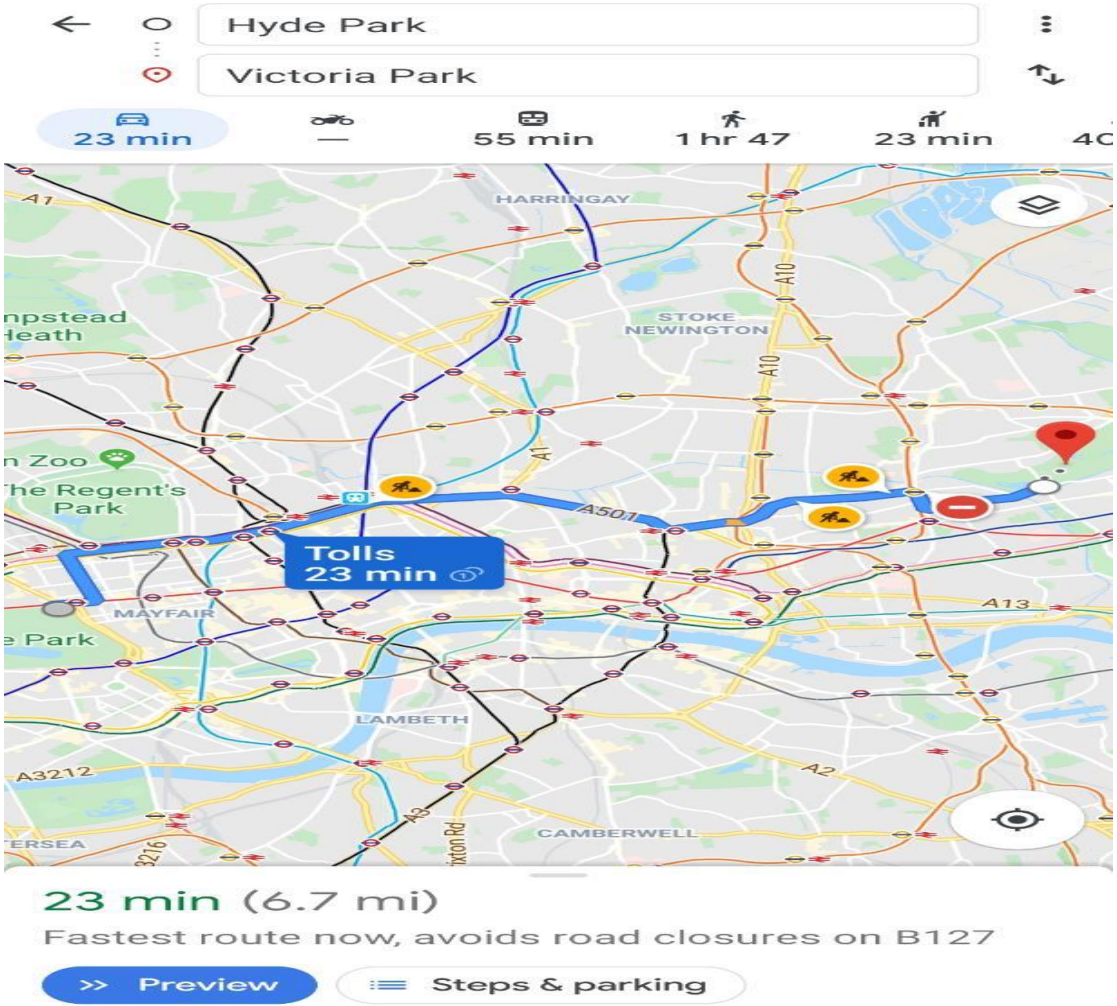
BeWayre

Team Random

Best way forward

We have created a Webpage for finding the safest route, by creating models that use various causes of accidents such as type of vehicle, weather conditions, street light availability and time of day to predict the best possible way in terms of time as well as safety.


Hyde Park
Victoria Park
23 min
55 min
1 hr 47
23 min
HARRINGAY
STOKE NEWINGTON
A10
A1
A501
Tolls
23 min
MAYFAIR
LAMBETH
CAMBERWELL
A13
A2
A3212
A3
23 min (6.7 mi)
Fastest route now, avoids road closures on B127
Preview
Steps & parking

1. Objective

How and What?

