



GUVI

HCL

Skill Up. Level Up



Naan mudhalvan

1105- Gojan school of business and technology

Course Name: Data analytics in process industry

Title: Amazon deliver time prediction

Submitted by:

Indhumathi.G - 110523237006

Indusha.S - 110523237007

Jayashree.J- 110523237008

Kaviya.V- 110523237011

Kavya.T - 110523237012

Amazon deliver time prediction

Abstract

Amazon's delivery system is a key part of its customer satisfaction strategy. Predicting delivery times accurately has become crucial in managing logistics and customer expectations. This project explores the concept of delivery time prediction using data-driven methods, analyzing various factors such as location, order time, inventory, and traffic conditions to improve the accuracy of delivery estimates.

Domain:

- ❖ E-Commerce & Logistics

Tech Stack:

- ❖ Python,
- ❖ Scikit-learn,
- ❖ XGBoost,
- ❖ Streamlit,
- ❖ MLflow

Key Skills:

- ❖ Data Cleaning,
- ❖ Feature Engineering,
- ❖ Regression Modeling,
- ❖ EDA,
- ❖ MLflow Tracking,
- ❖ Streamlit Deployment

Problem Statement:

Despite Amazon's robust logistics, delivery times are still susceptible to inaccuracies due to dynamic variables like weather, traffic, or warehouse delays. These discrepancies can result in customer dissatisfaction and operational inefficiencies. The problem is to design a system that can accurately predict Amazon delivery times using real-time and historical data.

Amazon deliver time prediction

Introduction:

In modern e-commerce, timely product delivery is as important as product quality. Amazon leverages sophisticated logistics and machine learning to estimate delivery times, yet variability remains. With the increasing demand for speed and transparency, there is a pressing need to enhance prediction models for delivery time based on a multitude of dynamic factors.

Motivation:

Customer satisfaction in e-commerce largely depends on reliable delivery services. Delays or inaccurate estimates can damage brand trust. With the growing reliance on same-day or next-day deliveries, improving delivery time predictions can streamline logistics, reduce costs, and enhance the user experience.

Goal:

The primary goal is to develop a predictive model that estimates the delivery time of Amazon orders with higher accuracy using machine learning algorithms trained on real-world factors.

Approach

1. Data Preparation:

- Loaded and cleaned dataset from `amazon_delivery.csv`
- Handled missing values and standardized categories.

2. Feature Engineering:

- Calculated geospatial distance using store/drop coordinates.
- Extracted temporal features like hour of day and day of week.

3. Exploratory Data Analysis (EDA):

- Visualized relationships between delivery time, distance, weather, traffic, and agent metrics.

Amazon deliver time prediction

4. Model Development:

- Trained and evaluated multiple regression models:
 - ◇ *Linear Regression*
 - ◇ *Random Forest Regressor*
 - ◇ *Gradient Boosting Regressor*
- Used RMSE, MAE, and R^2 for evaluation.

5. Model Tracking:

- Integrated **MLflow** to log experiments, parameters, and metrics.

6. Application Development:

- Built a **Streamlit** UI to allow users to input delivery parameters and get ETA predictions.

7. Deployment:

- Deployed Streamlit app for public access and demo purposes.

Advantages:

