Food Delivery Time Prediction

Objective

The goal is to predict whether a food delivery will be "Fast" or "Delayed" based on features like customer location, restaurant location, weather, traffic conditions, and more. This dataset can be used to explore **clustering and neural network models** for predictive analytics.

Phase 1 - Data Preprocessing and Feature Engineering

- **Data Import and Cleaning:**
    - Load the dataset (`Food_Delivery_Time_Prediction.csv`).
    - Handle missing values through imputation or removal.
    - Encode categorical variables (e.g., weather, traffic conditions) using One-Hot Encoding or Label Encoding.
    - Normalize numerical features (e.g., Distance, Delivery Time).
- **Feature Engineering:**
    - Calculate geographical distance using the Haversine formula if not provided.
    - Derive time-related features such as "rush hour" or "non-rush hour".

Phase 2 - Clustering using K-Means and Hierarchical Clustering

1. **K-Means Clustering**
    - **Objective** Use K-Means Clustering to group similar delivery times and characteristics based on features such as customer and restaurant locations, weather, traffic conditions, and more.
    - **Implementation**
        - Apply the K-Means algorithm to segment data into clusters based on delivery time patterns.
        - Use **Elbow Method** to determine the optimal number of clusters.
    - **Evaluation:**
        - Visualize clusters using 2D or 3D scatter plots.
        - Analyze how delivery times are influenced by different clusters.
2. **Hierarchical Clustering**
    - **Objective** Perform hierarchical clustering to understand how the features like traffic, weather, and location group together in a hierarchical structure.
    - **Implementation**
        - Apply Agglomerative Clustering to the dataset.
        - Use dendrograms to visualize the clusters and decide on the optimal number of clusters.
    - **Evaluation**
        - Visualize the hierarchy using a dendrogram.
        - Compare cluster analysis with K-Means results to find patterns.

Phase 3 - Neural Networks for Prediction

1. **Introduction to Neural Networks**
    - **Objective** Predict if a delivery will be "Fast" or "Delayed" using a neural network model.
    - **Implementation**
        - Build a simple feedforward neural network using Keras/TensorFlow.
        - Input features: Distance, Traffic, Weather, Order Priority, etc.
        - Output: "Fast" (0) or "Delayed" (1).

    - **Evaluation Metrics** Accuracy, Precision, Recall, F1-score.
2. **Model Improvement**
    - Tune hyperparameters such as the number of layers, number of neurons, activation functions, and learning rate.
    - Evaluate the neural network's performance compared to traditional machine learning models (e.g., Logistic Regression).

Phase 4 - Reporting and Insights

- **Model Comparison**
    - Compare the clustering results and neural network accuracy to determine the most effective method for delivery time prediction.
    - Discuss the insights derived from K-Means and Hierarchical Clustering and how they relate to improving delivery prediction.
- **Actionable Insights**
    - Based on clustering and neural network predictions, provide recommendations on improving food delivery times, optimizing delivery routes, and better managing resources.

Final Deliverables

1. **Jupyter Notebook (.ipynb)**
    - Full code for data preprocessing, clustering, and neural network modeling.
2. **Data Visualizations**
    - Visualizations such as Elbow Method graphs, dendrograms, `cluster scatter plots,` and neural network performance metrics.
3. **Final Report**
    - A comprehensive report summarizing the methodology, model evaluations, key findings, and actionable recommendations.
