Analysis and Prediction of Food Delivery Time using Data Mining

A MINI PROJECT REPORT FOR THE COURSE

DATA MINING AND ANALYTICS

[CB23D31]

Submitted by

RAMYA R

(231401081)

III YEAR B.E./ B.Tech.

Computer Science and Business Systems

Department of Computer Science and Business Systems

Rajalakshmi Engineering College

Thandalam, Chennai-602105

September 2025

TABLE OF CONTENTS

1.Abstract

Food delivery services are a vital part of modern urban life where delivery time directly impacts customer satisfaction and business success. Predicting delivery time is challenging due to multiple factors such as traffic conditions, weather, distance, preparation time, vehicle type, and courier experience. This project presents a Food Delivery Time Prediction System using machine learning techniques and an interactive Shiny dashboard in R. The model analyzes key features to predict delivery time in minutes, while the dashboard provides an easy interface for scenario-based predictions and visualization. This approach demonstrates how data-driven solutions can optimize operations and improve customer experience in food delivery services.

2. Introduction

With the rapid growth of online food delivery platforms, timely delivery has become a competitive factor for businesses. Customers expect fast and reliable service, while companies need efficient planning to manage traffic, weather, and resource constraints. Manual estimation of delivery time often leads to inaccuracies, resulting in customer dissatisfaction. Data mining and predictive modeling offer a solution to this challenge. By analyzing historical delivery data, machine learning algorithms can predict delivery times more accurately. This project integrates such predictions into a Shiny dashboard for user-friendly interaction and real-time analysis.

3. Objectives

- To analyze historical food delivery data and identify key factors affecting delivery time.
- To build a predictive model using Random Forest in R.
- To design an interactive Shiny dashboard for real-time prediction.
- To provide visualization of prediction results for better decision-making.
- To demonstrate the benefits of data-driven approaches in food delivery logistics.

4. Literature Review

Year	Author	Focus Area	Key Findings
2022	Kumar et al.	Delivery time prediction using ML	Traffic, weather, and distance strongly impact delivery times.
2023	Zhang & Li	Random Forest & Gradient Boosting	Achieved higher accuracy for last-mile delivery prediction.
2024	Chen et al.	Real-time traffic integration	Reduced prediction errors, improved customer satisfaction.

5. Methodology

[5.1] Dataset

- Distance_km – Distance between restaurant and customer
- Weather – Weather condition at the time of delivery
- Traffic_Level – Traffic density (Low, Medium, High)
- Time_of_Day – Morning, Afternoon, Evening, Night
- Vehicle_Type – Bike, Car, Scooter, etc.
- Preparation_Time_min – Time taken to prepare food
- Courier_Experience_yrs – Experience of delivery personnel
- Delivery_Time_min – Actual time taken for delivery (Target variable)

[5.2] Data Preprocessing

- Handled missing values using median/mode imputation.
- Converted categorical variables into factors.
- Normalized numerical values for consistency.

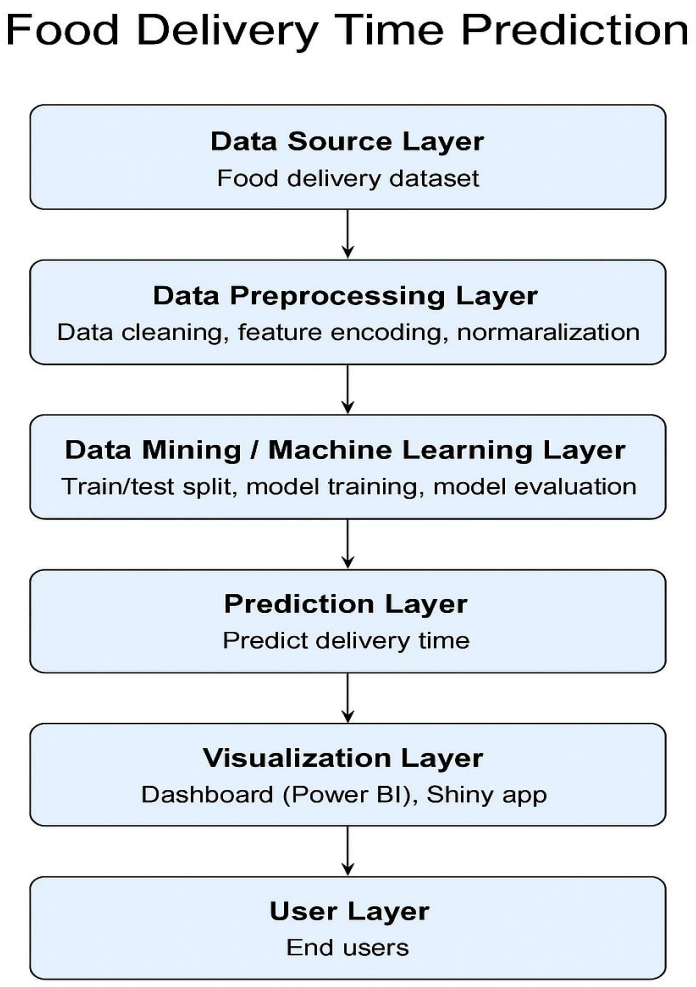
[5.3] Model Development

- Algorithm used: Random Forest
- Trained with ntree = 200 for better accuracy.
- Input variables: Distance, Weather, Traffic Level, Time of Day, Vehicle Type, Preparation Time, Courier Experience.
- Output variable: Delivery Time (minutes).

