# Transportation Mode Detection with GPS Trajectory Data using Classification


**GROUP 13 MEMBERS:**

Gayathri Sundareshwar
Keerthana Gopikrishnan
Deepasha Jenamani
Bobby Brady


# Problem Definition

Why Travel Mode Detection?

An increase in personal transport vehicles leads to various social and ecological issues.

Detecting travel mode can help:

- Improve City Planning
- Targeted Marketing
- Regulate Carbon Emission

# Problem Background

- Private vehicle ownership has been on a constant rise over the years.
- Increased vehicles lead to traffic congestion and high carbon emission.
- Identifying the preferred mode of transportation in a certain area can help develop ways to avoid social and ecological issues.


China statistics retrieved from CEICdata are showcased since the project uses data collected from Beijing.

# Executive Summary

- Identify GPS data sources with adequate features and volume
- Analyze the trajectories recorded through smartphone sensors.
- Identify features and their correlations.
- Extract quantifiable insights and derive additional features
- Use multilevel classifiers to detect and classify vehicle modes.
- Implement various ML algorithms and apply suitable hyperparameter tuning

# Business and Data Understanding

- Identify Project goals
- Background research on related topics
- Gathering Project and Data Requirements
- Identifying the potential data sources.
- Finalizing the dataset and exploring it further to enhance its usage.

- Multilevel classification of various transportation modes.
- Scope definition of the project.
- Experimenting and identifying best performing ML Model

# Background research

| Author | Data | Objective | Models | Results | Limitations |
|---|---|---|---|---|---|
| Gao et al.(2020) | GIS data | Utilizing public transport network data to detect travel modes. | Normal RF, Criteria-based RF | 94% overall accuracy | Not cost-effective due to transport network modeling and analysis |
| Zhu et al. (2021) | Real time GPS data | Detection of the LDV travel modes using highly imbalanced data. | Regular SVM, OCSVMs with exhaustive feature extraction methods | EFF-OCSVM : 92%<br>Regular SVM(Full features) : 99%<br>Regular SVM(Three features) : 97% | Work better only with large-scale homogeneous data. |
| Rezaie et al. (2017) | GPS and GTFS data, city zone map | Implementation of semi-supervised models to detect travel modes using validated and un-validated data. | DT, RF, Label Propagation using KNN | Semi-supervised work better when more than 70% of data is unlabeled if not supervised is better. | Errors in the estimations due to the scarcity of a detailed dataset |
| F. de S. Soares et al. (2019) | Sensor data | Identification of travel modes using RNN with feature extraction. | Deep RNN architecture with LSTM | Accuracy, precision, recall, and f1-score is nearly 90%. | Not all models were generalized better; only a specific few were. |
| Jahangiri and A. Rakha (2015) | Mobile phone sensor data | Detecting travel modes with data captured from mobile phone sensors. | KNN, SVM, RF, DT, Bagging | SVM - 94.62%<br>Bagging - 95.1% | Misclassification of car mode |
| Tišljari´ et al. (2021) | Cellular network data | Classifying travel modes using origin destination data | RF, DT, LR, NB, KNN | RF - 99.35%<br>DT - 98.93% | Ineffective in real time prediction |

# Project Requirements

| Functional Requirements |
| --- |
| •Qualified dataset with ample features and diverse travel modes. |
| •Source trajectories with target labels attached either in a separate file or to itself. |
| •Standardization of data<br>•Combine data from multiple .plt files into a single common .csv file<br>•Extensive data cleaning and preparation phase to yield optimal results |
| • Identification and extraction of derived features that are quantifiable |
| • Python libraries and packages that are required for model execution |
| • System compatibility to run intensive machine learning jobs |
| • Development of a proper game plan to tackle any roadblock that could occur |

