ELITE TECHNO® GROUPS

Project – Predict the City Taxi Trip Duration

Challenge

Business Goal: To improve the efficiency of electronic taxi dispatching systems it is important to be able to predict how long a driver will have his taxi occupied. If a dispatcher knew approximately when a taxi driver would be ending their current ride, they would be better able to identify which driver to assign to each pickup request.

ML Goal: To build a model that predicts the total ride duration of taxi trips in New York City

Data: The taxi trip records include fields capturing pick-up and drop-off dates/times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types, and driver-reported passenger counts. The data used in the attached datasets were collected and provided to the NYC Taxi and Limousine Commission (TLC) by technology providers authorized under the Taxicab

About Data

id	a unique identifier for each trip
vendor_id	a code indicating the provider associated with the trip record
pickup_datetime	date and time when the meter was engaged
dropoff_datetime	date and time when the meter was disengaged
passenger_count	the number of passengers in the vehicle (driver entered value)
pickup_longitude	the longitude where the meter was engaged
pickup_latitude	the latitude where the meter was engaged
dropoff_longitude	the longitude where the meter was disengaged
dropoff_latitude	the latitude where the meter was disengaged
store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor because the vehicle did not have a connection to the server Y=store and forward; N=not a store and forward trip
trip_duration	duration of the trip in seconds

Project Tasks

- Data Gathering
- Data pre-processing
  - Data Cleaning
  - Missing value analysis
  - Exploratory Data Analysis
  - Feature Creation
- Train Model
- Test Model

Project Submission

1. Create a presentation documenting all the project steps
2. Presentation Template can be found in GitHub Repository
3. You are encouraged to use your own templates but ensure all the steps are clearly explained
4. Attach the presentation using the Google form in next section

Thank You!

Keep learning Keep exploring