The Bengaluru Last Mile Challenge, 2025

Detailed Problem Description

Problem 1: Given the live location of buses of a set of routes up to time *t* on a particular day, predict the arrival time of these buses at various bus stops along their routes up to the end of the route.

You will be given a series of data packets consisting of locations and time stamps of a bus from the starting point of the route up to time *t* on a particular day. The task is to predict the arrival time of the bus at all future stops. The training data will consist of historical data packets for a large number of trips during a period of three weeks. Each data packet will contain information about the latitude and longitude, time stamp, and the bus route information. It should be noted that there may be several trips corresponding to a given bus route on any given day. Also, in some cases, the data packet may have null or missing values. It is also recommended that some of the training data be kept aside for validation purposes. The submitted solution will be evaluated on the basis of how well it can predict the arrival times at future bus stops of a set of buses at different times of the day on different days. The evaluation metric will give progressively higher scores for accurate prediction of arrival times at bus stops that are farther away, i.e., the farther the bus stop, the higher the score for the correct prediction of the arrival time. The exact scoring formula will be announced separately.

Problem 2: Predict the time required to travel between specified pairs of locations in an autorickshaw for departures falling between time *t* and time *t* +1 hour.

You will be given autorickshaw availability information at different times of the day (at a 15-minute granularity) in different parts of the city based on a hexagonal grid pattern that can be overlaid on a road network map Bengaluru. The information will include the number of trips that started inside the hexagon as well as the number of trips that ended inside the hexagon within 15-minute intervals over a 24-hour period. This information is computed based on the ridership data provided by Namma Yatri for a period of two weeks. In addition, you will also be provided information on the average autorickshaw speed on various Open Street Map (OSM) road segments of the city during 1-hour intervals over a 24 hour period. Roads in OSM are represented as segments, and each of them has an ID. This ID will be used to provide relevant information about that particular road segment. This information is again provided by Namma Yatri based on historical travel times. You can make use of the OSMnx package to superimpose this information on the road network of Bengaluru. This package will also allow you to find a path between the origin and destination for which you need to estimate the travel time. Optionally, you may use other approaches as well. The path will be a sequence of OSM IDs corresponding to the road segments that will be used in the journey. Based on this information, you are expected to estimate the total travel time from the point a person requests an autorickshaw ride on the app until the destination is reached. This will include the request acceptance time, the pickup time (during which the passenger needs to wait for the autorickshaw to pick him/her up), and the travel time. The submitted entries will be judged on the basis of how well the predicted times match actual times on the basis of actual passenger trips that would have taken place. The exact scoring formula will be announced separately.

Context about OSM and OSMnx:

- **OpenStreetMap (OSM):** This is a free, community-driven map of the world that has details about roads, intersections, and places. All of these have OSM IDs. In this hackathon, the

speed details are provided based on the OSM IDs corresponding to road segments of the road network of Bengaluru.

- **OSMnx:** This is a Python library that makes it easy to use OSM data. With OSMnx, you can download the road network of a city, visualize it, and run routing algorithms (like finding the shortest or fastest path). For this challenge, it helps you match autorickshaw data with real roads and estimate travel times.

Problem 3: Predict the total time required to complete a multi-modal journey, which may use a combination of bus and autorickshaw travel.

In this task, you are expected to combine the results of Problem 1 and 2 and estimate the most likely bi-modal journey time of a trip in which the initial leg is an autorickshaw ride, the middle leg is a bus ride, and the final leg is again an autorickshaw ride. The journey time should factor in the waiting times at each point and also the probability of missing a bus so that you may need to take the next one. The entries will be judged on the basis of how well the predicted times match actual times as obtained from passengers that would have taken the actual trips. The exact scoring formula will be announced separately.