# Project Resource Requirements

| Hardware Requirements | |
|---|---|
| **Hardware** | **Configuration** |
| macOS Local Client (2x) | X86 64-bit CPU, minimum 8GB available system storage, internet connection |
| Google Drive | 2GB available storage |
| Google Research Colaboratory | TPU, GPU |
| Windows Local Client (2x) | X86 64-bit CPU, minimum 8GB available system storage, Internet connection |

| Software Requirements | |
|---|---|
| **Software/ Tools** | **Configuration / License** |
| macOS | macOS 10.15 or higher |
| Windows OS | Windows 10 or higher |
| Anaconda | Open Source |
| Python | 3.7 or higher |
| Google Research Colaboratory | Free licensed |
| Jira | Free licensed |
| Jupyter Notebook | Open Source |

**Project Cost Justification**

| Functionality | Resource Type | Resource | Duration | Cost | Justification |
|---|---|---|---|---|---|
| Project Development | Hardware | 8GB RAM Cost | 4 Months | $27 | Basic necessities of overall project development |
| | Hardware | 64-Bit Intel i7 processor | 4 Months | $470 | |
| | Hardware | 500 GB Hard Disk | 4 Months | $50 | |
| | Tool/Software | Windows OS (10) | 4 Months | $130 | |
| | Hardware | Broadband Connection | 4 Months | $80 | |
| | Hardware | Modem | 4 Months | $150 | |
| Project Management | Tool/Software | Jira | 3 Months | $0 | WBS and Gantt Chart creation |
| Data Storage | Tool/Software | Google Drive Storage | 4 Months | $0 | Storing the dataset |
| Data Processing | Tool/Software | Python, Libraries, Numpy, Scikit Learn, pandas | 3 Months | $0 | Creation of python scripts during the exploratory analysis, data pre-processing, tranformation and preparation |
| Model Implementation | Tool/Software | Anaconda | 3 Months | $0 | Model Training, Development and Evaluation |
| | Tool/Software | Google Research Collaboratory | 1 Month | $0 | |
| | Tool/Software | Jupyter Notebook, Scikit Learn, Keras, Tensorflow | 2 Months | $0 | |

# Data Collection and Dataset Description

- Collected by: Microsoft Research Asia
- File format: .plt (trajectories) and .txt (labels)
- Dense Representation – 91.5% of the dataset
- Dataset Size: 1.67 GB

Data Collection Summary

| | |
|---|---|
| Time span of the collection | 04/2007 – 8/2012 |
| Number of trajectories | 18,670 |
| Number of points | 24,876,978 |
| Total distance | 1,292,951km |
| Total duration | 50,176hour |
| Effective days | 11,129 |