- ➔ **Safest Path**
  Allow users to input an origin and a destination and find the safest driving route that connects the two.
- ➔ **Car Accident Risk Prediction**
  Predict the accident probability and make clusters for Accident hotspots & Crime hotspots
- ➔ **Assigning weights for Route planning using Here API**

Features used for Route selection

- Weather
- Vehicle crime spots
- Accident HotSpots
- Junctions
- Real Time Traffic Data
- Time of travel

Data Selection

- Possible paths from Here API: https://route.cit.api.here.com/routing/7.2/calculateroute.json?waypoint0=52.5214,13.4155&waypoint1=52.5253,13.3693&mode=fastest;car&alternatives=4&app_id=FnxPXB5eCeTGeFJYZjGO&app_code=U2odxl9gB5Zv6EZTMw30nvGQzG9B39C1D6h-Xzarg4M
- Road Transport Accidents Data from Kaggle: https://www.kaggle.com/akshay4/road-accidents-incidence
- Real time Weather data from Openweathermap: http://api.openweathermap.org/data/2.5/weather?q=Canberra&APPID=fd88f5a3b637e917d5ed48d9d6b53335
- Vehicle Crime Data from UK Police dataset: https://data.police.uk/data/
- Real Time Traffic Incident Data: https://traffic.ls.hereapi.com/traffic/6.0/incidents.json?corridor=51.5072%2C-0.1275%3B51.50781%2C-0.13112%3B51.51006%2C-0.1346%3B1000&apiKey=U2odxl9gB5Zv6EZTMw30nvGQzG9B39C1D6h-Xzarg4M

FEATURES FOR ML MODELS

- Weather (temperature, wind speed, visibility, rainy/snowy/icy, snow depth, rain depth, etc)
- Time features: hour of day, day of the week, month of the year
- Static features such as speed limit, junction details, proximity to intersections, road surface type, etc
- Location features: Latitude, Longitude, District
- Vehicle Crime

