ANALYSIS OF ONLINE FOOD ORDERING BEHAVIOUR - ECS784P

Abstract– This report presents an analysis of online food ordering behaviour using a dataset containing information about customers such as age, gender, marital status, occupation, income level, educational qualifications, family size, location, and feedback on their ordering experience. The objective is to understand the factors influencing customers' decisions to order food online and their overall satisfaction with the service. The analysis involves data preprocessing, exploratory data analysis (EDA), and the application of machine learning models to predict customer behaviour. The findings from this study can provide valuable insights to online food ordering businesses for enhancing their services and targeting specific customer segments effectively.

I. Introduction

In recent years, the online food ordering industry has witnessed tremendous growth, driven by increasing internet penetration, convenience, and changing consumer preferences. With a plethora of online food delivery platforms available, understanding customer behaviour and preferences has become crucial for businesses operating in this domain. This report aims to analyse a dataset containing customer information and feedback to uncover patterns and trends that can aid in improving online food ordering services.

Objective

The primary objective of this analysis is to gain insights into the factors that influence customers' decisions to order food online and their satisfaction with the service. Specifically, the analysis aims to:

- Preprocess and prepare the dataset for analysis by handling missing values, converting categorical data to numerical form, and addressing any data quality issues.
- Perform exploratory data analysis (EDA) to understand the distribution and relationships between different variables in the dataset.
- Investigate the impact of factors such as age, gender, marital status, occupation, income level, educational qualifications, and family size on customers' online food ordering behaviour.
- Analyse customer feedback to identify areas for improvement and potential factors contributing to positive or negative experiences.
- Apply machine learning models to predict customer behaviour and evaluate their performance.
- Derive actionable insights and recommendations for online food ordering businesses based on the analysis findings.

II. Literature Review

Recent years have seen a sharp increase in the online food ordering market, fueled by convenience, shifting customer tastes, and rising internet coverage. Numerous investigations have looked into what influences customer behaviour in this area. Customers' decisions to purchase meals online are influenced by a number of important aspects, including price, convenience, time savings, and range of selections [1]. Online meal ordering has been a popular option among customers since it allows them to explore menus, make orders, and get delivery all from the comfort of their homes or places of business. Age, income, and level of education are examples of demographic characteristics that have been found to influence preferences for ordering meals online. Pigatto et al. (2017)[2] suggest that younger generations and individuals with higher income levels are more likely to embrace online food ordering services. Higher educated people also have a tendency to be more open to embracing new services and technology, such as online meal ordering. Factors including order accuracy, delivery time, and quality of service have been connected to customer happiness and loyalty in the online food ordering space [3]. Yeo et al. (2017)[4] emphasise the importance of providing a seamless and consistent experience to customers, from the ordering process to the actual delivery of the food, in order to foster customer satisfaction and repeat business.

Studies have also explored the factors influencing consumer adoption of online food ordering services. Kapoor and Vij (2018)[1] identify perceived usefulness, ease of use, and trust as key determinants of consumer adoption. Furthermore, social influences—like peer recommendations and online reviews—can have a big impact on how consumers interpret and make judgments about ordering takeout online [5].

III. Data Preprocessing

The dataset used in this analysis is taken from Kaggle. It contains 388 records and 13 columns, including customer information such as age, gender, marital status, occupation, monthly income, educational qualifications, family size, location (latitude and longitude), pin code, output (whether they ordered food or not), feedback, and an unnamed column. Data Source: Kaggle-Online Food Dataset (Refer Fig. 1).