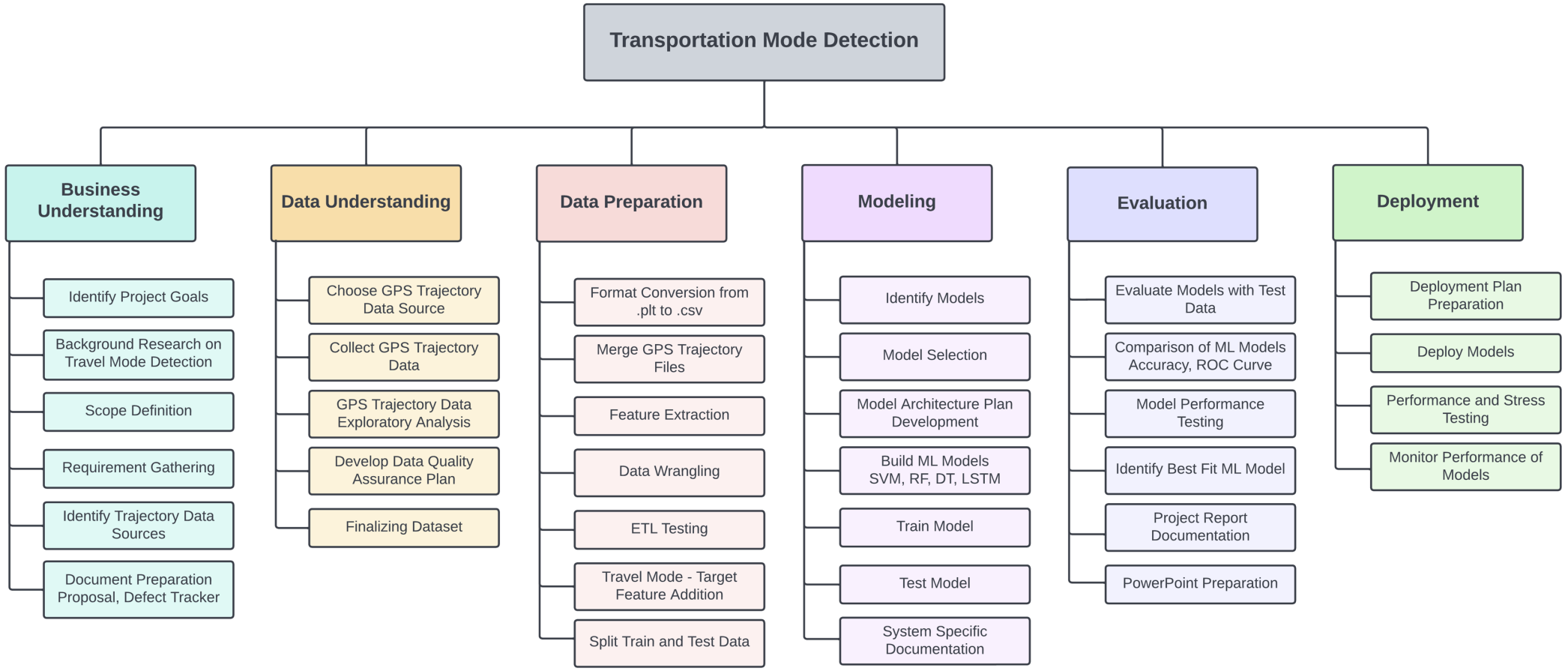
Sample raw .plt file

```
Geolife trajectory
WGS 84
Altitude is in Feet
Reserved 3
0,2,255, My Track,0,0,2,8421376
0
39.981731,116.327236,0,63,39650.9753935185,2008-07-21,23:24:34
39.981748,116.327094,0,133,39650.9754282407,2008-07-21,23:24:37
39.981801,116.327159,0,127,39650.9754513889,2008-07-21,23:24:39
39.981828,116.327253,0,107,39650.975474537,2008-07-21,23:24:41
39.981803,116.327267,0,94,39650.9754976852,2008-07-21,23:24:43
```

