Project Proposal

# SG5326 Machine Learning

# for Software Engineers

## Group 8

Ang Seng Ngiap Melvin

Shi Jingli

Sivalingarajah Ajanthan

Wong Chee Vui

**Business scenario**

There are many property websites advertising their property deals online with different variants of features. Such features include property size, number of bedrooms, number of washrooms, location attractiveness, condition of existing houses and grandeur or grade of housing etc.

Such deals are based on demand and supply of the buyers and sellers. As a perfect market exists only in economic theory, it is therefore impossible for every sellers and buyers to possess perfect knowledge. Sellers will sometimes overestimated or underestimated the selling price of their existing homes. Concurrently, buyers will also offer either low or high prices to sellers for their homes.

To mitigate this issue, we can employ machine-learning techniques here to predict the housing prices of a certain district. With this, buyers and sellers could make an informed decision in buying or selling their houses respectively, thereby creating a much fairer and transparent property transaction environment.

To illustrate our point, we would be using the sample dataset on the house sales between May 2014 and May 2015 in King County, Seattle, USA in this project. The dataset can be obtained from this [url](https://www.kaggle.com/harlfoxem/housesalesprediction).

The proposed system would be subjected to a couple of machine learning algorithms and the most appropriate one would then be chosen to predict the sale price of a house. The programming language chosen is python for the machine learning engine and QT for the user interface module.

User would be able to determine the percentage of training and testing set before loading into the system for prediction.