DOI https://doi.org/10.30525/978-9934-26-597-6-68

DEVELOPMENT OF A MODEL FOR PREDICTING FOOD DELIVERY TIME BASED ON ORDER DATA

Yekaterina Kim[1*], Ilya Syssa[1], Yalkunzhan Yussupov[1], Kirill Shestakov[1]
[1]Turan University, 050013, Republic of Kazakhstan,
Almaty, Satpayev st., 16a
**Corresponding author's e-mail: e.kim@turan-edu.kz*
Received 12 March 2025, www.isma.lv

Abstract

This paper focuses on developing a model for predicting food delivery time based on order data. Machine learning methods are applied. The system architecture includes layers for data collection, processing, storage, and analysis, as well as integration with client interfaces. The developed model improves prediction accuracy, optimizes logistics, and enhances the user experience by providing precise delivery time estimates.

Keywords: machine learning, delivery time prediction, data analysis, logistics, forecasting models

1 Introduction

Modern web technologies and machine learning methods are actively used to solve various problems in logistics and delivery. One such task is predicting the time of order delivery, which is especially relevant given the rapid growth of e-commerce and online services. Developing accurate and efficient delivery time prediction systems is crucial for optimizing delivery processes and improving user convenience.

Delivery time prediction systems analyze data on transportation status, weather conditions, traffic congestion, and other factors affecting delivery speed. This enables more precise forecasts and highly accurate order arrival time predictions. Machine learning is one of the most effective tools for such predictions, as it is increasingly used to analyze large volumes of data and automate processes [1, 2].

The article presents the development of a delivery time prediction model using machine learning methods.

2 Literature review

In the growing food delivery market, predicting order completion time is crucial for improving service quality. The accuracy of prediction directly impacts customer satisfaction and company efficiency [1].

Modern delivery time prediction technologies rely on data analysis, machine learning, and integration with navigation systems [2, 3]. Large platforms such as Uber Eats and DoorDash use real-time data processing algorithms to determine optimal delivery times.

Machine learning algorithms help predict delivery time by analyzing historical data [3]. These technologies consider:

– Meal preparation time in the restaurant;

– Courier route time based on location;

– Route deviations due to weather conditions or traffic congestion.

An example is the use of neural networks for data analysis and automating the prediction process [3]. Delivery services integrate with navigation systems (Google Maps, Waze) to account for traffic conditions and dynamically adjust courier routes.

3 Architecture of the food delivery time prediction system

To build the food delivery time prediction system, an architecture was developed that effectively processes data, applies machine learning models, and provides real-time forecasts. This architecture includes several key components, integrated into a unified system (Figure 1):

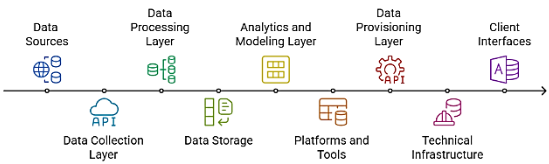

Figure 1. System architecture

I. Data Sources

Collecting data from various sources provides a foundation for analysis and forecasting:

1. Order Data – customer information, order content, order time.

2. Restaurants – meal preparation time, kitchen workload, ingredient availability.

3. Couriers – location, status, type of transport.

4. External Data – weather conditions, road conditions, time of day, day of the week.

II. System Architecture Layers

1. Data Collection Layer

– API Integration: Geolocation services (e.g., Google Maps, Yandex.Maps), weather services for obtaining weather data.

– Data Input Modules: Connection to restaurant and logistics databases, GPS data collection from couriers.

2. Data Processing Layer

– Data Cleaning and Normalization: Removing missing values, filtering outliers, standardizing data for model accuracy.

– Data Enrichment: Merging input data with additional information, such as traffic congestion forecasts or weather indices.

3. Data Storage

– Relational Databases: For structured data (PostgreSQL, MySQL).

– Big Data Storage: For historical and unstructured data (Hadoop, Amazon S3).

4. Machine Learning Model Layer

– Regression: For evaluating delivery time.

– Neural Networks: For identifying complex relationships between factors.

– Gradient Boosting: For analyzing large datasets and improving prediction accuracy.

5. Platforms

– TensorFlow, PyTorch – for model development and deployment.

– MLFlow or DVC – for tracking experiments.

6. Analytical Algorithms

– Route optimization, peak demand prediction.

7. Data Provisioning Layer (API)

– RESTful API or GraphQL for delivering prediction results to client applications.

– Integration with interfaces: Displaying estimated delivery time on user screens, push notifications about delivery status.