Clustering

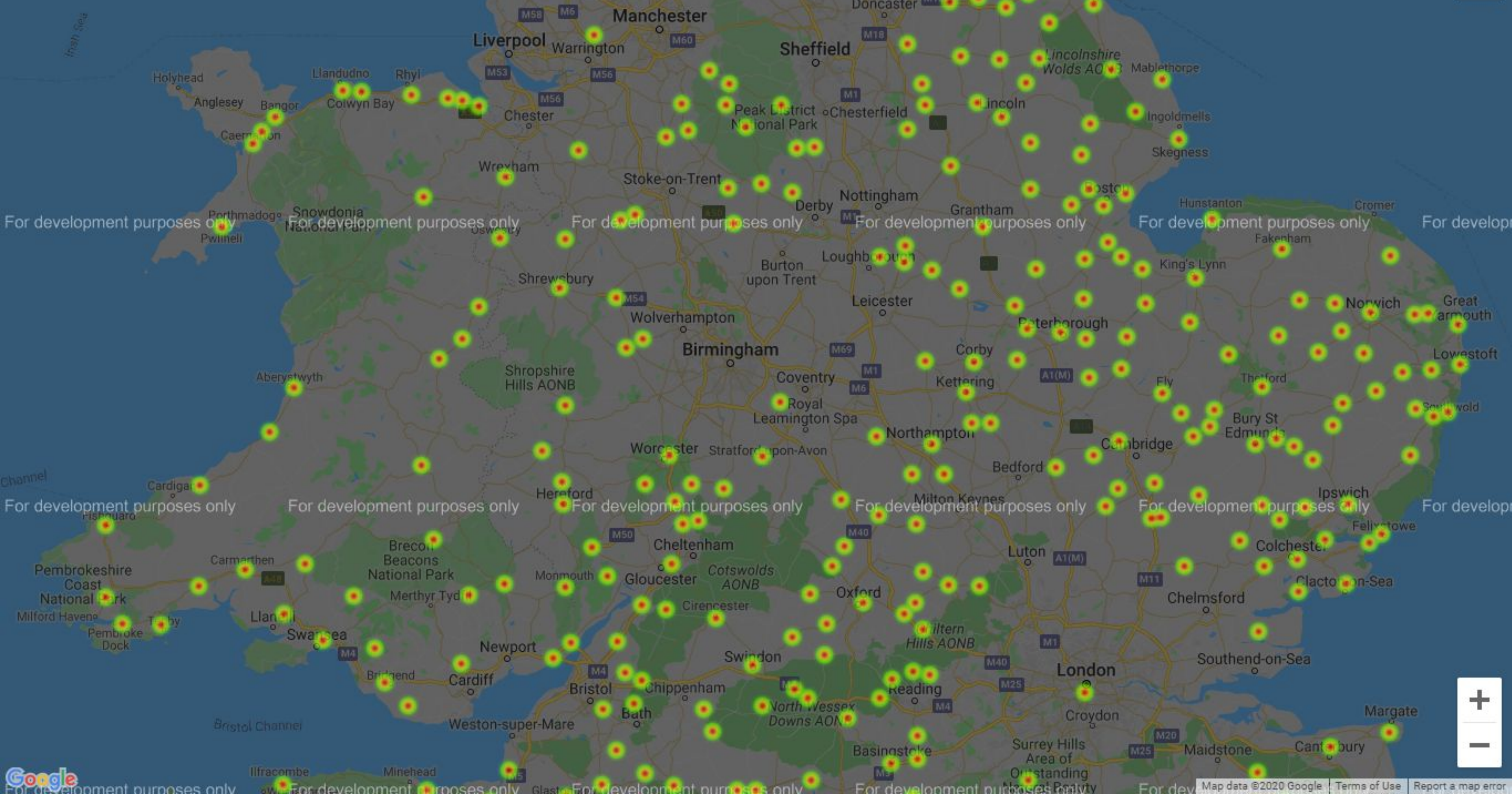
Method: **DBSCAN algorithm**

The main concept of DBSCAN algorithm is to locate regions of high density that are separated from one another by regions of low density.

