IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ※ Peer-reviewed & Refereed journal ※ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

Forecasting Food Delivery Time: An Exploration of Predictive Models and Factors Impacting Delivery Time Estimation

Dr. Ayyappa Chakravarthi M[1], Shaik Eesa Ruhulla Haq[2], Madapakula Venkata Anil[3], Bathula Venkata Vamsi[4], Gogireddy Venkata Reddy[5]

Associate Professor, Department of CSE-Data Science, KKR & KSR Institute of Technology and Sciences. [1]

BTech CSE-Data Science, KKR & KSR Institute of Technology and Sciences, Guntur, Andhra Pradesh, India.[2-5]

Abstract: With the exponential growth of the online food delivery industry, ensuring timely deliveries has become paramount to customer satisfaction and business success. However, the persistent challenge of delayed delivery continues to plague food delivery companies, resulting in customer dissatisfaction and potential revenue loss. This project delves into the domain of food services to investigate the complexities surrounding the accurate estimation of delivery times. Through the exploration of predictive models and the analysis of various factors influencing delivery time estimation, including geographical variables, traffic patterns, order complexity, and operational dynamics, the study aims to develop robust forecasting mechanisms. By leveraging historical data and employing advanced analytical techniques, this research seeks to uncover insights that can enhance operational efficiency and mitigate delivery delays effectively. Ultimately, the outcomes of this study are poised to contribute valuable insights to the online food delivery industry, fostering improved customer satisfaction, retention, and sustained growth in this dynamic market landscape.

Keywords: Delivery time estimation, Customer satisfaction, Predictive models, Geographical variables, Traffic patterns

I. INTRODUCTION

The online food delivery industry has witnessed exponential growth, underscoring the critical importance of timely deliveries for ensuring customer satisfaction and sustaining business success. However, persistent challenges with delayed deliveries have plagued food delivery companies, leading to customer dissatisfaction and potential revenue loss. In response to these challenges, this research endeavors to delve into the complexities surrounding the accurate estimation of delivery times within the food services domain.

Through the exploration of predictive models and the comprehensive analysis of various factors influencing delivery time estimation, including geographical variables, traffic patterns, order complexity, and operational dynamics, this study aims to develop robust forecasting mechanisms. Leveraging historical data and employing advanced analytical techniques, the research seeks to uncover valuable insights that can enhance operational efficiency and effectively mitigate delivery delays.

Ultimately, the outcomes of this study are poised to contribute significant advancements to the online food delivery industry. By fostering improved customer satisfaction, retention, and facilitating sustained growth in this dynamic market landscape, the research endeavors to address critical challenges and pave the way for enhanced service delivery in the food services sector.

A. MOTIVATION

The motivation for this project lies in improving the accuracy of food delivery time predictions, addressing existing inefficiencies in the online food delivery sector. Our goal is to enhance customer satisfaction and operational efficiency through advanced predictive models, contributing to a more seamless and optimized food delivery experience.

B. PROBLEM DEFINATION

The problem addressed in this project is the accurate estimation of food delivery times in the online food delivery industry. The challenge lies in developing effective predictive models that consider various factors such as geographical variables, traffic patterns, and other relevant parameters to provide precise estimates. The goal is to optimize and enhance the prediction process, contributing to improved operational efficiency, customer satisfaction, and overall performance in the dynamic landscape of food delivery services.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

II. DATA ANALYSIS AND PREPROCESSING

In this section, we elucidate our approach, starting with a presentation of both the existing and derived features. Subsequently, we provide an overview of the preprocessing techniques applied.