III. Technical Infrastructure

1. Server-side Infrastructure

– Cloud Platforms: Amazon Web Services (AWS), Google Cloud Platform (GCP), or Microsoft Azure for scalability and big data processing.

– Containerization: Docker and Kubernetes for deploying and managing microservices.

2. Client-side Infrastructure

– Mobile applications and web interfaces (Figure 2): Integration with interfaces to display delivery time predictions.

– Courier tracking modules: Maps showing real-time courier routes.

– Real-time data processing: Technologies such as Apache Kafka and Spark Streaming for processing real-time data, such as courier statuses and road conditions.

4 Machine learning model for delivery time prediction

To build and train the machine learning model for delivery time prediction, a dataset called Food Delivery Time Prediction Case Study.xlsx was used. This dataset includes various parameters influencing food delivery time.

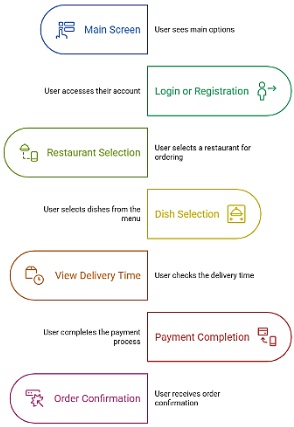

Figure 2. The order process

During data analysis, three key features were selected for prediction: weather, distance, and order time, as they directly affect delivery speed and can significantly improve prediction accuracy.

– Weather plays a critical role in delivery processes since adverse conditions (rain, snow, strong winds) can slow down movement and impact overall delivery time.

– Distance is an obvious factor – longer distances require more time for delivery.

– Order time is essential because different times of the day and peak hours influence delivery completion time.

A machine learning model was trained using Google Colab, employing various algorithms (Figure 3).

```
model = LinearRegression()
model.fit(X_train, y_train)
```

```
LinearRegression
LinearRegression()
```

```
y_pred = model.predict(X_test)

mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
print(f'R^2 Score: {r2}')
```

```
Mean Squared Error: 71.63221754013398
R^2 Score: 0.1830093542965805
```

Figure 3. Building and training a forecasting model

This approach enabled efficient data processing and pattern recognition based on selected features.

5 Conclusions

The developed machine learning-based model for food delivery time prediction demonstrates high accuracy, contributing to improved service quality and customer satisfaction. Integrating this model into the delivery system optimizes routes, considers dynamic factors, and reduces delays, making the process more transparent and efficient.

References

[1] Y. Geetha Sri, K. Vatsalya, G. Harshitha Devi, J. Sai Chandra. Precision delivery: machine learning algorithms for real-time food delivery estimates // Journal of Nonlinear Analysis and Optimization, 2025. – № 16(01). – P. 104-113. DOI: 10.36893/JNAO.2025.V16I01.013

[2] Abdul Muqtadir Mohammed. Machine learning-driven last-mile delivery time prediction using mgrs block analysis // International journal of computer engineering & technology, 2025. – № 16(1). – P. 2619-2630. – DOI: 10.34218/IJCET_16_01_185

[3] Sami Serkan Özarık, Paulo da Costa, Alexandre Florio. Machine Learning for Data-Driven Last-Mile Delivery Optimization // Transportation Science, 2023. – № 58(1). – DOI: 10.1287/trsc.2022.0029

Authors

Yekaterina Kim, 06.08.1977, Uzbekistan
Current position, grades: Cand.Tech.Sc., Associate Professor at the Higher School of Information Technology, Turan University
University studies: Kazakh National Technical University named after K.I. Satpayev **Scientific interest:** information technologies in education, computer modelling, machine learning, artificial intelligence
Publications (number or main): more than 100
Experience: Over 20 years of teaching at universities

Ilya Syssa, 25.11.2003, Almaty
Current position, grades: 3rd year student, GPA – 3.01
University studies: Turan University
Scientific interest: information technologies and engineering
Publications (number or main): -
Experience: participated in the abated university robotics competition developed work projects, front-end development of web applications and sites

PHOTO
OF
CONTRIBUTOR

Yalkunzhan Yussupov, 21.03.2003, Almaty
Current position, grades: 3rd year student, GPA – 2.5 **University studies:** Turan University **Scientific interest:** information technologies **Publications** (number or main): –
Experience:

PHOTO
OF
CONTRIBUTOR

Kirill Shestakov, 18.05.2004, Almaty
Current position, grades: 3rd year student, GPA – 2.23 **University studies:** Turan University **Scientific interest:** information technologies **Publications** (number or main): –
Experience: