AI-Powered Delivery Time Prediction System

Machine Learning Solution for Food Delivery Optimization

HACKATHON PROJECT

TEAM ALPHA

NOVEMBER 2023

Problem Statement & Business Impact

Inaccurate Estimates

Traditional delivery time predictions are often static and fail to account for dynamic real-world variables, leading to customer dissatisfaction.

Operational Challenges

Inefficient resource allocation, suboptimal route planning, and increased operational costs stem from unreliable delivery time forecasts.

The Solution: ML Regression

Develop a data-driven machine learning regression model to accurately predict delivery times in minutes, enhancing overall efficiency.

Our primary objective is to predict delivery time with high accuracy, directly impacting customer satisfaction and optimizing logistics.

Dataset Overview & Feature Characteristics

Our model leverages a comprehensive dataset comprising crucial information to predict delivery times effectively. Understanding the nature of these features is paramount for building robust predictive models.

- **Delivery Partner:** Age, Ratings
- **Location Data:** Restaurant and Delivery Coordinates (Latitude/Longitude)
- **Order Details:** Type of Order, Type of Vehicle
- **Target Variable:** Delivery Time (in minutes)

The dataset underwent rigorous cleaning, ensuring no missing values or duplicates, which is vital for model integrity.

Delivery_person_Age	Numerical
Delivery_person_Ratings	Numerical
Restaurant_latitude	Numerical
Restaurant_longitude	Numerical
Delivery_location_latitude	Numerical
Delivery_location_longitude	Numerical
Type_of_order	Categorical
Type_of_vehicle	Categorical

Exploratory Data Analysis: Key Insights

- **Delivery Time Distribution**

  Delivery times largely follow a normal distribution, with a significant peak observed between 20-30 minutes, indicating a common range for most deliveries.

- **Distance Correlation**

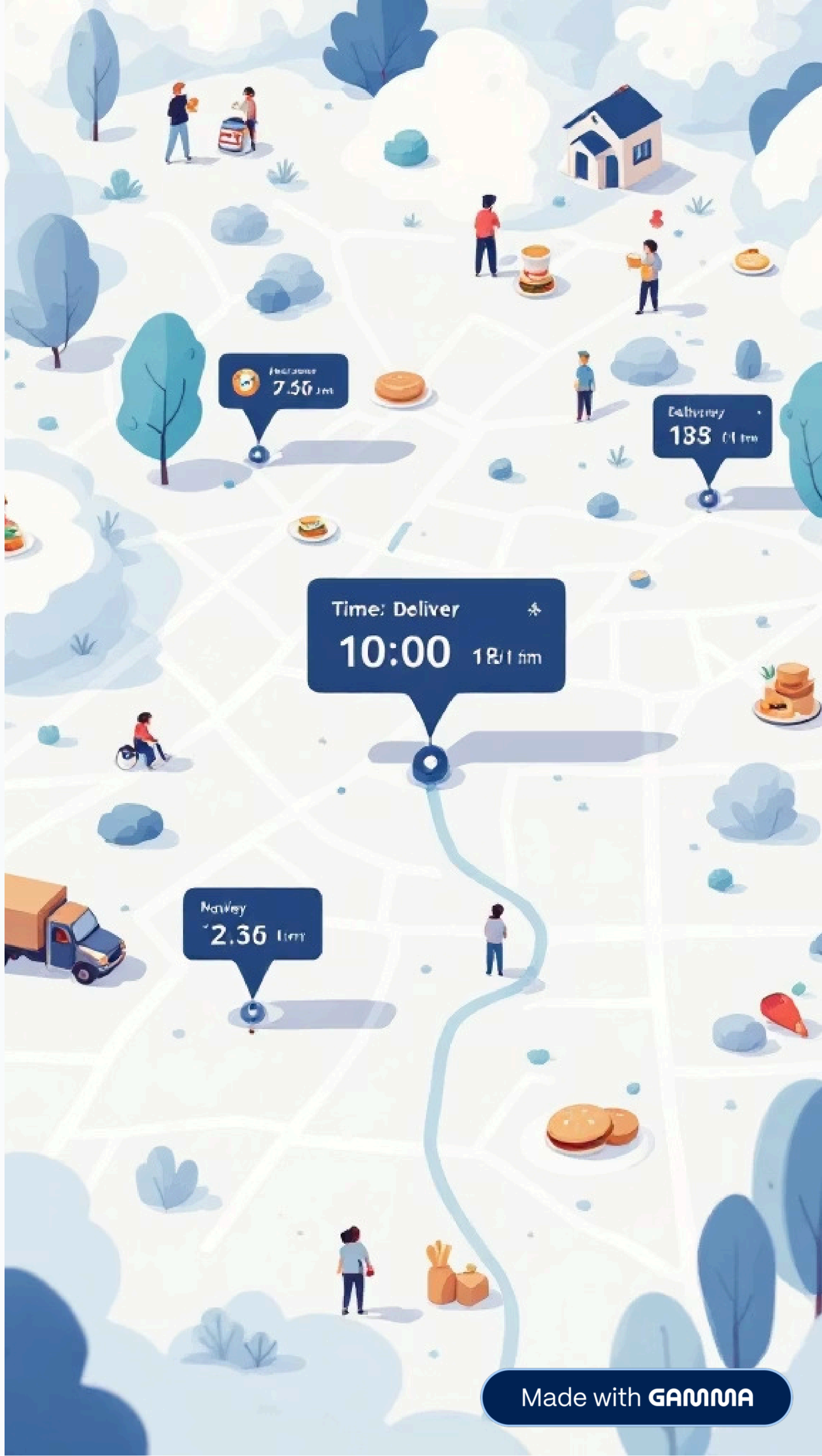
  A strong positive correlation exists between the delivery distance and the time taken, highlighting distance as a critical predictive factor.