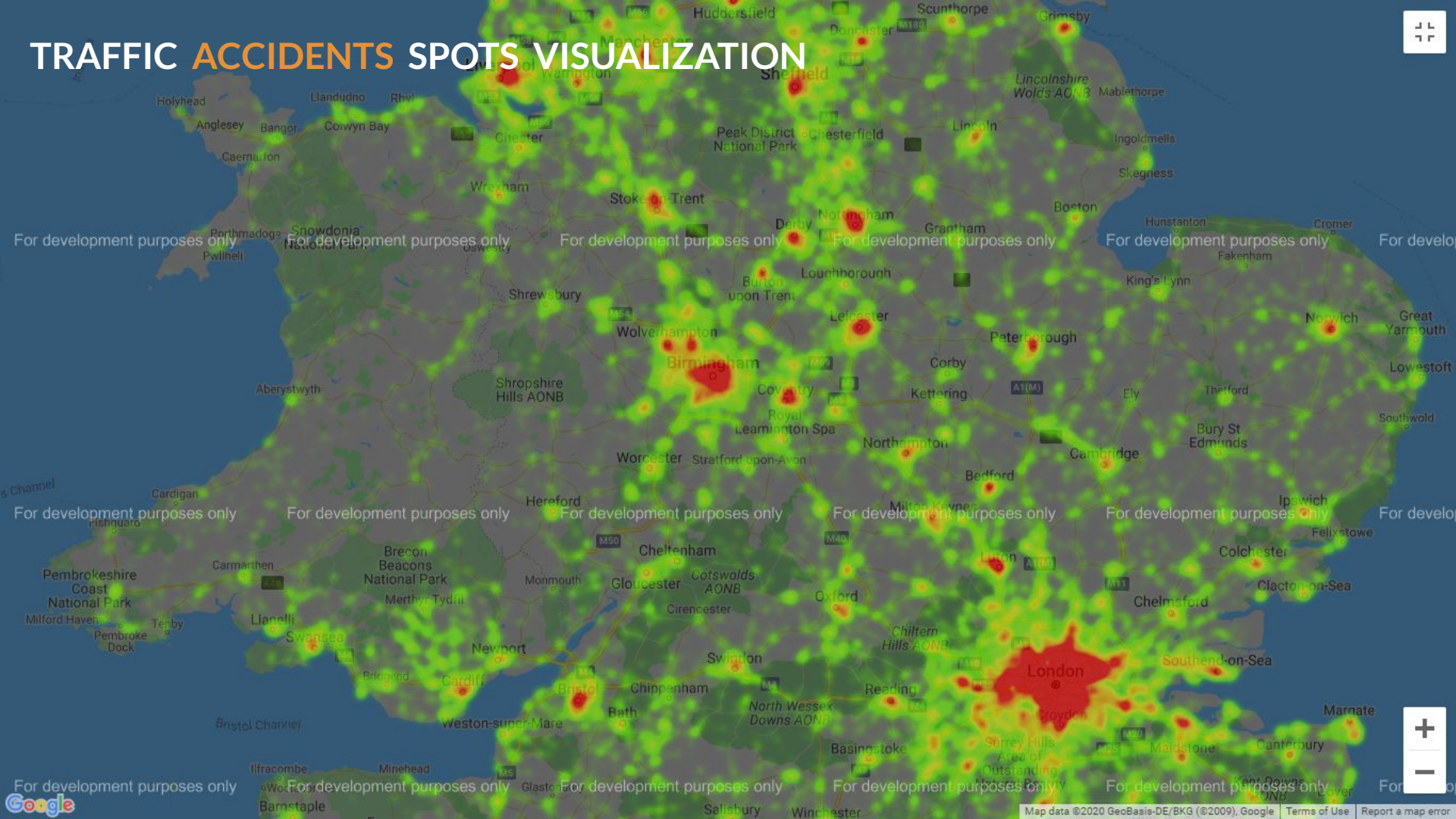
Purpose: To find the road accident and vehicle crime hotspots which will ultimately contribute to the probability of safety of a route.

