# PUBLIC TRANSPORTATION EFFICIENCY ANALYSIS

STEPS FOR IMPLEMENTING MACHINE LEARNING ALGORITHM: SUBMITTED BY: JAYAPRIYADHARSHINI.V

NM ID:au211521106068

**1. Model Training:**

* Train the selected machine learning models using the training data.

**2. Model Evaluation:**

* Evaluate the models using the testing data. Common evaluation metrics include Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) for regression problems, and accuracy, precision, recall for classification problems.

**3. Hyperparameter Tuning:**

* Fine-tune the hyperparameters of your models to improve their performance. This can be done using techniques like grid search or random search.

**4. Deployment:**

* Once you have a well-performing model, deploy it to make predictions on new, unseen data. This could involve setting up a real-time prediction system or integrating it into existing software infrastructure.

[1]:

%matplotlib inline

**import numpy as np** *# linear algebra*

**import pandas as pd** *# data processing, CSV file I/O (e.g. pd.read\_csv)*

**import matplotlib.pyplot as plt import datetime**

**import os**

**from math import** sqrt

**import warnings**

*## For Multiple Output in single cell*

**from IPython.core.interactiveshell import** InteractiveShell InteractiveShell.ast\_node\_interactivity = "all" warnings.filterwarnings('ignore')

[2]:

data = pd.read\_csv('../content/20140711.CSV') data.shape

data.head(10)

[2]: (31767, 6)

[2]: TripID RouteID StopID StopName WeekBeginning \ 0 23631 100.0 14156.0 181 Cross Rd 2013-06-30 00:00:00

1 23631 100.0 14144.0 177 Cross Rd 2013-06-30 00:00:00

2 23632 100.0 14132.0 175 Cross Rd 2013-06-30 00:00:00

3 23633 100.0 12266.0 Zone A Arndale Interchange 2013-06-30 00:00:00

4 23633 100.0 14147.0 178 Cross Rd 2013-06-30 00:00:00

5 23634 100.0 13907.0 9A Marion Rd 2013-06-30 00:00:00

6 23634 100.0 14132.0 175 Cross Rd 2013-06-30 00:00:00

7 23634 100.0 13335.0 9A Holbrooks Rd 2013-06-30 00:00:00

8 23634 100.0 13875.0 9 Marion Rd 2013-06-30 00:00:00

9 23634 100.0 13045.0 206 Holbrooks Rd 2013-06-30 00:00:00

NumberOfBoardings 0 1.0

1 1.0

2 1.0

3 2.0

|  |  |
| --- | --- |
| 4 | 1.0 |
| 5 | 1.0 |
| 6 | 1.0 |
| 7 | 1.0 |
| 8 | 1.0 |
| 9 | 1.0 |

[3]:

out\_geo = pd.read\_csv('../content/output\_geo.csv') out\_geo.shape

out\_geo.head()

[3]: (4165, 10)

[3]: accuracy formatted\_address \

1. ROOFTOP 181 Cross Rd, Westbourne Park SA 5041, Australia
2. ROOFTOP 177 Cross Rd, Westbourne Park SA 5041, Australia
3. ROOFTOP 175 Cross Rd, Westbourne Park SA 5041, Australia
4. GEOMETRIC\_CENTER Zone A Arndale Interchange - South side, Kilke…
5. ROOFTOP 178 Cross Rd, Malvern SA 5061, Australia

google\_place\_id input\_string latitude \

1. ChIJKT7I9rbPsGoRVHMHkIy-Oyk 181 Cross Rd -34.966656
2. ChIJ-VFZ87bPsGoRyfVgC5qbPpE 177 Cross Rd -34.966607
3. ChIJIztlirbPsGoR38KRk76kPFI 175 Cross Rd -34.966758
4. ChIJn0C1hCPGsGoRIWvCdhF1RIg Zone A Arndale Interchange -34.875160
5. ChIJycNiylvOsGoRdhfq9GKnpq0 178 Cross Rd -34.964960

longitude number\_of\_results postcode status \

|  |  |  |  |
| --- | --- | --- | --- |
| 0 138.592148 | 1 | 5041 | OK |
| 1 138.592301 | 1 | 5041 | OK |
| 2 138.592715 | 1 | 5041 | OK |
| 3 138.551628 | 1 | 5009 | OK |
| 4 138.611477 | 1 | 5061 | OK |

type

1. street\_address
2. street\_address
3. street\_address
4. bus\_station,establishment,point\_of\_interest,tr…
5. street\_address