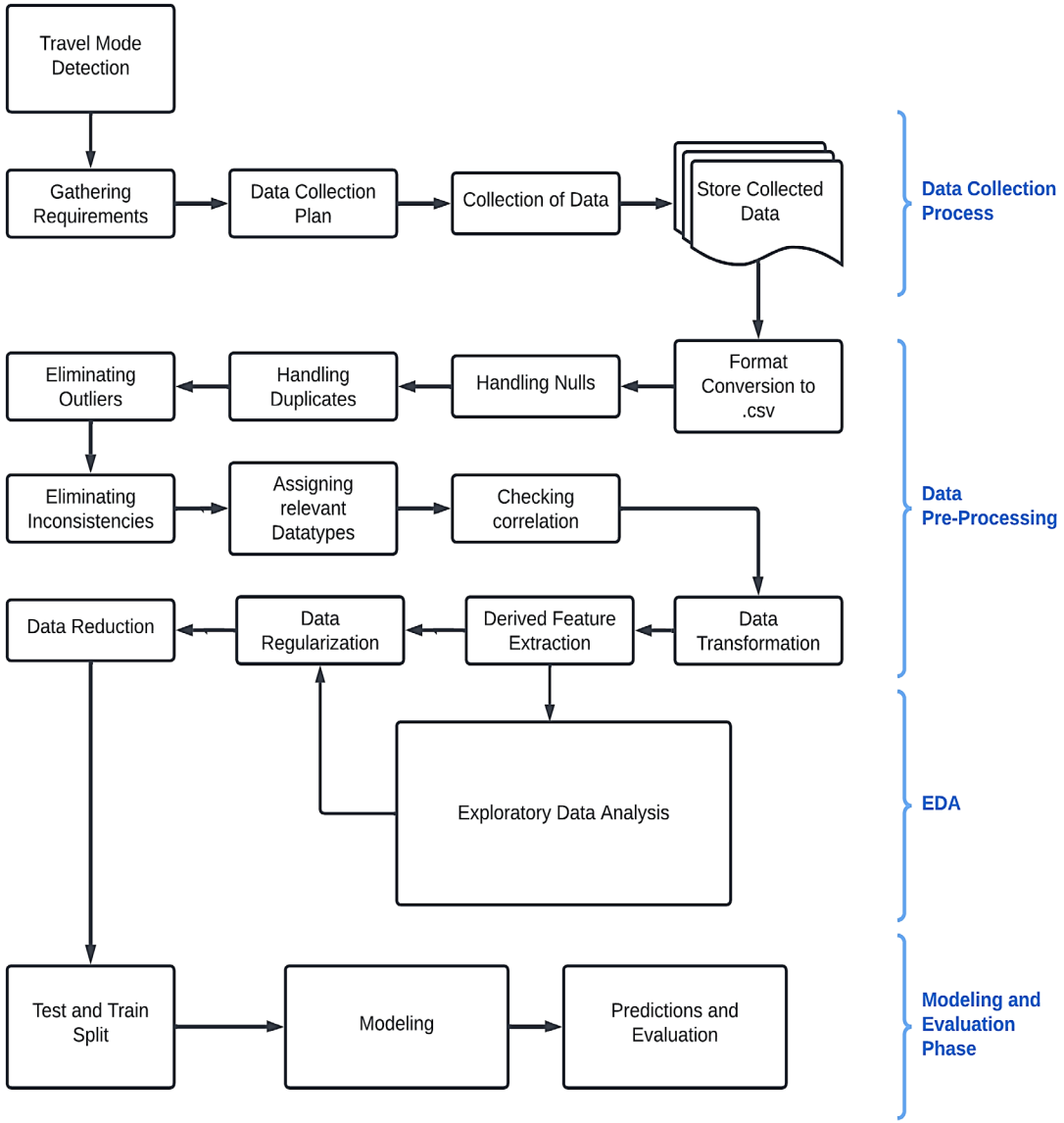
Sample raw .txt file

| Start Time | End Time | Transportation Mode |
|---|---|---|
| 2008/07/22 07:19:54 | 2008/07/22 07:43:02 | taxi |
| 2008/07/22 07:50:32 | 2008/07/22 08:54:40 | bus |
| 2008/07/22 08:54:42 | 2008/07/22 09:10:56 | walk |
| 2008/07/22 10:07:43 | 2008/07/22 10:17:51 | walk |
| 2008/07/22 10:22:13 | 2008/07/22 11:22:23 | bus |

# Project Management Plan



# Workflow Diagram



# Data Cleaning

Sample raw .plt file

(2819105 X 7)

```
Geolife trajectory
WGS 84
Altitude is in Feet
Reserved 3
0,2,255, My Track,0,0,2,8421376
0
39.981731,116.327236,0,63,39650.9753935185,2008-07-21,23:24:34
39.981748,116.327094,0,133,39650.9754282407,2008-07-21,23:24:37
39.981801,116.327159,0,127,39650.9754513889,2008-07-21,23:24:39
39.981828,116.327253,0,107,39650.975474537,2008-07-21,23:24:41
```

Sample raw .txt file

(2819105 X 3)

```
Start Time       End Time Transportation Mode
2008/07/22 07:19:54       2008/07/22 07:43:02       taxi
2008/07/22 07:50:32       2008/07/22 08:54:40       bus
2008/07/22 08:54:42       2008/07/22 09:10:56       walk
2008/07/22 10:07:43       2008/07/22 10:17:51       walk
```