A. DATASET DESCRIPTION

The "Forecasting food delivery time" dataset released UberEATS, an American professional services firm. The data is containing about 45,593 entries and 20 features, which are explained as below:

1. **ID:** Unique identifier for each record.
2. **Delivery_person_ID:** Identifier for delivery personnel.
3. **Delivery_person_Age:** Age of the delivery person.
4. **Delivery_person_Ratings:** Ratings assigned to the delivery person.
5. **Restaurant_latitude:** Latitude coordinates of the restaurant.
6. **Restaurant_longitude:** Longitude coordinates of the restaurant.
7. **Delivery_location_latitude:** Latitude coordinates of the delivery location.
8. **Delivery_location_longitude:** Longitude coordinates of the delivery location.
9. **Order_Date:** Date of the order.
10. **Time_Ordered:** Time when the order was placed.
11. **Time_Order_picked:** Time when the order was picked up.
12. **Weatherconditions:** Weather conditions at the time of delivery.
13. **Road_traffic_density:** Density of road traffic during delivery.
14. **Vehicle_condition:** Condition of the delivery vehicle.
15. **Type_of_order:** Type of order placed (e.g., buffet, drinks, meal).
16. **Type_of_vehicle:** Type of vehicle used for delivery.
17. **multiple_deliveries:** Indicator if multiple deliveries were made.
18. **Festival:** Whether it was a festival day or not.
19. **City:** City where the delivery was made.
20. **Time_taken(min):** Time taken for delivery in minutes.

B. DATASET ANALYSIS

In this part, we are going with the analysis, they are following, starting with the Univariate analysis and then go with the Multivariate analysis.

1. Univariate analysis

we analyze the number of orders in relation to various features in the univariate analysis