A) Looking at top 10 records

```
[3] df.head(10)
```

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitude	Pin code	Output	Feedback	Unnamed: 12
0	20	Female	Single	Student	No Income	Post Graduate	4	12.9766	77.5993	560001	Yes	Positive	Yes
1	24	Female	Single	Student	Below Rs.10000	Graduate	3	12.9770	77.5773	560009	Yes	Positive	Yes
2	22	Male	Single	Student	Below Rs.10000	Post Graduate	3	12.9551	77.6593	560017	Yes	Negative	Yes
3	22	Female	Single	Student	No Income	Graduate	6	12.9473	77.5616	560019	Yes	Positive	Yes
4	22	Male	Single	Student	Below Rs.10000	Post Graduate	4	12.9850	77.5533	560010	Yes	Positive	Yes
5	27	Female	Married	Employee	More than 50000	Post Graduate	2	12.9299	77.6848	560103	Yes	Positive	Yes
6	22	Male	Single	Student	No Income	Graduate	3	12.9770	77.5773	560009	Yes	Positive	Yes
7	24	Female	Single	Student	No Income	Post Graduate	3	12.9828	77.6131	560042	Yes	Positive	Yes
8	23	Female	Single	Student	No Income	Post Graduate	2	12.9766	77.5993	560001	Yes	Positive	Yes
9	23	Female	Single	Student	No Income	Post Graduate	4	12.9854	77.7081	560048	Yes	Positive	Yes

Fig 1. Dataset Sample

The data preprocessing steps involved:

- Checking for missing values, and no missing values were found in the dataset.
- Dropping the unnamed column as it was not relevant to the analysis.
- Converting categorical variables such as gender, marital status, occupation, income level, educational qualifications, and feedback into numerical form for analysis and modelling.

IV. Exploratory Data Analysis (EDA)

The EDA phase involved exploring the dataset through various visualisations and statistical summaries to gain insights into the data. Key findings from the EDA include:

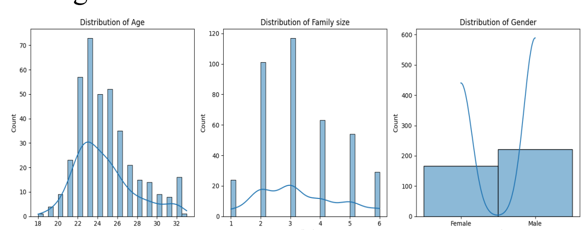

Fig 2. Data distribution of some key Features

- Age Distribution: The dataset covered customers aged between 18 and 33 years, with a mean age of approximately 25 years. (Refer Fig. 2)
- Gender Representation: The dataset contained records for both males (57%) and females (43%).

- Marital Status: A majority of the customers were single (69%), followed by married individuals (28%).
- Occupation: Students (53%) and employees (30%) constituted the majority of the customer base.
- Income Levels: A significant portion of customers (48%) reported no income, likely due to the high representation of students. (Refer Fig. 3)

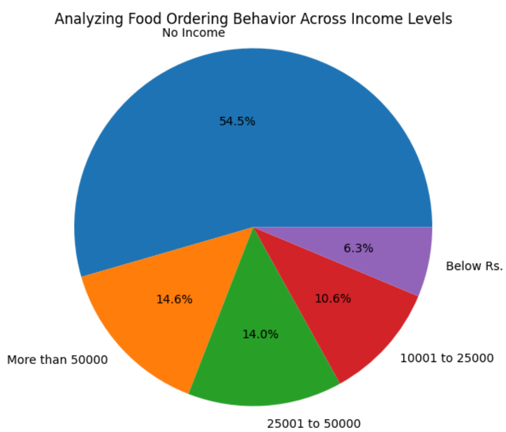

Fig 3. Food Ordering behaviour across various Income levels

- Educational Qualifications: The dataset covered individuals with varying educational backgrounds, including graduates (46%), postgraduates (45%), and Ph.D. holders (6%).
- Family Size: The average family size was around 3 members.
- Income and Ordering Behaviour: Customers with higher income levels (above Rs. 50,000) exhibited a higher tendency to order food online compared to those with lower or no income.
- Age and Ordering Behaviour: Customers in the age range of 20-26 years were more likely to order food online. (Refer Fig. 4)

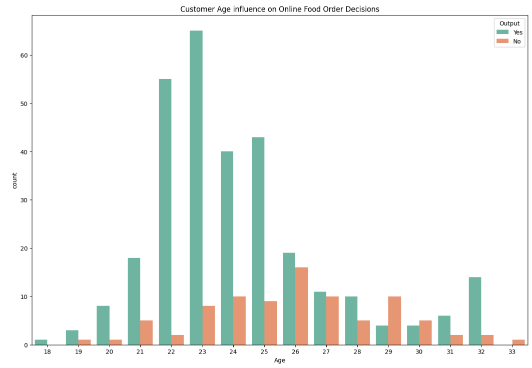

Fig 4. Customer Age influence on Online food ordering

- Gender and Ordering Behaviour: There was a slightly higher percentage of males (58.1%) ordering food online compared to females (41.9%). (Refer Fig. 5)