7 Users' Records Sampled from Source Dataset

| | People_Num | Time | Travel Start Time | Travel End Time | Lat | Lon | Alt | Transportation Mode |
|---|---|---|---|---|---|---|---|---|
| 0 | 104 | 3/28/08 8:44 | 3/28/08 8:42 | 3/28/08 9:50 | 40.942837 | 180.065000 | 0.0 | bus |
| 1 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 39.942548 | 180.030000 | 0.0 | bus |
| 2 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 42.942260 | 181.020000 | -777.0 | bus |
| 3 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 38.941900 | 180.064000 | 0.0 | bus |
| 4 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 39.941663 | -179.070000 | -777.0 | bus |
| 5 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 39.940685 | 180.087000 | 0.0 | bus |
| 6 | 104 | 3/28/08 8:48 | 3/28/08 8:42 | 3/28/08 9:50 | 39.940363 | -178.090000 | 0.0 | bus |
| 7 | 167 | 6/5/08 1:57 | 6/5/08 1:56 | 6/5/08 2:17 | 40.006297 | 116.320673 | 221.0 | bike |
| 8 | 167 | 6/5/08 1:57 | 6/5/08 1:56 | 6/5/08 2:17 | 40.006306 | 116.320783 | 221.0 | bike |
| 9 | 104 | 3/28/08 8:49 | 3/28/08 8:42 | 3/28/08 9:50 | 39.947393 | 116.303697 | 0.0 | bus |

Format Conversion and Merging to Single .csv File (2819105 X 8)

**Data Cleaning Operations**

- Handling Nulls
- Handling Redundancies
- Eliminate Abnormalities
- Eliminate Simultaneous trajectories
- Eliminate trajectories with the same start and end time

| People_Num | Time | Travel Start Time | Travel End Time | Lat | Lon | Alt | Transportation Mode |
|---|---|---|---|---|---|---|---|
| 104 | 2008-03-28 08:44:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.962098 | 116.301595 | 0.0 | bus |
| 104 | 2008-03-28 08:48:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.948270 | 116.303298 | 0.0 | bus |
| 104 | 2008-03-28 08:49:20 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.947045 | 116.303850 | 0.0 | bus |
| 104 | 2008-03-28 08:50:18 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.940685 | 116.304043 | 0.0 | bus |
| 104 | 2008-03-28 08:51:08 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.936278 | 116.303712 | 0.0 | bus |
| 104 | 2008-03-28 08:52:00 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.932202 | 116.303695 | 0.0 | bus |
| 104 | 2008-03-28 08:52:52 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.930165 | 116.304010 | 0.0 | bus |
| 104 | 2008-03-28 08:53:52 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.927862 | 116.304065 | 0.0 | bus |
| 104 | 2008-03-28 08:54:44 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.925368 | 116.304213 | 0.0 | bus |
| 104 | 2008-03-28 08:55:36 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.923573 | 116.304010 | 0.0 | bus |

Sample Records After Data Cleaning Operations (2778059 X 8)

*Before*

```
People_Num : 2819105
Time : 2819105
Travel Start Time : 2819105
Travel End Time : 2819105
Lat : 2819105
Lon : 2819105
Alt : 2819105
```

*After*

```
People_Num : 2778059
Time : 2778059
Travel Start Time : 2778059
Travel End Time : 2778059
Lat : 2778059
Lon : 2778059
Alt : 2778059
```

