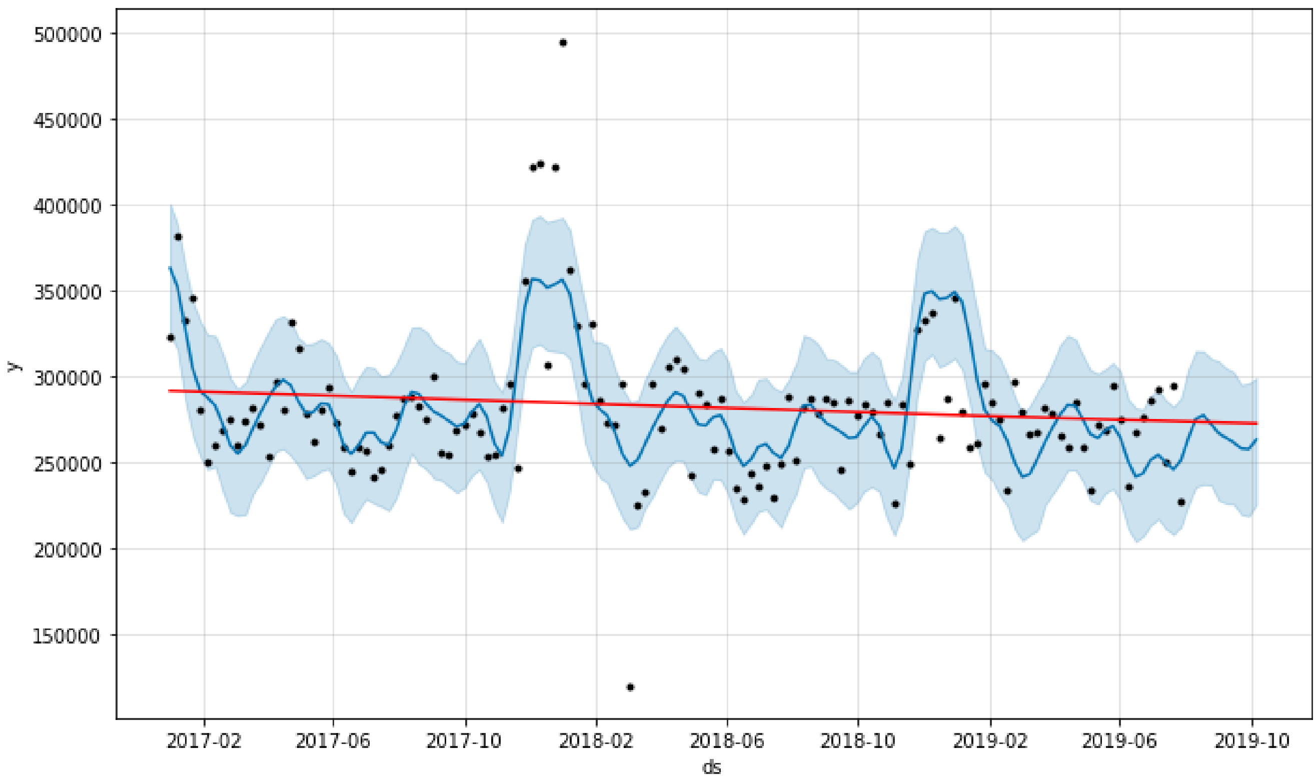

Figure 9: Forecasting plot

regression models are the right choice. The Gradient Boosting, XGBoosting, LightGBM, and CatBoost models were built to assess the performance of each other. Thus, the model XGBoost and LightGBM outperformed all the existing boosting algorithms with an RMSE of 0.45 and 0.49 respectively. Moreover, These boosting algorithms work superior to the boosting models in the research paper (Tanizaki et al.; 2019). Similar to (Menculini et al.; 2021), The Facebook prophet model is a less accurate model. Overall, random forest regression, XGBosst, and LightGBM models outperform all other models with respect to their RMSE, are better suited for the train and test sets, and are consistent with the research questions. When the splitting has been done between 70% and 30% and 80% and 20%, significant changes in the model training time have been identified, but there have been no changes in performance.

7 Conclusion and Future Work

To answer the research question objectives in the section 1 were implemented stating from the critical investigation of the food-based industries, food demand forecasting models (multiple linear regression, lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and Gradient boosting regression models such as Gradient Boosting, XGBoosting, LightGBM, and CatBoost) were implemented, evaluated and critically analyzed. All the regression models were compared with each other and also compared with existing state of art based n the evaluation metrics such as RMSE, MAE, and R^2.