Parameters:
1.Density at a point P: Number of points within a circle of Radius Eps (ϵ) from point P.
2.Dense Region: For each point in the cluster, the circle with radius ϵ contains at least minimum number of points (MinPts).

TRAFFIC ACCIDENT CLUSTER VISUALIZATIONS



TRAFFIC ACCIDENTS SPOTS VISUALIZATION



PREDICTING CAR ACCIDENT RISK

Objective

- To create an interactive model that can identify accidents hotspots and predict a safe and better path within London. Taking into account the local weather, junctions, accidents and vehicle crime data etc.

Data Processing

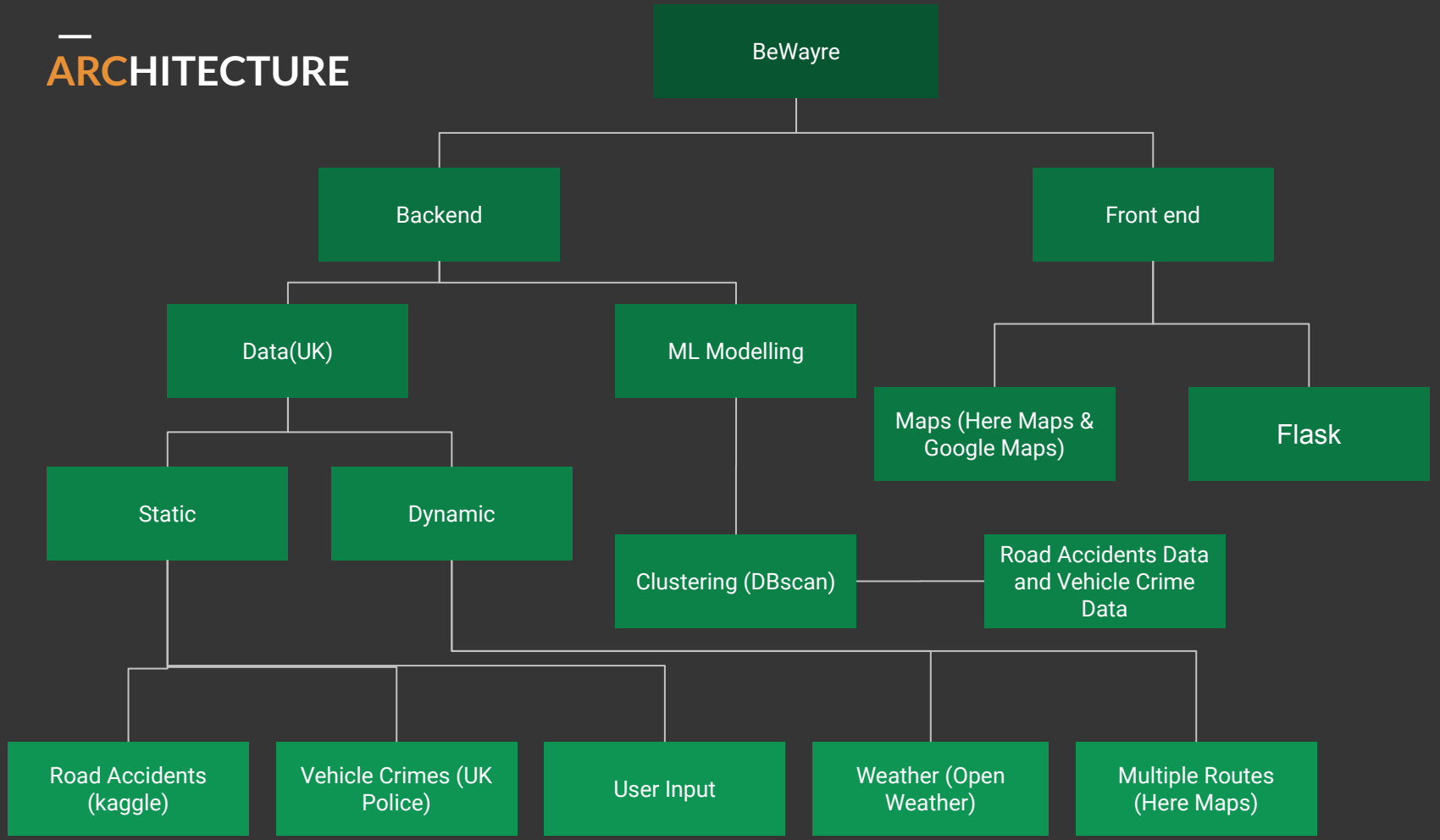
- Cleaning and imputation
- DBSCAN clustering
- Negative Sampling
- Train-test split

Exploration

Supervised Learning

- Model Training
- Logistic Regression
- Model Evaluation
- Feature Importance
- Random Forest

ARCHITECTURE



HERE API

- Uses HERE REST API for geocoding of source & destination and finding alternative routes between them by corresponding to Route_Mode entered for Car.
- Obtaining the waypoints , that were also coordinates of junctions and turns on the each route.
- Uses the clustered data points centroids and a radius of .3 km, for better analysis of Accident and Vehicle Crime spots.
- Analyses the intersection points of path and cluster circle and checks that intersection point actually lie on the paths.

OpenWeathermap API

- Uses OpenWeathermap API for real time weather prediction for given latitude and longitude.
- Obtaining the features values according to different values for model input.
- Uses a limit of 2 kms radius for recalling the weather condition for next coordinates or say junctions
- Conditions: High winds(>45 kmph)
  Rain/Drizzle (%clouds>85)
  Low Visibility(<1000m)

NEGATIVE SAMPLING:

A classic binary classification problem: we needed to find out the combination of factors(e.g. Weather, time, etc) that would "activate" a hotspot.

We followed a method described by *Yuan et al.* **In essence, this method entails the random generation of three negative samples for every positive sample in the cluster.**

