To train our predictive model, we read the .csv file and compiled into a list of 1) time, 2) seats occupied and 3) day of the week. Subsequently, we fit our data onto k-Nearest Neighbours classifiers, split the dataset into training and test sets and find the k-value that will return the best prediction rate. From our **GUI**, users can enquire about the seats’ occupancies for a specific day and time of the week. This input will be parse to the backend’s server through **Firebase** and run in the k-NN classifier with the best k-value. As such, we will be able to predict the number of seats occupied from the weighted average of seats’ occupancies across the week. This predicted value is then parse back to the GUI through firebase and displayed to the user.

For our predictive model, we did not consider normalising our data because all the data points are substantially distributed (in the x and y-axis) and we want the difference between each seat occupied and each time to have equal weightage.

To further improve our model, our backend server is able to update new records of time; seats and day into our existing database at every hourly interval when the device is in operation. This will increase the amount of data in our database and provide a better and more accurate prediction in the future.