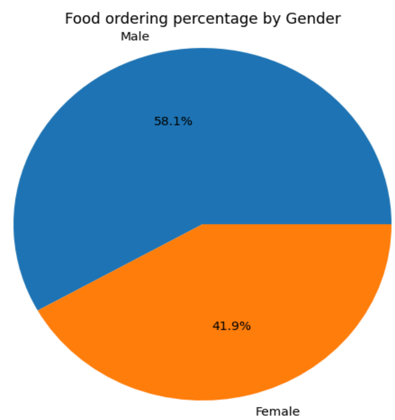

Fig 5. Food Ordering percentage by Gender

V. Methodologies

To analyse the online food ordering behaviour and predict customer decisions, two machine learning models were employed:

Random Forest Classifier: This ensemble learning method combines multiple decision trees to improve prediction accuracy and handle overfitting. Random Forest Classifiers are known for their robustness and ability to capture non-linear relationships in the data.

Classification problems like the one in this analysis, which predict whether or not a customer will order food online based on various features like age, gender, income level, and educational qualifications, are a good fit for Random Forest Classifiers.

<u>Logistic Regression:</u> A widely used statistical model for binary classification problems, Logistic Regression models the probability of an event occurring based on the input features. It is a simple and interpretable model that can provide insights into the relative importance of different features.
Even with huge datasets, logistic regression may be trained fairly quickly and is computationally efficient. It can still function as a useful baseline for comparison and as a more straightforward, understandable model, even though it might not always perform as well as more sophisticated models.

<u>Train Test Split:</u>
The dataset was split into training and testing sets (75% for training and 25% for testing) to evaluate the performance of the models.

VI. Results

The results of the analysis are summarised below:
<u>Random Forest Classifier:</u> The Random Forest Classifier achieved an accuracy of 86.60% on the test dataset, indicating its effectiveness in predicting online food ordering behaviour based on the given features.
<u>Logistic Regression:</u> The Logistic Regression model obtained an accuracy of 77.32% on the test dataset, which is lower than the Random Forest Classifier.
<u>Feature Importance:</u> The correlation matrix revealed that factors such as age, income level, and educational qualifications had a relatively stronger correlation with the target variable (online food ordering decision) compared to other features like gender, marital status, and family size. (Refer Fig. 6)
<u>Customer Feedback:</u> The analysis of customer feedback showed that approximately 82% of the customers provided positive feedback, while 18% expressed negative feedback. This information can be valuable for identifying areas of improvement and addressing customer concerns.
<u>Confusion Matrix:</u> The confusion matrix for the Random Forest Classifier showed that it correctly classified 68 instances as positive (ordered food) and 16 instances as negative (did not order food). However, it misclassified 6 positive instances as negative and 7 negative instances as positive.
<u>Cross-Validation:</u> The cross-validation scores for the Random Forest Classifier (mean accuracy of 90.37% with a standard deviation of 2.80%) were higher than the Logistic Regression model (mean accuracy of 77.66% with a standard deviation of 0.30%), further reinforcing the superior performance of the Random Forest Classifier on this dataset.

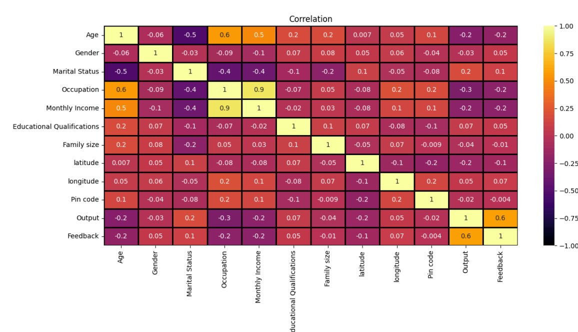

Fig 6. Feature Correlation matrix

VII. Conclusion

The analysis of the online food ordering dataset provided valuable insights into customer behaviour and preferences. The key findings and conclusions are as follows:

- Age, income level, and educational qualifications emerged as crucial factors influencing customers' decisions to order food online.

