PREPARED BY hIMANSHU Kumar jHA JULY 10

report

Delivery Time Prediction

# Dataset Summary

Dataset Used: [[Food\_Delivery\_Times.csv](https://myoffice.accenture.com/:x:/r/personal/h_d_jha_accenture_com/Documents/model2.0/Food_Delivery_Times.csv?d=w3ad342ff0916494295ff52ac53d6b9e6&csf=1&web=1&e=HXd9l4)](https://myoffice.accenture.com/:x:/g/personal/h_d_jha_accenture_com/EdNDFhxPIrZCucFwwbGi3r0B5OkAP3P4S6G5tGmSIJuw4g?e=bw4pVd)

Initial Info

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 1000 entries, 0 to 999

Data columns (total 8 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Distance\_km 1000 non-null float64

1 Weather 970 non-null object

2 Traffic\_Level 970 non-null object

3 Time\_of\_Day 970 non-null object

4 Vehicle\_Type 1000 non-null object

5 Preparation\_Time\_min 1000 non-null int64

6 Courier\_Experience\_yrs 970 non-null float64

7 Delivery\_Time\_min 1000 non-null int64

dtypes: float64(2), int64(2), object(4)

memory usage: 62.6+ KB

**Categorical Columns:**

['Weather', 'Traffic\_Level', 'Time\_of\_Day', 'Vehicle\_Type']

**Numerical Columns:**

['Distance\_km', 'Preparation\_Time\_min', 'Courier\_Experience\_yrs']

# Data – PREPOCESSING

* **Missing Values:** Handled using SimpleImputer (mean for numerical, most frequent for categorical).
* **Feature Scaling:** Applied RobustScaler to numerical features.
* **Feature Selection:** Used SelectKBest with f\_regression to retain top 3 numerical features.
* **Encoding**: One-hot encoding for categorical variables using OneHotEncoder.

# MODEL TRAINING

Three regression models were evaluated:

* **Random Forest Regressor**
* **Gradient Boosting Regressor**
* **XGBoost Regressor**

# MODEL INFERENCES

| **Model** | **MAE** | **RMSE** | **R² Score** | **CV R² Mean** | **CV R² Std** |
| --- | --- | --- | --- | --- | --- |
| Random Forest | 6.85 | 9.58 | 0.795 | 0.696 | 0.053 |
| **Gradient Boosting** | **6.60** | **9.39** | **0.803** | **0.713** | **0.067** |
| XGBoost | 8.05 | 10.95 | 0.732 | 0.651 | 0.027 |

* **Best Model**: Gradient Boosting Regressor
* **Best Hyperparameters** (via GridSearchCV):
  + n\_estimators: 50
  + learning\_rate: 0.1
  + max\_depth: 3

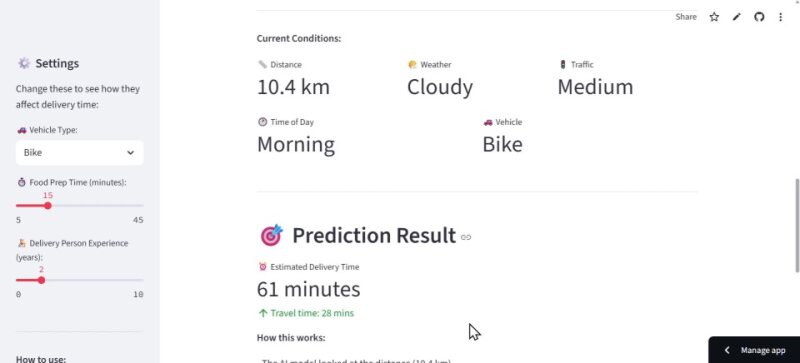
# DEPLOYMENT

* The best model was saved as best\_gradient\_boosting\_model.pkl using joblib.

**WEB APPLICATION: DELIVERY TIME PREDICTOR**

**DEMO** [**https://deliverytimepredictionmodel-v1.streamlit.app/**](https://deliverytimepredictionmodel-v1.streamlit.app/)

**VIDEO:**

**[](https://myoffice.accenture.com/personal/h_d_jha_accenture_com/Documents/model2.0/demo.mp4)**

**Built With**

* **Framework**: Streamlit
* **APIs Used**:
  + Google Maps Distance Matrix API (for real-time distance and traffic)
  + OpenWeatherMap API (for current weather conditions)
* **Features**
  + Interactive map to select delivery location
  + Sidebar to configure:
  + Vehicle type
  + Food preparation time
  + Courier experience
* Real-time data integration for:
  + Distance
  + Traffic level
  + Weather
  + Time of day
* Displays:
  + Estimated delivery time (via ML model)
  + Quick estimate (based on distance + prep time)
  + Comparison between both estimates
* **Demo Data**
  + A synthetic dataset of delivery pickup location was generated to simulate predictions.
  + The model successfully predicted delivery times for various combinations of inputs, demonstrating its robustness and adaptability.