

## Accessibility Rating and Public Car Spaces

Method:

The following characteristics of a suburb are used to determine the accessibility level:

1. Working age population
2. Number of transport stops (trains, wharves, bus interchange) in a suburb
3. Population per mode of transport (Suburb level)
4. Number of public parking spaces available in a suburb

We are using a two-step modelling process.

The first step involves applying Machine Learning via BIRCH Clustering, an Unsupervised Learning technique. Through this we are able to create three clusters that will identify accessibility levels for each suburb in our dataset.

Next, the result of the clustering model is used as input to a Logistic Regression Model. The regression model can be used to predict the accessibility level of a suburb that is not in our existing dataset given the suburb's characteristics.

Looking at the feature importance from the Regression model, we are also able to infer the impact of each suburb characteristics to its accessibility level.

Suburbs that are classified under the low accessibility level cluster exhibit the following characteristics based on the regression model coefficients:

1. Likely to have low working age population
2. Less likely to have available public parking spaces
3. Less likely to have population that use the modes of transport considered in the model
4. Most of these suburbs are located outside the urban areas

Suburbs that are classified under the high accessibility level cluster exhibit the following characteristics based on the regression model coefficients:

1. More likely to have high working age population
2. More likely to have available public parking spaces
3. More likely to have population that use the public transport
4. Most of these suburbs are located in the major train interchange
5. Most of these suburbs are located in the major urban areas

Suburbs classified as medium are those that have characteristics in between the ones mentioned above. These suburbs are also likely to be closer to major urban areas but are likely to have less public parking spaces close to public transport. Hence, people from these suburbs are more likely to use their own vehicles than use public transport. This can be a focus of city planners when aiming to reduce car use by population.

Further improvements are planned to increase the reliability of the model. Some of the issues encountered when implementing the project that need to be addressed are:

1. Availability of similar data for all suburbs in NSW
2. We identified that the Public Transport Accessibility Level (PTAL) will be useful data to help improve the model, however, access to this data in a format that can be used is difficult. This will be included in future iteration of the project.
3. City planning subject matter expertise will help improve the assignment of the clusters and interpretation of drivers that affect classification.

Addition of new features will help us improve the way we determine accessibility scores for the suburbs. The following features will also be considered in the modelling in future iteration:

1. Car sharing
2. Cycling
3. Essential Services.

The aim of the project is to be able to aid city planners to design better transport systems. Future iterations of the project will include other capabilities such as Scenario-based Planning and Car Park Spaces Demand Forecasting which will leverage the machine learning models.