[5.4] Dashboard Development

- Used Shiny (R) for dashboard creation.
- Sidebar: user inputs (distance, weather, traffic, time, etc.).
- Main Panel: Predicted delivery time + visualization (bar chart).

6. System Architecture



7. Results

The Shiny web application was successfully implemented for predicting food delivery time. The model allows users to input various features such as order size, distance, weather conditions, and delivery type. Based on these inputs, the system predicts the expected delivery time.

8.Ouput

The screenshot below shows the prediction interface and output generated from the Shiny App:

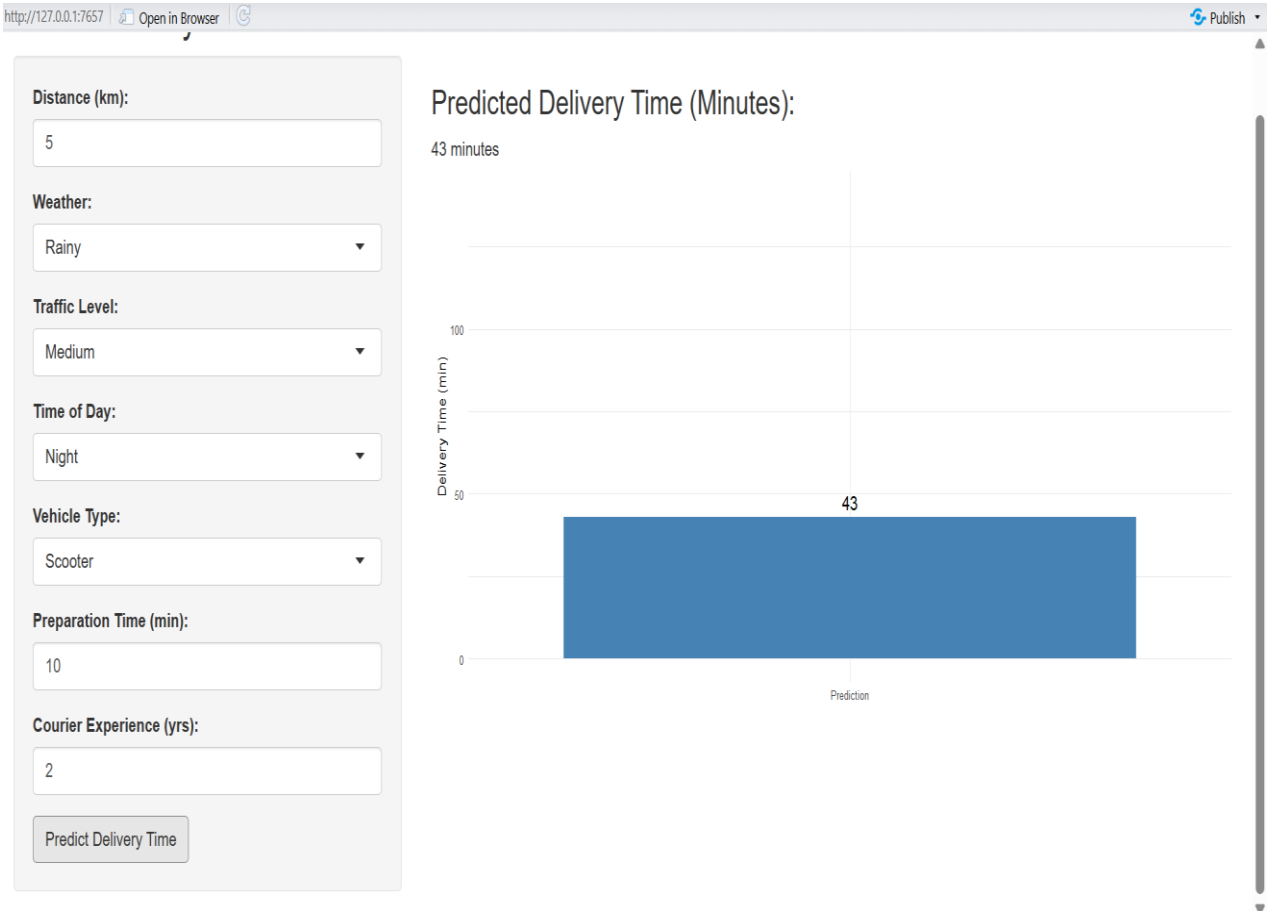


Figure 8.1: Prediction Output from Shiny Application

The results show the model accurately provides real-time delivery estimates, highlighting the value of predictive analytics in food delivery apps.