Record count before and after

# Data Pre-Processing

| People_Num | Time | Travel Start Time | Travel End Time | Lat | Lon | Alt | Transportation Mode |
|---|---|---|---|---|---|---|---|
| 104 | 2008-03-28 08:44:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.962098 | 116.301595 | 0.0 | bus |
| 104 | 2008-03-28 08:48:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.948270 | 116.303298 | 0.0 | bus |
| 104 | 2008-03-28 08:49:20 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.947045 | 116.303850 | 0.0 | bus |
| 104 | 2008-03-28 08:50:18 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.940685 | 116.304043 | 0.0 | bus |
| 104 | 2008-03-28 08:51:08 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.936278 | 116.303712 | 0.0 | bus |
| 104 | 2008-03-28 08:52:00 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.932202 | 116.303695 | 0.0 | bus |
| 104 | 2008-03-28 08:52:52 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.930165 | 116.304010 | 0.0 | bus |
| 104 | 2008-03-28 08:53:52 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.927862 | 116.304065 | 0.0 | bus |
| 104 | 2008-03-28 08:54:44 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.925368 | 116.304213 | 0.0 | bus |
| 104 | 2008-03-28 08:55:36 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.923573 | 116.304010 | 0.0 | bus |

Records retrieved after Data Cleaning Operations (2778059 X 8)

| | Unnamed: 0 | People_Num | Time | Travel Start Time | Travel End Time | Lat | Lon | Alt | Transportation Mode | Travel Count | Time Gap(s) | Distance(m) | Speed(m/s) | Acceleration(m/s^2) | Total Time(s) | Total Distance(m) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 104 | 2008-03-28 08:44:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.962098 | 116.301595 | 0.0 | bus | 1 | 240 | 1542.2920914409665 | 6.43 | -0.01 | N.A | N.A |
| 1 | 1 | 104 | 2008-03-28 08:48:30 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.948270 | 116.303298 | 0.0 | bus | 1 | 50 | 143.9636385353144 | 2.88 | 0.19 | N.A | N.A |
| 2 | 2 | 104 | 2008-03-28 08:49:20 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.947045 | 116.303850 | 0.0 | bus | 1 | 58 | 706.3668121584448 | 12.18 | -0.04 | N.A | N.A |
| 3 | 3 | 104 | 2008-03-28 08:50:18 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.940685 | 116.304043 | 0.0 | bus | 1 | 50 | 490.10067854245034 | 9.8 | -0.02 | N.A | N.A |
| 4 | 4 | 104 | 2008-03-28 08:51:08 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.936278 | 116.303712 | 0.0 | bus | 1 | 52 | 452.6519866162259 | 8.7 | -0.08 | N.A | N.A |
| 5 | 5 | 104 | 2008-03-28 08:52:00 | 2008-03-28 08:42:00 | 2008-03-28 09:50:00 | 39.932202 | 116.303695 | 0.0 | bus | 1 | 52 | 227.727881360944 | 4.38 | -0.0 | N.A | N.A |

Additional Features Extraction (148489 X 15)

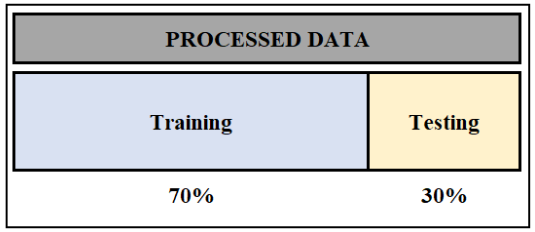
```
[[-0.57725638 -0.36064775 -0.68514532 -0.71842107 -0.63942502 -0.75000622]
 [-0.60200882 -0.35953161 -0.67269288 -0.71842107 -0.63942502 -0.75000622]
 [ 0.83363229  1.00328162  0.5632115   0.42896792  0.1636005   0.08555144]
 [-0.47747314 -0.29367904 -0.41430481 -0.58343413 -0.37174985 -0.471487  ]
 [-0.07292557 -0.12737341 -0.12478565  0.0240071   0.1636005   0.36407066]
 [-0.41481854 -0.24345251 -0.50458498 -0.58343413 -0.63942502 -0.75000622]
 [-0.60510287 -0.36288004 -0.69448465 -0.71842107 -0.63942502 -0.75000622]
 [-0.34210827 -0.13183799 -0.277328   -0.31346025 -0.37174985 -0.471487  ]
```