- **Ratings Impact**

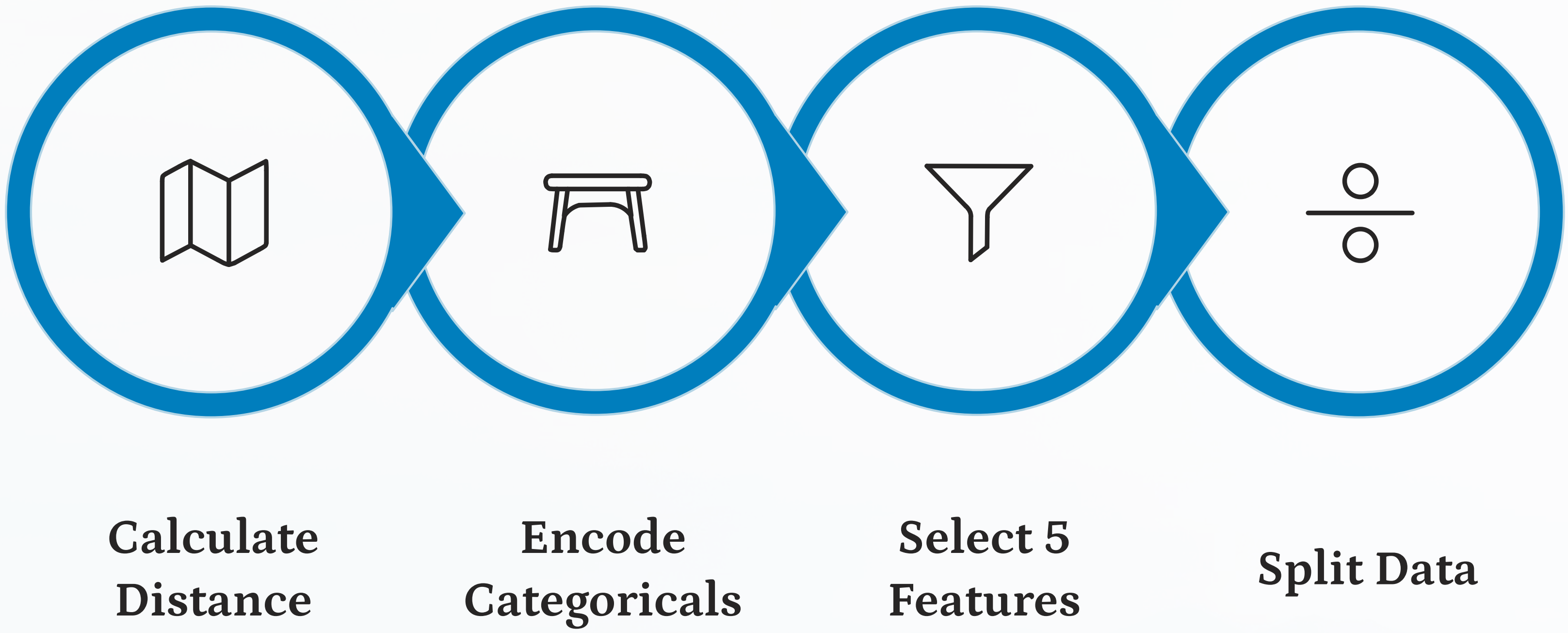
  Delivery partners with higher ratings tend to complete deliveries faster, suggesting efficiency directly links to customer satisfaction scores.