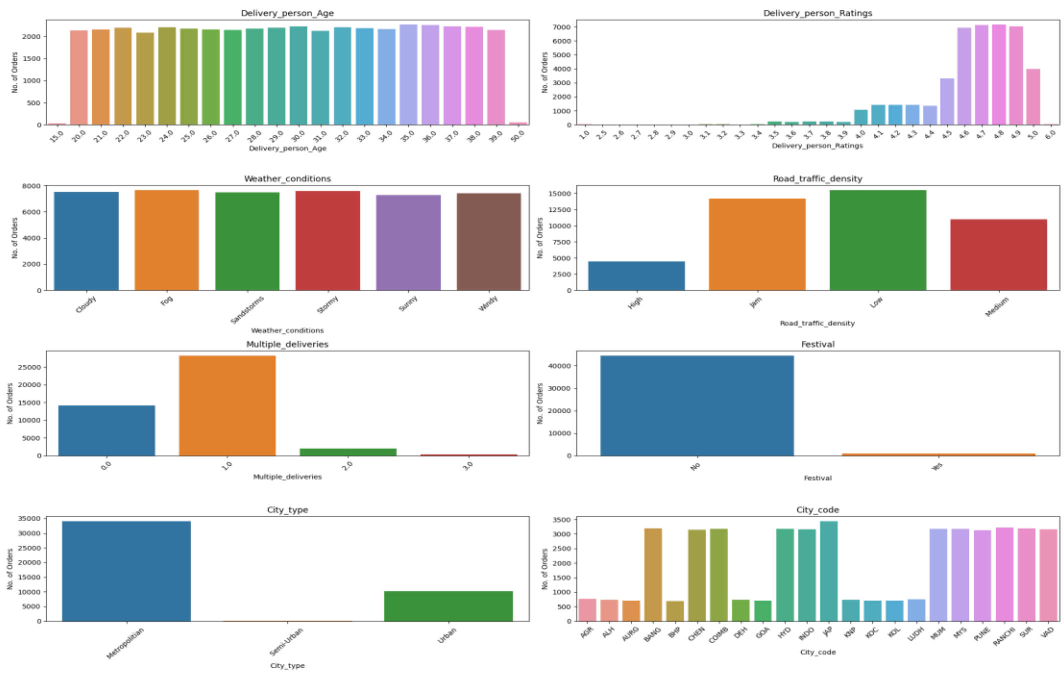

Fig: 1 Univariate Result

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

2. Multivariate analysis

we analyze two or more features at a time in multivariate analysis

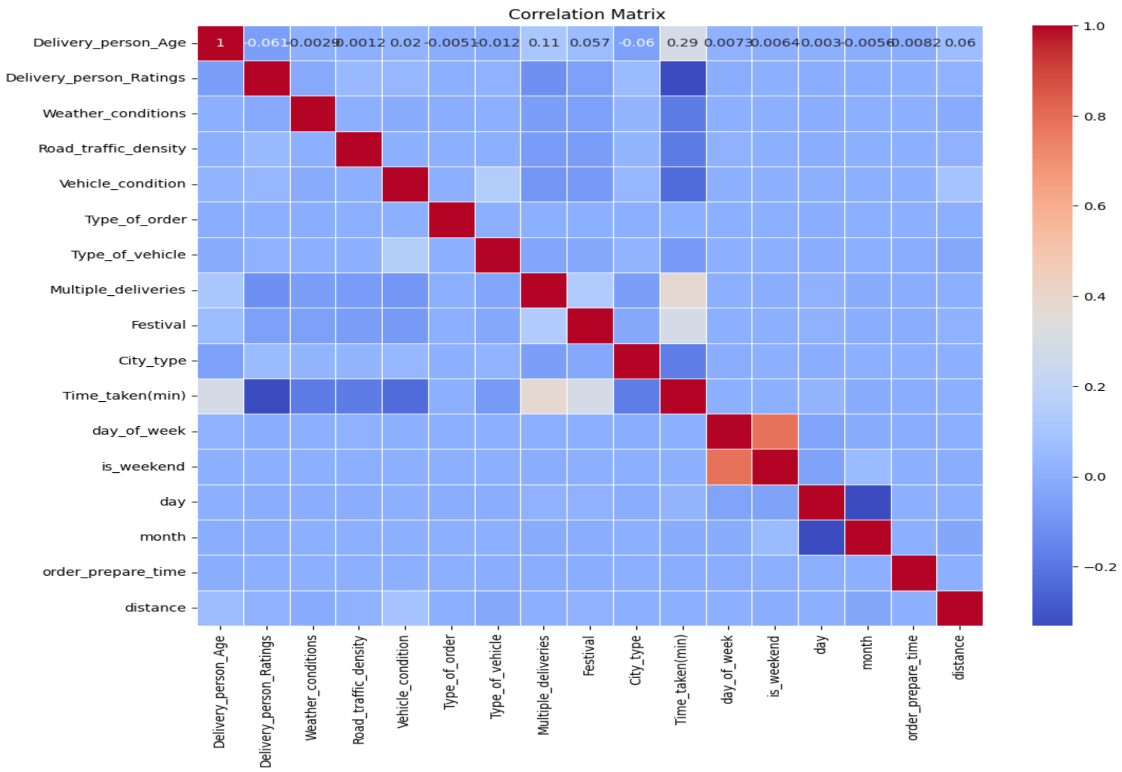

Fig: 2 Multivariate Result

C. FEATURE ENGINEERING

We derive additional features from the existing ones to expanding the dataset's information, facilitating valuable insights and enhancing overall analysis. The newly calculated features contribute to a more comprehensive understanding of the dataset.

1. **City_code:** Derived from Delivery_person_ID, this unique identifier categorizes delivery locations by city.
2. **day_of_week:** Numeric representation obtained from Order_Date, providing insight into the day for temporal analysis.
3. **is_weekend:** Binary indicator derived from day_of_week, signaling whether the corresponding day is a weekend.
4. **day:** Extracted from Order_Date, specifying the day of the month for chronological reference.
5. **month:** Extracted from Order_Date, denoting the month in which the order was placed.
6. **order_prepare_time:** Calculated as the duration between Time_Ordered and Time_Order_picked, reflecting the time taken for order preparation.
7. **distance:** Computed using the geodesic library, this metric measures the spatial separation between Restaurant_coordinates and Delivery_coordinates.

D. PRE-PROCESSING TECHNIQUES

1. Label Encoding for Categorical Variables: Label Encoding is a preprocessing technique employed to convert categorical variables into numerical representations. It assigns a unique integer label to each category, allowing machine learning algorithms to work with categorical data more effectively.