Sample Results of Data Standardization

| | Travel Count | Transportation Mode | Max Speed(m/s) | 95% Speed(m/s) | 75% Speed(m/s) | Mean Speed(m/s) | Speed Std | Max Acceleration(m/s^2) | 95% Acceleration(m/s^2) | 75% Acceleration(m/s^2) |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | bus | 12.18 | 10.46 | 5.87 | 3.41 | 3.83 | 0.19 | 0.11 | 0.04 |
| 1 | 2 | bus | 15.97 | 15.84 | 15.51 | 11.80 | 5.18 | 0.19 | 0.14 | 0.07 |
| 2 | 5 | walk | 1.83 | 1.72 | 1.51 | 1.34 | 0.29 | 0.01 | 0.01 | 0.01 |
| 3 | 7 | bus | 12.78 | 11.97 | 9.67 | 6.78 | 3.56 | 0.16 | 0.16 | 0.08 |
| 4 | 8 | walk | 1.50 | 1.50 | 1.45 | 1.34 | 0.14 | 0.01 | 0.01 | 0.00 |

| Mean Acceleration(m/s^2) | Acceleration Std | Non 0 Mean Speed(m/s) | Non 0 Mean Acceleration(m/s^2) | Total Time(s) | Total Distance(m) |
|---|---|---|---|---|---|
| 0.03 | 0.05 | 3.41 | 0.03 | 3811 | 6465.72 |
| 0.04 | 0.05 | 11.80 | 0.04 | 939 | 9533.53 |
| 0.00 | 0.00 | 1.34 | 0.00 | 601 | 791.24 |
| 0.06 | 0.05 | 6.78 | 0.06 | 866 | 5771.59 |
| 0.00 | 0.00 | 1.34 | 0.00 | 588 | 783.50 |

Data Aggregation




Splitting Training and Testing Data

# Support Vector Machine

PROS

- Productive in high dimensional cases
- The risk of overfitting is less.
- Works well when there is clear margin of separation

CONS

- Not suitable when target classifiers overlap
- More Features, More Complexities
- Longer training period

| SUPPORT VECTOR MACHINE | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Classifier** | **kernel** | **gamma** | **c** | **Weighted F1 score** | **Weighted Precision** | **Weighted Recall** | **Accuracy** |
| Model 1 | rbf | None | None | 0.7938 | 0.78 | 0.81 | 0.81 |
| Model 2 | rbf | 0.5 | 5 | 0.83 | 0.83 | 0.84 | 0.84 ⬆ |

# Random Forest

PROS

- Parallel Architecture
- Normalization is not required due to the rule-based approach.
- Flexible to handle both regression and classification

CONS

- Higher complexity
- Fails to determine the significance of each variable
- Longer training period

| RANDOM FOREST | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Classifier** | **random_state** | **max_depth** | **n_jobs** | **criterion** | **Weighted F1 score** | **Weighted Precision** | **Weighted Recall** | **Accuracy** |
| Model 1 | None | None | None | None | 0.82 | 0.81 | 0.83 | 0.83 |
| Model 2 | 233 | 10 | 2 | entropy | 0.86 | 0.86 | 0.87 | 0.87 ⬆ |

# Decision Tree

PROS

- Non-parametric algorithm
- Normalization is not required
- Easy to understand

CONS

- Prone to overfitting
- Time consuming
- Unstable and small changes can have a large impact

| DECISION TREE | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Classifier** | **criterion** | **max_depth** | **random_state** | **Weighted F1 score** | **Weighted Precision** | **Weighted Recall** | **Accuracy** |
| Model 1 | gini | None | None | 0.7545 | 0.75 | 0.75 | 0.75 |
| Model 2 | gini | 4 | 233 | 0.7813 | 0.78 | 0.79 | 0.79 ⬆ |

# Long Short Term Memory

PROS

- Learns short and long term dependencies
- Less prone to vanishing gradient compared to other RNN Architecture
- Simple Implementation