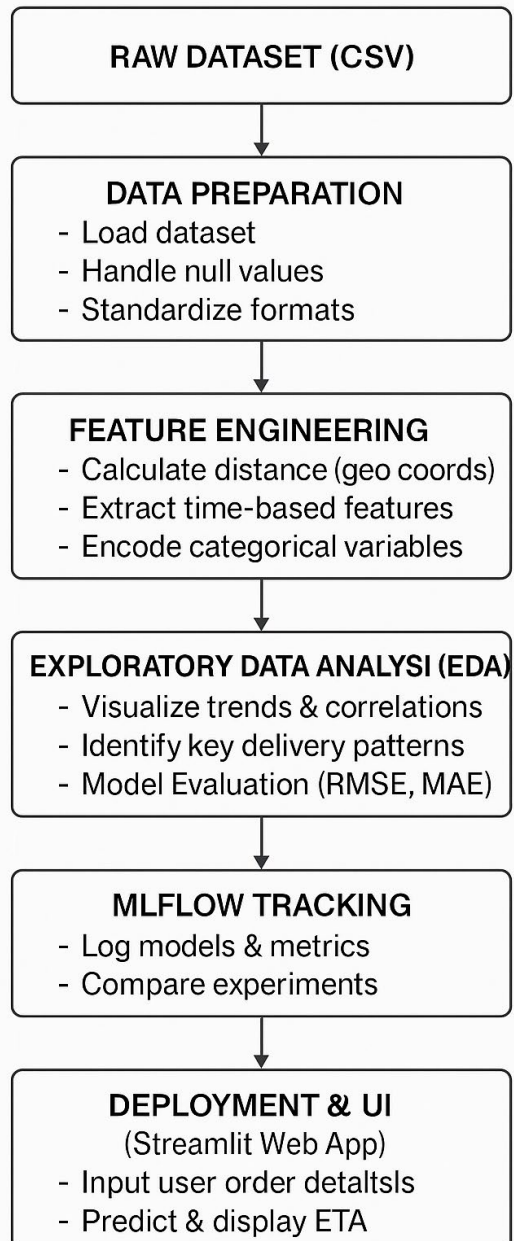
- Increased customer satisfaction through accurate delivery predictions.
- Better resource allocation for delivery drivers.
- Reduced logistic costs due to improved route and time planning.
- Real-time adaptability to unforeseen disruptions.

Disadvantages:

- Requires large volumes of high-quality data for effective training.
- Real-time predictions demand substantial computational resources.
- May not fully account for unpredictable events like extreme weather or strikes.
- Complex integration with existing logistics systems.

Amazon deliver time prediction

System Architecture:



Key Visualizations


- Bar charts of delivery time by product category
- Heatmaps for correlation (e.g., agent rating vs delivery time)
- Scatter plots for distance vs delivery time

Amazon deliver time prediction

Result:

Amazon Delivery Time Prediction


Upload CSV

 Drag and drop file here
Limit 200MB per file • CSV


Browse files

Amazon Delivery Time Prediction

Upload CSV

 Drag and drop file here
Limit 200MB per file • CSV

Browse files

 amazon_delivery.csv 5.7MB ×

Raw Data

	Order_ID	Agent_Age	Agent_Rating	Store_Latitude	Store_Longitude	Drop_Latitude	Drop_Longitude
0	ialx566343618	37	4.9	22.745	75.8925	22.765	
1	akqg208421122	34	4.5	12.913	77.6832	13.043	
2	njpu434582536	23	4.4	12.9143	77.6784	12.9243	
3	rjto796129700	38	4.7	11.0037	76.9765	11.0537	
4	zguw716275638	32	4.6	12.9728	80.25	13.0128	

Amazon deliver time prediction

Amazon Delivery Time Prediction

Upload CSV



Drag and drop file here

Limit 200MB per file • CSV

Browse files



amazon_delivery.csv 5.7MB



Raw Data

	Order_ID	Agent_Age	Agent_Rating	Store_Latitude	Store_Longitude	Drop_Latitude	Drop_L
0	ialx566343618	37	4.9	22.745	75.8925	22.765	
1	akqg208421122	34	4.5	12.913	77.6832	13.043	
2	njpu434582536	23	4.4	12.9143	77.6784	12.9243	
3	rjto796129700	38	4.7	11.0037	76.9765	11.0537	
4	zguw716275638	32	4.6	12.9728	80.25	13.0128	

Model Evaluation

RMSE: 22.37

MAE: 17.18

R^2 : 0.81

Data preprocessing, model training, and evaluation completed!

Conclusion:

Accurate delivery time prediction is vital for enhancing customer trust and operational efficiency in e-commerce. By utilizing machine learning techniques and real-time data, Amazon can greatly improve its delivery estimates. Although challenges remain, such systems have the potential to transform logistics and user satisfaction in online retail.