In this context, Label Encoding is applied to categorical features like 'Weather_conditions,' 'Road_traffic_density,' 'Type_of_order,' 'Type_of_vehicle,' 'Festival,' and 'City_type' in the dataset.

2. Standardization of Features: Standardization is a preprocessing step that involves scaling numerical features to have zero mean and unit variance. In the project, it is utilized to standardize features such as 'Age,' 'Ratings,' 'Order_prepare_time,' and 'Distance.' Standardized features ensure that all numerical inputs are on a consistent scale, preventing certain features from dominating the model training process.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

III. METHODOLOGY AND METRICS

In this section, we present an outline of the forecasting methodologies implemented in our study. The models undergo training and validation on 80% of the data, followed by testing on the remaining 20%.

The hardware configuration includes a minimum of 8GB RAM. The primary aim of this section is to list all the models employed before undertaking a metric-based evaluation of their individual outcomes.

IV. A. MACHINE LEARNING MODELS

A. MACHINE LEARNING MODELS

In this project, we used the following machine learning models they are namely,

LinearRegression(), RandomForestRegressor(), AdaBoostRegressor(), KNeighborsRegressor(), DecisionTreeRegressor(), GradientBoostingRegressor(), XGBRegressor(), LGBMRegressor().

B. METRICS RESULT

In this part, the metrics results are shown to the all models.

Metrics are used in our models are Mean absolute error**(MAE)** Mean square error**(MSE)**, Root mean squared error(**RMSE),**R-Squared**(R_2)**, Accuracy and the respective each model metric results are given below.

	Modelling Name	MAE	MSE	RMSE	R_2	Accuracy
0	LinearRegression	5.70300	51.02061	7.142872	0.418092	41.809195
1	RandomForestRegressor	3.23028	16.72567	4.089703	0.809238	80.923786
2	AdaBoostRegressor	5.01770	37.66702	6.137346	0.570394	57.039438
3	KNeighborsRegressor	5.21195	44.27212	6.653730	0.495061	49.506087
4	DecisionTreeRegressor	4.20688	30.74572	5.544883	0.649334	64.933421
5	GradientBoostingRegressor	3.58849	20.42712	4.519637	0.767022	76.702151
6	XGBRegressor	3.20604	16.26452	4.032929	0.814497	81.449748
7	LGBMRegressor	3.15380	15.62794	3.953219	0.821758	82.175785

Fig: 3 Metrics Result

C. HYPERPARAMETER TUNING RESULT

Hyperparameter tuning is the optimization process of selecting the best external configurations, known as hyperparameters, for a machine learning model. These settings, like learning rates or the number of hidden layers, are crucial for a model's performance but are not learned during training.

In our project, **we are applied hyperparameter tuning using GridSearchCV to experiment** with different values and find the optimal hyperparameter sets for all the models. This process enhances model performance and ensures better generalization to new data. And the following are the hyper-parameters results to the all the machine learning models applied.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

	Modelling Name	Best Parameters	Best R2 Score
0	LinearRegression	{}	0.416317
1	RandomForestRegressor	{'max_depth': None, 'n_estimators': 300}	0.813927
2	AdaBoostRegressor	{'learning_rate': 0.1, 'n_estimators': 50}	0.565688
3	KNeighborsRegressor	{'n_neighbors': 7}	0.501603
4	DecisionTreeRegressor	{'max_depth': None}	0.650988
5	GradientBoostingRegressor	{'learning_rate': 0.5, 'n_estimators': 150}	0.789939
6	XGBRegressor	{'learning_rate': 0.1, 'max_depth': 9, 'n_esti...	0.818343
7	LGBMRegressor	{'learning_rate': 0.5, 'max_depth': None, 'n_e...	0.818534

Fig: 4 Best Parameters for every Model

V. RESULTS AND DISCUSSION

In the results and discussion section, the food delivery time prediction involves taking specific inputs. These inputs consist of various features crucial for the prediction process. The chosen features serve as essential input parameters for the accurate estimation of food delivery times.