CONS

- Longer training time
- Data must be sequential
- Harder to interpret the results

| LONG SHORT TERM MEMORY | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Classifier** | **LSTM Units** | **Dropout Unit** | **Bayesian Probability Units** | **Weighted F1 score** | **Weighted Precision** | **Weighted Recall** | **Accuracy** |
| Model 1 | 32 | None | None | 0.23 | 0.25 | 0.3 | 0.30 |
| Model 2 | 32 | 0.5 | 14 | 0.14 | 0.66 | 0.29 | 0.29 |

# Model Evaluation and Comparison

| Alogrithm | Weighted F1 score | Weighted Precision | Weighted Recall | Accuracy |
|---|---|---|---|---|
| Decision Tree | 0.7813 | 0.78 | 0.79 | 0.79 |
| Random Forest | 0.86 | 0.86 | 0.87 | 0.87 |
| Support Vector Machine | 0.83 | 0.83 | 0.84 | 0.84 |
| Long- Short Term Memory | 0.23 | 0.25 | 0.30 | 0.30 |
| K-Nearest Neighbour | 0.806 | 0.81 | 0.81 | 0.81 |
| Logistic Regression | 0.76 | 0.75 | 0.78 | 0.78 |

# Limitations and Future Scope

## Limitations:

- The imbalanced count of records for each travel mode in the dataset made it hard to predict the modes with lesser recorded trajectories.
- The model works well with the recorded trajectories, but the utilization of real-time data was not explored.

## Future Scope:

- Utilization of real-time data can be considered
- Exploring ways to extract insights on how this can be utilized in research areas like city planning and targeted marketing.

# References


- Gao, L., Chen, X., Zhu, Z., & Chang, T. H. (2020, September). Effectiveness of Public Transport Networks in Motorized Mode Detection: A Case Study of a Planning Survey in Nanjing. *2020 IEEE 5th International Conference on Intelligent Transportation Engineering (ICITE)*, *22*(9), 5473–5485. https://doi.org/10.1109/icite50838.2020.9231462

- Tišljari´, L., Cvetek, D., Vareskic, V., & Greguric, M. (2021). Classification of Travel Modes from Cellular Network Data Using Machine Learning Algorithms. *2021 International Symposium ELMAR*. https://doi.org/10.1109/elmar52657.2021.9550817

- Jahangiri, A., & Rakha, H. A. (2015). Applying Machine Learning Techniques to Transportation Mode Recognition Using Mobile Phone Sensor Data. *IEEE Transactions on Intelligent Transportation Systems*, *16*(5), 2406–2417. https://doi.org/10.1109/tits.2015.2405759

- Zhu, L., Borlaug, B., Lin, L., Holden, J., & Gonder, J. (2021, April). Identifying Light-Duty Vehicle Travel from Large-Scale Multimodal Wearable GPS Data with Novelty Detection Algorithms. *2021 IEEE Green Technologies Conference (GreenTech)*. https://doi.org/10.1109/greentech48523.2021.00034

- Rezaie, M., Patterson, Z., Yu, J., & Yazdizadeh, A. (2017, September). Semi-Supervised travel mode detection from smartphone data. *2017 International Smart Cities Conference (ISC2)*, *19*(5), 1547–1558. https://doi.org/10.1109/isc2.2017.8090800

- F. de S. Soares, E., Salehinejad, H., Campos, C. A. V., & Valaee, S. (2019, December). Recurrent Neural Networks for Online Travel Mode Detection. *2019 IEEE Global Communications Conference (GLOBECOM)*, *22*(9), 5473–5485. https://doi.org/10.1109/globecom38437.2019.9013316

- *Global Economic Data, Indicators, Charts & Forecasts.* (n.d.). CEIC. https://www.ceicdata.com/en

- Microsoft. (n.d.). GeoLife GPS Trajectories. *Microsoft Downloads*. Retrieved September 27, 2022, from https://www.microsoft.com/en-us/download/details.aspx?id=52367&from=https%3A%2F%2Fresearch.microsoft.com%2Fen-us%2Fdownloads%2Fb16d359d-d164-469e-9fd4-daa38f2b2e13%2F