[4]:

*#DistanceFromCentre: Distance measure from the city centre*

*#For Calculating Distance between centre with other bus stops by using*␣

↪*Longitude and Latitude*

*#we have used the Haversine formula*

**from math import** sin, cos, sqrt, atan2, radians

**def** calc\_dist(lat1,lon1):

*## approximate radius of earth in km*

R = 6373.0

dlon = radians(138.604801) - radians(lon1) dlat = radians(-34.921247) - radians(lat1)

a = sin(dlat / 2)\*\*2 + cos(radians(lat1)) \* cos(radians(-34.921247)) \*␣

↪sin(dlon / 2)\*\*2

c = 2 \* atan2(sqrt(a), sqrt(1 - a))

**return** R \* c

[5]:

out\_geo['dist\_from\_centre'] = out\_geo[['latitude','longitude']].apply(**lambda** x:␣

↪calc\_dist(\*x), axis=1)

[6]:

out\_geo.head()

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [ | 6]: |  | accuracy | formatted\_address | | \ |
|  |  | 0 | ROOFTOP | 181 Cross Rd, Westbourne Park SA 5041, Australia | |  |
|  |  | 1 | ROOFTOP | 177 Cross Rd, Westbourne Park SA 5041, Australia | |  |
|  |  | 2 | ROOFTOP | 175 Cross Rd, Westbourne Park SA 5041, Australia | |  |
|  |  | 3 | GEOMETRIC\_CENTER | Zone A Arndale Interchange - South side, Kilke… | |  |
|  |  | 4 | ROOFTOP | 178 Cross Rd, Malvern SA 5061, Australia | |  |
|  | | | google\_place\_id input\_string | | latitude | \ |
| 0 | | | ChIJKT7I9rbPsGoRVHMHkIy-Oyk 181 Cross Rd | | -34.966656 |  |
| 1 | | | ChIJ-VFZ87bPsGoRyfVgC5qbPpE 177 Cross Rd | | -34.966607 |  |
| 2 | | | ChIJIztlirbPsGoR38KRk76kPFI 175 Cross Rd | | -34.966758 |  |
| 3 | | | ChIJn0C1hCPGsGoRIWvCdhF1RIg Zone A Arndale Interchange | | -34.875160 |  |
| 4 | | | ChIJycNiylvOsGoRdhfq9GKnpq0 178 Cross Rd | | -34.964960 |  |

longitude number\_of\_results postcode status \

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 138.592148 | 1 | 5041 | OK |  |
| 1 138.592301 | 1 | 5041 | OK |  |
| 2 138.592715 | 1 | 5041 | OK |  |
| 3 138.551628 | 1 | 5009 | OK |  |
| 4 138.611477 | 1 | 5061 | OK |  |
|  |  |  |  | type dist\_from\_centre |

|  |  |
| --- | --- |
| 0 street\_address | 5.180961 |
| 1 street\_address | 5.172525 |
| 2 street\_address | 5.180709 |
| 3 bus\_station,establishment,point\_of\_interest,tr… | 7.057549 |
| 4 street\_address | 4.900099 |

[7]:

*#exp\_data = out\_geo.head(10)*

*##Fill the missing values with mode*

out\_geo['type'].fillna('street\_address',inplace=**True**)

out\_geo['type'] = out\_geo['type'].apply(**lambda** x: str(x).split(',')[-1])

[8]:

out\_geo['type'].unique()

[8]: array(['street\_address', 'transit\_station', 'premise', 'political', 'school', 'route', 'intersection', 'point\_of\_interest', 'subpremise', 'real\_estate\_agency', 'university', 'travel\_agency', 'restaurant', 'supermarket', 'store', 'post\_office'], dtype=object)

[9]:

data['WeekBeginning'] = pd.to\_datetime(data['WeekBeginning']).dt.date data['WeekBeginning'][1]

[9]: datetime.date(2013, 6, 30)