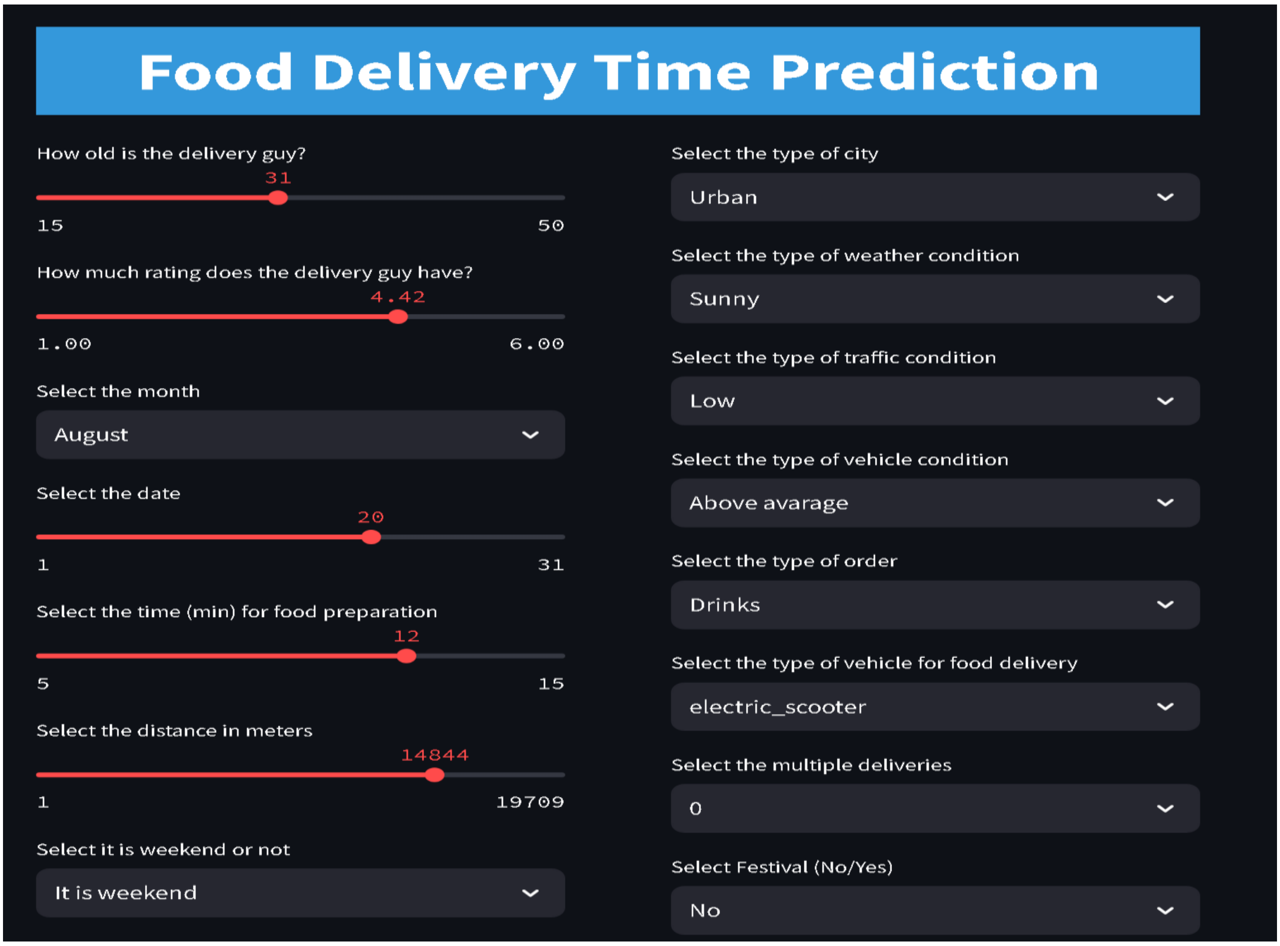

Fig: 5 Intake of inputs

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 13, Issue 3, March 2024
DOI: 10.17148/IJARCCE.2024.13318

In the results and discussion section, the process begins with the intake of inputs for food delivery time prediction. After data processing, the final output emerges as the predicted time for food delivery, represented in minutes.