- Students and young professionals (20-26 years) were more likely to order food online, potentially due to convenience and lifestyle preferences.
- Customers with higher income levels exhibited a greater tendency to order food online compared to those with lower or no income.
- The Random Forest Classifier outperformed the Logistic Regression model in terms of prediction accuracy, achieving an accuracy of 86.60% on the test dataset.
- Customer feedback analysis revealed a high satisfaction rate (82% positive feedback) but also highlighted areas for improvement to enhance the overall customer experience.
- Based on these findings, online food ordering businesses can tailor their marketing strategies, service offerings, and customer support to cater to the preferences and needs of their target customer segments more effectively.

Recommendations

Based on the analysis and conclusions, the following recommendations can be made for online food ordering businesses:

- Target Marketing: Create focused marketing campaigns aimed at the age and socioeconomic groups that are most likely to order takeaway online.
- Student-centric Offerings: To meet the needs of the vast majority of students who order meals online, investigate student-centric pricing schemes, discounts, and meal plans.
- Convenience and Reliability: When promoting services, put special emphasis on order accuracy, speedy delivery, and convenience as these attributes are likely to appeal to the intended customer segments.
- Customer service: Take action to resolve issues raised by clients and enhance areas like poor service quality, order errors, and delivery delays that were mentioned in complaints.
- Personalisation: To improve customer satisfaction and retention, use customer data and preferences to create customised promotions, loyalty programmes, and recommendations.
- Continuous Monitoring: To stay ahead of the competition and adjust to shifting customer needs and preferences, keep a close eye on competitor offerings, market trends, and customer feedback.

Limitations and Future Work

While this analysis provides valuable insights, it is important to note the following limitations:

The results may not apply to other areas or localities because the dataset only covers a fraction of the community data available. Data on variables that may affect consumer satisfaction and ordering behaviour, such as delivery time, order accuracy, as well as service quality, are missing from the dataset. The analysis does not take into account the influence of outside variables like marketing trends or promotional campaigns, nor does it take into account possible shifts in consumer behaviour over time.

<u>Future research in this domain could focus on:</u>

Incorporating additional data sources and variables, such as social media sentiment, competitor analysis, and market trends, to enhance the understanding of customer behaviour.

Carrying out long-term research to track shifts in consumer behaviour and preferences.

Investigating cutting-edge machine learning methods to extract more meaningful insights from unstructured data and customer feedback, such as deep learning and natural language processing.
Analysing how the online meal ordering experience is affected by new technologies such as voice-activated ordering and augmented reality.

VIII. References

1. Kapoor, A. P., & Vij, M. (2018). Technology at the dinner table: Ordering food online and restaurant patronage. Journal of Retailing and Consumer Services, 43, 215-229. URL:https://ideas.repec.org/a/eee/joreco/v43y2018icp342-351.html
2. Pigatto, G., Machado, J. G. C. F., Negreti, A. D. S., & Machado, L. M. (2017). Have you chosen to eat at the restaurant today? Human Factors on the Choice of Consumption Products and Services. British Food Journal, 119(8), 1870-1882. URL:https://www.researchgate.net/publication/314244684_Have_you_chosen_your_request_Analysis_of_online_food_delivery_companies_in_Brazil
3. Qin, H., & Prybutok, V. R. (2008). Determinants of customer-perceived service quality in fast food restaurants and their relationship to customer satisfaction and behavioural intentions. Quality Management Journal, 15(2), 35-50. URL:https://www.researchgate.net/publication/285661523_Determinants_of_Customer-Perceived_Service_Quality_in_Fast-Food_Restaurants_and_Their_Relationship_to_Customer_Satisfaction_and_Behavioral_Intentions
4. Yeo, V. C. S., Goh, S. K., & Rezaei, S. (2017). Consumer experiences, attitude and behavioural intention toward online food delivery (OFD) services. Journal of Retailing and Consumer Services, 35, 150-162. URL: https://doi.org/10.1016/j.jretconser.2016.12.013
5. Ioanăs, E., & Stoica, I. (2014). Social media and its impact on consumers behaviour. International Journal of Economic Practices and Theories, 4(2), 295-303. URL:https://www.researchgate.net/publication/313421625_Social_Media_and_its_Impact_on_Consumers_Behavior
6. Kaggle-Online Food Dataset