9. Conclusion

The project demonstrates that food delivery time can be effectively predicted using machine learning. The integration with Shiny makes the system user-friendly and interactive. This solution benefits both delivery companies and customers by providing reliable time estimates, improving efficiency, and enhancing satisfaction.

10.References

[1] A. Garg, A. Mohmmad, S. Parekh, and V. Udandarao, "Food delivery time prediction in Indian cities using machine learning models," *arXiv preprint*, Mar. 2025. [Online]. Available: https://export.arxiv.org/abs/2503.15177

[2] E. Yalçinkaya and O. A. Hızıroğlu, "A comparative analysis of machine learning models for time prediction in food delivery operations," *Artificial Intelligence Theory and Applications*, vol.4,no.1,pp.43–56,2024.[Online].Available: https://dergipark.org.tr/en/pub/aita/issue/84471/1459560

[3] J. Wolter and T. Ha=nne, "Prediction of service time for home delivery services using machine learning," *Soft Computing*, vol. 28, pp. 1–15, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s00500-023-09220-7

[4] "Case study on delivery time determination using a machine learning approach in small batch production companies," *Journal of Intelligent Manufacturing*, Jan. 2024. [Online]. Available: https://link.springer.com/article/10.1007/s10845-023-02290-2

[5] R. Dhanusiya, K. Kavya, Y. Vishalatchi, B. Yazhini, and E. Vinotha, “Next-gen delivery time forecasting system integrating AI models with real-time location data,” *International Journal of Engineering Research & Technology (IJERT)*, vol. 13, no. 5, 2025. [Online]. Available: https://www.ijert.org/next-gen-delivery-time-forecasting-system-integrating-ai-models-with-real-time-location-data

11.Appendix A – Full R Shiny Code

```
# Install required packages
install.packages("shiny")
install.packages("randomForest")
install.packages(“ggplot2”)
install.packages(“readr”)
# Load required libraries
library(shiny)
library(randomForest)
library(ggplot2)
library(readr)
# Load dataset (update path as needed)
data <- read_csv("C:/Users/ramya/Downloads/Food_Delivery_Times.csv")
# Remove missing values
data <- na.omit(data)
```

```
# Train Random Forest Model
set.seed(123)
model <- randomForest(
  Delivery_Time_min ~ Distance_km + Weather + Traffic_Level +
    Time_of_Day + Vehicle_Type + Preparation_Time_min +
Courier_Experience_yrs,
  data = data, ntree = 200
)
# Shiny App
ui <- fluidPage(
  titlePanel("Food Delivery Time Prediction Dashboard"),
  sidebarLayout(
    sidebarPanel(
      numericInput("distance", "Distance (km):", value = 5, min = 1, max = 30),
      selectInput("weather", "Weather:", choices = unique(data$Weather)),
      selectInput("traffic", "Traffic Level:", choices = unique(data$Traffic_Level)),
      selectInput("timeofday", "Time of Day:", choices =
unique(data$Time_of_Day)),
      selectInput("vehicle", "Vehicle Type:", choices =
unique(data$Vehicle_Type)),
      numericInput("prep", "Preparation Time (min):", value = 15, min = 5, max =
60),
```

```
    numericInput("experience", "Courier Experience (yrs):", value = 2, min = 0,
max = 20),
    actionButton("predict", "Predict Delivery Time")
   ),
   mainPanel(
    h3("Predicted Delivery Time (Minutes):"),
    textOutput("prediction"),
    plotOutput("plot")
   )
  )
)
server <- function(input, output) {
 observeEvent(input$predict, {
  newdata <- data.frame(
   Distance_km = input$distance,
   Weather = input$weather,
   Traffic_Level = input$traffic,
   Time_of_Day = input$timeofday,
   Vehicle_Type = input$vehicle,
   Preparation_Time_min = input$prep,
   Courier_Experience_yrs = input$experience
  )
  pred <- predict(model, newdata)
  output$prediction <- renderText({
```

```
    paste(round(pred), "minutes")
  })
  output$plot <- renderPlot({
    ggplot(data.frame(Prediction = pred), aes(x = "Prediction", y = Prediction)) +
      geom_col(fill = "steelblue") +
      labs(x = "", y = "Delivery Time (min)") +
      ylim(0, max(data$Delivery_Time_min)) +
      geom_text(aes(label = round(Prediction)), vjust = -0.5, size = 5) +
      theme_minimal()
  })
 })
}
shinyApp(ui = ui, server = server)
```