MODEL

Logistic Regression

Accuracy: 0.755960264901
Precision: 0.530735235034
Recall: 0.288743169399

Random Forest Classifier

Accuracy: 0.968929359823

Ideas dropped but worth rethinking #Future Plans

Average fuel consumption

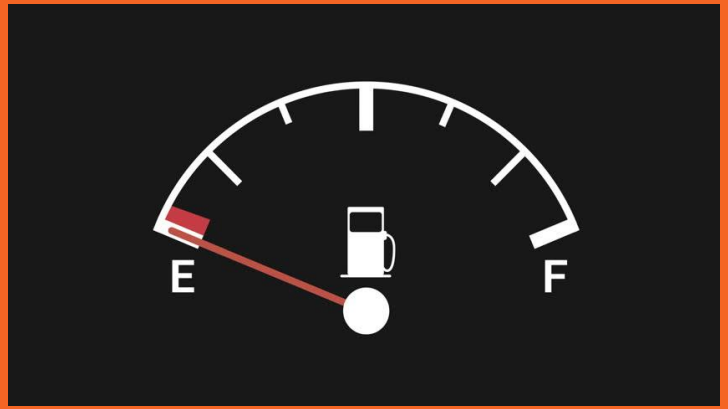
Population Density

Street light coordinates

Mobile Network Coverage

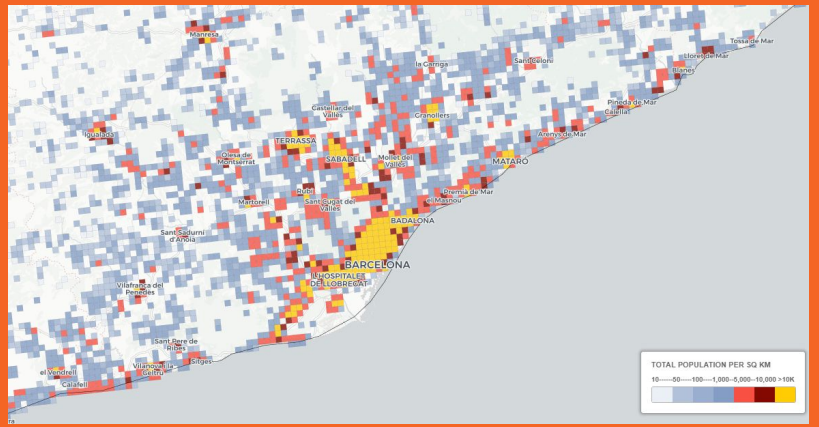
Average fuel consumption

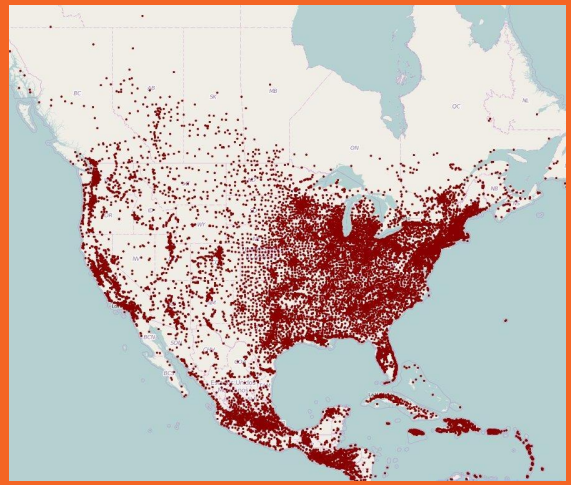
- with this concept, we can make good use of generated data by reading anonymously average fuel consumption at some speed on a certain path of a mercedes car. With repeated data, we can rate the roads and intersection if the fuel consumption is higher than average on that path, the main reason being the quality of that road.
- **Challenges**
  Issue of privacy
  data management
  Improved fuel gauge



Population Density

- This could be traced by the number of mobile users and trend of their movement during day or night.
  The present data of gov open for public is insufficient and not upto the scale that could be implemented for route prediction.
- **Challenges**
  Issue of privacy of mobile users.
  Collaboration with telecom companies or government organisations for data.

