

Intracity Fare Estimation

Problem Introduction :

Chennai (formerly known as Madras) , located on the Coromandel Coast off the Bay of Bengal . People from all over the world come to the marvelous here to spend their holidays, enjoy the natural splendor and to collect unforgettable memories. Recently, Devesh Sethia went to Chennai to met his friend Kartikay Singh , who is a reputed minister of Ministry of Transport and Highways , India . While he was there , many times he was being cheated by drivers . He asked his friend Kartikay to look into the matter and suggested him to declare a fare rate for different modes of transports in a city. He liked his idea . But analyzing these fares was difficult task for him. For this , he called one of his college friend , Mohammad Ausaf Jafri who has deep knowledge in Machine Learning and provided him dataset containing 20k samples of data for intracity fares in different cities of India . He got a month period to come up with the best possible solution . But these days , he is so busy in his office works . He is not getting much time to spend on analysing the dataset .So , he need help from others . Help him to solve the problem .

Data Description :

We are providing a training dataset(of 20k samples) describing trips/samples for past 2 years (from 01/01/2015 to 31/12/2016) for different modes of transport in different cities of India . Each data sample corresponds to one complete trip . it contains 11 features as follows :-

1. **ID(Integer)** : It is a unique identifier for different samples .
2. **TIMESTAMP(Datetime)** : It is trip start time . It's format is like (Year) - (Month) - (Day) (Hours) : (Minutes) : (Seconds) .
3. **STARTING_LATITUDE(Float)** : It is trip start time position's latitude in degree North .
4. **STARTING_LONGITUDE(Float)** : It is trip start time position's longitude in degree East .
5. **DESTINATION_LATITUDE(Float)** : It is trip stop time position's latitude in degree North.
6. **DESTINATION_LONGITUDE(Float)** : It is trip stop time position's longitude in degree East.
7. **VEHICLE_TYPE(String)** : It tells different transport vehicle type used for the trip .
8. **TOTAL_LUGGAGE_WEIGHT(Float)** : It is total luggage carried by the passenger in kilograms .
9. **WAIT_TIME(Float)** : It is the time for which driver waited for the passenger before start of the trip in minutes .
10. **TRAFFIC_STUCK_TIME(Integer)** : It is the time for which vehicle waited in traffic in minutes .
11. **DISTANCE(Integer)** : It is total distance covered in a trip in kilometres .

Moreover , file contains TARGET VARIABLE - FARE(Integer) which tells complete trip cost .

File Descriptions :

intracity_fare_train.csv - The training dataset .

intracity_fare_test.csv - The test set .

intracity_fare_sample_submission.csv - A sample submission file .