The implementation has made it possible for us to comprehend the effects of the novelty feature (Food Demand) in the model and how it contributes to accuracy improvement and error rate reduction. The predictive models used were multiple linear regression, lasso, ridge, Bayesian ridge regression, SVR, decision tree, random forest, and gradient boosting regression models such as Gradient Boosting, XGBoosting, LightGBM, CatBoost, and Facebook prophet. The outcome demonstrates that random forest, XG-Boosting, and LightGBM regression models had the best performance, with RMSE scores of 0.25, 0.45, and 0.49, respectively. These models outperformed all the existing models with respect to RMSE scores. This research is limited to the weekly data and there are no external factors such as weekends, holidays, weather, and special events are not taken into account.

This research work can be extended to enrich the dataset by including external factors such as weekends, holidays, weather, and special events, and to improve its accuracy by implementing more deep learning models and neural network models like nbeats and Temporal Fusion Transformer. Additionally, instead of using weekly data, this research can be applied to daily data for better future predictions.

References

Anly Antony, M. and Kumar, R. S. (2021). A comparative study on predicting food quality using machine learning techniques, *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Vol. 1, pp. 1771–1776.

Bozkir, A. S. and Sezer, E. A. (2011). Predicting food demand in food courts by decision tree approaches, *Procedia Computer Science* **3**: 759–763. World Conference on Information Technology.

Costagliola, G., Fuccella, V., Giordano, M. and Polese, G. (2009). Monitoring online tests through data visualization, *IEEE Trans. Knowl. Data Eng.* **21**: 773–784.

Donselaar, K. H. V., Peters, J., Jong, A. D. and Broekmeulen, R. A. (2016). Analysis and forecasting of demand during promotions for perishable items, *International Journal of Production Economics* **172**.

FAO, I. (2019). The state of food and agriculture 2019. moving forward on food loss and waste reduction, *FAO, Rome* pp. 2–13.

Fattah, J., Ezzine, L., Aman, Z., Moussami, H. E. and Lachhab, A. (2018). Forecasting of demand using arima model, *International Journal of Engineering Business Management* **10**: 1847979018808673.

Fukase, E. and Martin, W. (2020). Economic growth, convergence, and world food demand and supply, *World Development* **132**: 104954.

Kakani, V., Nguyen, V. H., Kumar, B. P., Kim, H. and Pasupuleti, V. R. (2020). A critical review on computer vision and artificial intelligence in food industry, *Journal of Agriculture and Food Research* **2**: 100033.

Kim, T., Park, Y., Myung, J. and Han, E. (2018). Food price trends in south korea through time series analysis, *Public Health* **165**: 67–73.

Menculini, L., Marini, A., Proietti, M., Garinei, A., Bozza, A., Moretti, C. and Marconi, M. (2021). Comparing prophet and deep learning to arima in forecasting wholesale food prices, *Forecasting* **3**.

Pallathadka, H., Jawarneh, M., Sammy, F., Garchar, V., Sanchez, D. T. and Naved, M. (2022). A review of using artificial intelligence and machine learning in food and agriculture industry, *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 2215–2218.

Posch, K., Truden, C., Hungerländer, P. and Pilz, J. (2022). A bayesian approach for predicting food and beverage sales in staff canteens and restaurants, *International Journal of Forecasting* **38**.

Tanizaki, T., Hoshino, T., Shimmura, T. and Takenaka, T. (2019). Demand forecasting in restaurants using machine learning and statistical analysis, *Procedia CIRP* **79**: 679–683. 12th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 18-20 July 2018, Gulf of Naples, Italy.

Tanizaki, T., Hoshino, T., Shimmura, T. and Takenaka, T. (2020). Restaurants store management based on demand forecasting, *Procedia CIRP* **88**: 580–583. 13th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 17-19 July 2019, Gulf of Naples, Italy.

van Dijk, M., Morley, T., Rau, M. L. and Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050, *Nature Food* **2**.

Vithitsoontorn, C. and Chongstitvatana, P. (2022). Demand forecasting in production planning for dairy products using machine learning and statistical method, *2022 International Electrical Engineering Congress (iEECON)*, pp. 1–4.

Wibowo, A. and Yasmina, I. (2021). Food price prediction using time series linear ridge regression with the best damping factor, *Adv. Sci. Technol. Eng. Syst. J* **6**: 694–698.

Çetinkaya, Z. and Erdal, E. (2019). Daily food demand forecast with artificial neural networks: Kırıkkale university case, *2019 4th International Conference on Computer Science and Engineering (UBMK)*, pp. 1–6.