Fig: 6 Predicted time for food delivery

VI. CONCLUSION

In conclusion, this study meticulously explores predictive models for accurate food delivery time estimation. The project involves comprehensive data analysis, including data cleaning, feature engineering, and rigorous model evaluation, resulting in a robust predictive model optimized through parameter tuning.

The encoding of categorical variables enhances model effectiveness, with evaluation metrics emphasizing the model's accuracy. Notably, the project introduces innovative approaches, such as geospatial distance calculations for determining the physical distance between the restaurant and the delivery location, contributing valuable insights to optimize food delivery predictions in the dynamic online food delivery industry. The findings hold significant implications for improving operational efficiency, customer satisfaction, and sustained growth in this rapidly evolving market.

REFERENCES

[1]. Lidya Chitra Laoh; Timothy Adithia Pongantung; Carolin Mulalinda, Android Application Food Delivery Services, DOI: 10.1109/ICORIS50180.2020.9320843, 18 January 2021
[2]. Tejas Raibagi; Ashwin Vishwakarma; Jahnavi Naik; Rujata Chaudhari; Geetanjali Kalme, Orderista - AI-based Food Ordering Application, DOI: 10.1109/ICAIS50930.2021.9396040, 12 April 2021
[3]. Cristina-Edina Domokos; Barna Séra; Károly Simon; Lajos Kovács; Tas-Béla Szakács, Netfood: A Software System for Food Ordering and Delivery, DOI: 10.1109/SISY.2018.8524854, 08 November 2018
[4]. I Gusti Made Karmawan; Irenea Jessica Dwitahthea Susanto; Anderes Gui; Muhammad Shabir Shaharudin; Yudi Fernando; Ridho Bramulya Ikhsan, Factors Influencing Purchase Intention of Cloud-based Food Delivery, DOI: 10.1109/ICITACEE53184.2021.9617543, 03 December 2021
[5]. Albin Antony; P. Sivraj, Food Delivery Automation in Restaurants Using Collaborative Robotics, DOI: 10.1109/ICIRCA.2018.8597280, 03 January 2019
[6]. Krishna Kumar Kottakki; Sunil Rathee; Kranthi Mitra Adusumilli; Jose Mathew; Bharat Nayak; Saket Ahuja, Customer Experience Driven Assignment Logic for Online Food Delivery, DOI: 10.1109/IEEM45057.2020.9309783, 12 January 2021
[7]. Md. Mosfikur Rahman; Musfiqur Rahman Foysal; Nazmun Nessa Moon; Fernaz Narin Nur, TraFoo: An Android Application for Food Delivery in Train, DOI: 10.1109/SMARTGENCON51891.2021.9645900, 21 December 2021
[8]. Kequan Li; Yan Chen; Jiachen Liu; Lin Zhang; Xiangwei Mu, Online Food Delivery Platforms and Restaurants' Interactions in the Context of the Ban on Using Single-Use Plastics, DOI: 10.1109/ACCESS.2021.3095296, 07 July 2021
[9]. Ziboud Van Veldhoven; Paulien Aerts; Sanne Lies Ausloos; Jente Bernaerts; Jan Vanthienen, The Impact of Online

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2024

DOI: 10.17148/IJARCCE.2024.13318

Delivery Services on the Financial Performance of Restaurants, DOI: 10.1109/ICIM52229.2021.9417153, 30 April 2021

[10]. Ehsanur Rahman Rhythm; Rajvir Ahmed Shuvo; Md Sabbir Hossain; Md. Farhadul Islam; Annajiat Alim Rasel, Sentiment Analysis of Restaurant Reviews from Bangladeshi Food Delivery Apps, DOI: 10.1109/ESCI56872.2023.10100214, 19 April 2023