- **Vehicle Type Variations**

  Different vehicle types exhibit consistent median delivery times, but also reveal specific outlier patterns that need further investigation.



Feature Engineering & Preprocessing



Transforming raw data into meaningful features is crucial for model performance. Our meticulous approach ensures the model learns from the most relevant information.

Calculated Distance

Using the Haversine formula, we accurately computed `distance_km` from restaurant and delivery coordinates.

Categorical Encoding

Categorical variables like `Type_of_order` and `Type_of_vehicle` were converted into numerical representations.

Final Features

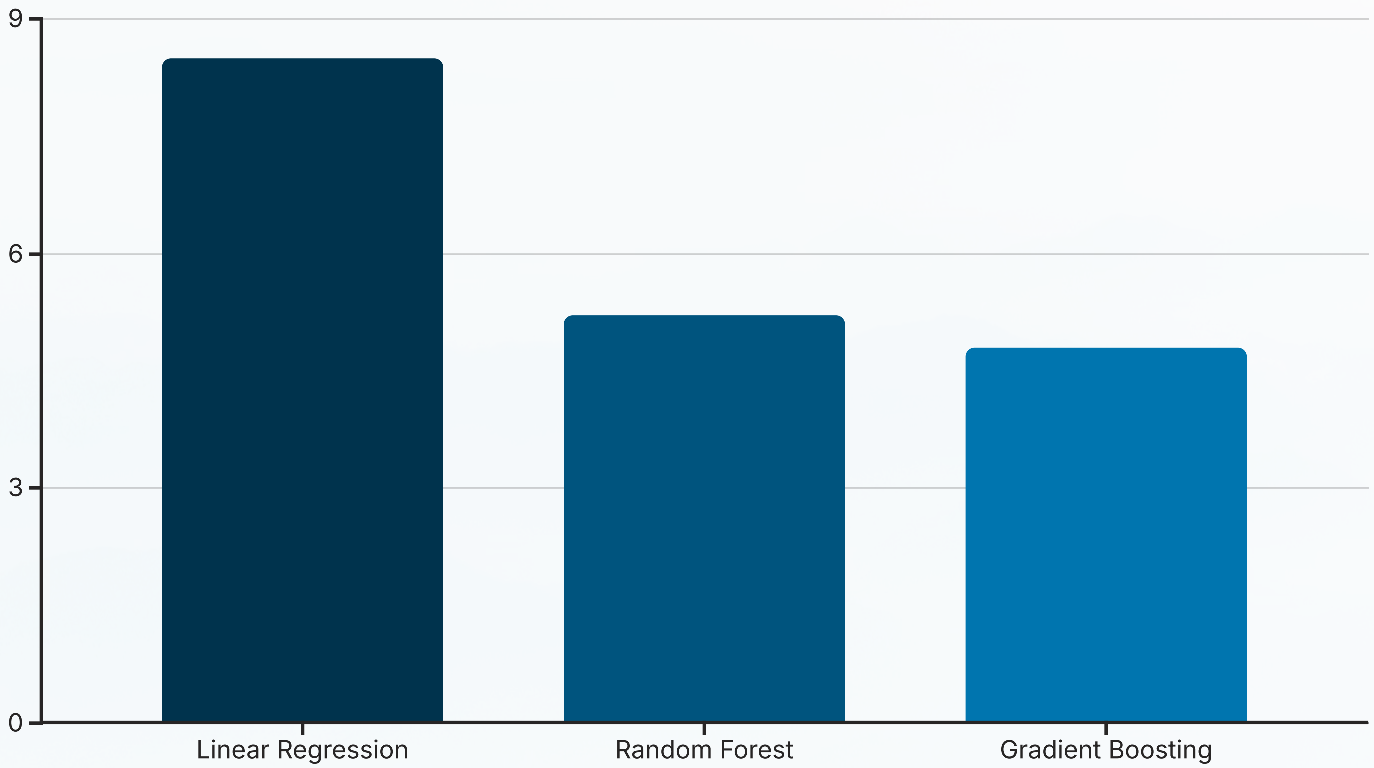
Selected 5 key features: `Delivery_person_Age`, `Delivery_person_Ratings`, `distance_km`, `Type_of_order_encoded`, `Type_of_vehicle_encoded`.

Train-Test Split

The dataset was split into an 80-20 ratio for training and testing, ensuring robust model validation.

Model Development & Results

After rigorous experimentation, Gradient Boosting emerged as the superior model, balancing accuracy and efficiency.



Comparison of Root Mean Squared Error (RMSE) across different models.

Models Explored

- Linear Regression (Baseline)
- Random Forest Regressor
- Gradient Boosting Regressor

Key Metrics

Evaluated models using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), focusing on minimizing prediction errors.

Best Performer

Gradient Boosting Regressor delivered the lowest RMSE, proving to be the most accurate model for our prediction task.

Top Predictors

