

Problem Description :-

In the Global Entrepreneurship Summit, Chh-OLA, the taxi-hailing startup led by Khanchandani, Harsh and Dwivedi garnered a lot of interest and funding from the investors. Five years later, Chh-OLA is successfully running a number of taxis on the streets of New Delhi and garnering significant profit.

Now, Chh-OLA has opened a Data Science division recently and you have been recruited into the division. As part of your training, Shruti, the head of the Data Science division, has given you a task of estimating total fare amount of trips using various trip parameters such as distance, passenger count etc.

Can you fulfil this task given by your head?

Data Description :-

The data contains information about 1.4 million trips in Chh-OLA cabs, in the form of 17 columns, each having specific information about the respective trip.

File descriptions :-

- train.csv - the training set
- test.csv - the test set
- sampleSubmission.csv - a sample submission file in the correct format

Columns Description :-

- ID - Unique Identifier
- vendor_id - Taxi data providing vendor; 1 = TaxiTech Inc. 2 = DataCollectors Inc.
- pickup_loc - Location ID from where passenger was picked up
- drop_loc - Location ID where passenger was dropped
- driver_tip - Tip given to driver
- mta_tax - Automatically triggered tax amount
- distance - Distance covered in the trip
- pickup_time - Date/Time when meter started
- drop_time - Date/Time when meter stopped
- num_passengers - Cab passenger count
- toll_Amt - Toll paid in the booths
- payment_Method - Method of payment symbolised by a numeric code (1 = Credit Card, 2 = Cash, 3 = Free ride, 4 = Disputed, 5 = Unknown, 6 = Void trip)
- rate_code - Rate code for the trip (1 = Standard, 2 = Airport, 3 = Connaught Place, 4 = Noida, 5 = Negotiated Fare, 6 = Pooled ride)
- stored_flag - Flag which signifies whether trip data was immediately sent to Chh-OLA's database or not (Y=Yes, N=No, because of connection error)
- extra_charges - Miscellaneous charges
- improvement_charge - Charge levied for improvement in infrastructure
- total_amount - Output label; Final amount to be paid including meter fare and all extra charges