[11]. YongChai Tan; BentFei Lew; KhimLeng Tan; KaeVin Goh; KienLoong Lee; ZhiChao Khor, A new Automated Food Delivery System using autonomous track guided centre-wheel drive robot, DOI: 10.1109/STUDENT.2010.5686986, 13 January 2011

[12]. Jie Tian, Analysis on the safety design of food delivery packaging in the post-epidemic era, DOI: 10.1109/CTMCD53128.2021.00062, 28 June 2021

[13]. Milan Markovic; Peter Edwards; Naomi Jacobs, Recording Provenance of Food Delivery Using IoT, Semantics and Business Blockchain Networks, DOI: 10.1109/IOTSMS48152.2019.8939250, 23 December 2019

[14]. Siddhartha Paul; Sunil Rathee; Jose Matthew; Kranthi Mitra Adusumilli, An Optimization Framework for On-Demand Meal Delivery System, DOI: 10.1109/IEEM45057.2020.9309922, 12 January 2021

[15]. Gurmeet Singh; Shavneet Sharma; Anushree Tandon; Puneet Kaur, Drone Food Delivery: A Solution to Crowding During the Global COVID-19 Pandemic, DOI: 10.1109/TEM.2022.3181251, 19 October 2022

[16]. Yan Liu; Bin Guo; Chao Chen; He Du; Zhiwen Yu; Daqing Zhang; Huadong Ma; FooDNet: Toward an Optimized Food Delivery Network Based on Spatial Crowdsourcing; DOI: 10.1109/TMC.2018.2861864, 01 August 2018

[17]. Siddhartha Paul; Goda Doreswamy, An Optimization Model for Priority-Based On-Demand Meal Delivery System, DOI: 10.1109/IEEM55944.2022.9989782, 26 December 2022

[18]. Jie Zheng; Ling Wang; Li Wang; Shengyao Wang; Jing-Fang Chen; Xing Wang, Solving Stochastic Online Food Delivery Problem via Iterated Greedy Algorithm With Decomposition-Based Strategy, DOI: 10.1109/TSMC.2022.3189771, 21 July 2022

[19]. Xiangyu Kong; Guangyu Zou; Heng Qi; Jiafu Tang, Optimization of O2O Food Delivery Strategy in Smart Cities, DOI: 10.1109/ISC255366.2022.9921961, 26 October 2022

[20]. Farhana Huq; Nahar Sultana; Sujan Sarkar; Md. Abdur Razzaque; Mosaddek Hossain Kamal Tushar, Optimal Worker Selection for Maximizing Quality-of-Service of Online Food Delivery System, DOI: 10.1109/STI47673.2019.9068059, 16 April 2020

[21]. K.Sasikala Rani; S. Vishali, Blockchain driven IoT based Delish2Go Decentralized Food Delivery Application, DOI: 10.1109/ICAIS50930.2021.9395873, 12 April 2021

[22]. Ayyappa Chakravarthi M, Dr M. Thillaikarasi, Dr Bhanu Prakash Battula, published “Classification of Image Spam Using Convolution Neural Network” in International Information and Engineering Technology Association (IIETA) - “Traitement du Signal” Volume 39, No. 1

[23]. Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published “Classification of Social Media Text Spam Using VAE-CNN and LSTM Model” in International Information and Engineering Technology Association (IIETA)-Ingénierie des Systèmes d’Information Volume 25, No. 6.

[24]. Ayyappa Chakravarthi M, Dr. M. Thillaikarasi, Dr. Bhanu Praksh Battula, published a paper “Social Media Text Data Classification using Enhanced TF_IDF based Feature Classification using Naive Bayesian Classifier” in IJAST 2020

[25]. Ayyappa Chakravarthi M, Dr. M.Thillaikarasi, Dr. Bhanu Prakash Battula published “Social Media Data Classification – a Survey” in “Journal of Information and Computational Science” Volume 10, issue 3, March 2020s