`distance_km` and `Delivery_person_Ratings` were identified as the most influential features.

Deployment & User Interface

Our solution culminates in an intuitive web application, making real-time delivery predictions accessible and actionable.

Interactive Web App

A user-friendly interface built with React.js frontend and a Python ML backend for seamless interaction.

Real-time Prediction

Instantly predict delivery times using dynamic inputs, offering immediate value to users.

Intuitive Design

Features easy-to-use sliders and dropdowns, with instant results delivered via email for convenience.

Production Ready

The trained model is saved as a `.pkl` file, ensuring scalability and ease of deployment in production environments.

Conclusion & Future Scope

Key Achievements

- Successfully developed an end-to-end ML pipeline from EDA to deployment.
- Achieved high accuracy with the Gradient Boosting model for delivery time prediction.
- Delivered a production-ready user interface for practical application.
- Improved customer experience and operational efficiency are direct benefits.

Future Enhancements

- **Traffic Data Integration:** Incorporate real-time traffic information for more precise predictions.
- **Weather Conditions:** Add weather features to account for environmental impacts on delivery times.
- **Batch Prediction:** Enable predictions for multiple orders simultaneously to support logistics planning.
- **Mobile App Integration:** Extend the solution to a dedicated mobile application for on-the-go access.

Our project lays a strong foundation for optimizing food delivery services, with exciting avenues for future development.