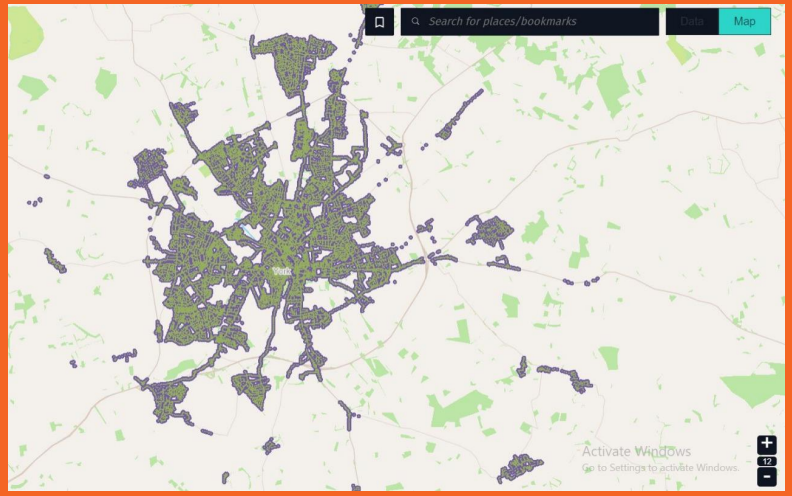


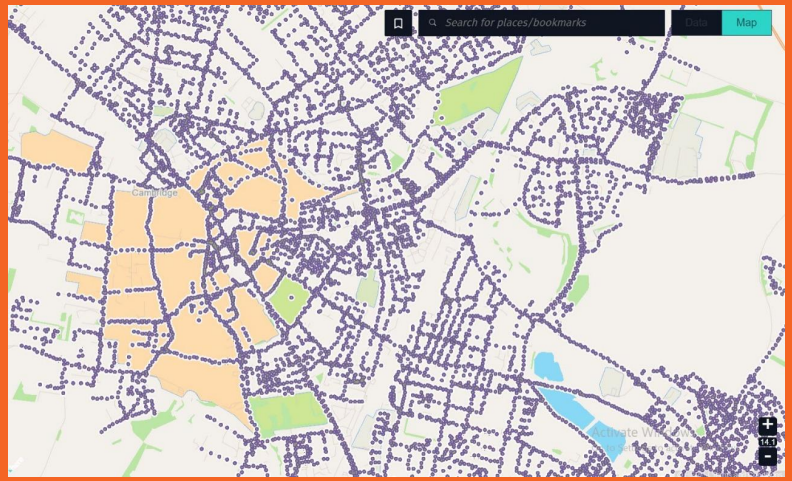


Representative images

Street light Coordinates

- This could be used to find dark stretches along the path and can be used for path selection.
- **Challenges**
  - Where data is available, we have street lights everywhere. So it makes of no use.
  - Where we have uncertainty and/or unavailability of street light, we don't have any data.
- Screenshots added of street light coordinates projected in the map of 1)York city, 2) Cambridgeshire Insight.

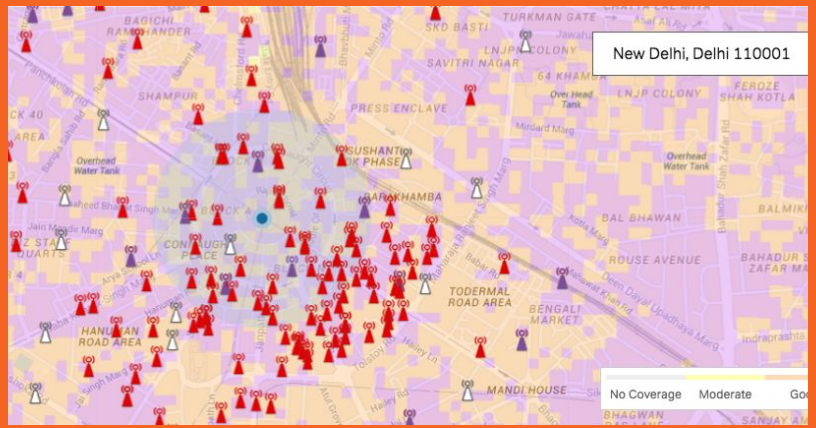




Source: data.gov.uk

Mobile Network Coverage

- This is a vital feature as if there is no connectivity to a certain distance in a path then path may be considered to be unsafe than those having a stable network.
- **Challenges**
  Open network data of each telecom operator in the locality is required.
  I.e. Tower coordinates, height and range, bandwidth etc.



Representative images: Source Airtel

Future plans.

- Implementation of population density either by Mobile phone users location data or data provided by any gov.
- User Feedback Systems, SoS & emergency response
- Reducing the Latency
- Use of Computer Vision (CV) to capture road signs.

Thank You!

We hope our idea has appealed to you and come across as practical and feasible. And we highly appreciate the time you have given to consider what we have put forward.