## Take Home Assignment - 2021/01/22

Hi, thank you for your time for interviewing with us. This assignment is going to take around 1~2 hours to complete.

### Dataset

The training and testing datasets are included in this folder. Please check.

The dataset we choose for this assignment is the boston housing dataset. Each record in the database describes a Boston suburb or town. The data was drawn from the Boston Standard Metropolitan Statistical Area (SMSA) in 1970. The attributes are deﬁned as follows (taken from the UCI Machine Learning Repository1): CRIM: per capita crime rate by town

1. ZN: proportion of residential land zoned for lots over 25,000 sq.ft.
2. INDUS: proportion of non-retail business acres per town
3. CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
4. NOX: nitric oxides concentration (parts per 10 million)
5. RM: average number of rooms per dwelling
6. AGE: proportion of owner-occupied units built prior to 1940
7. DIS: weighted distances to ﬁve Boston employment centers
8. RAD: index of accessibility to radial highways
9. TAX: full-value property-tax rate per $10,000
10. PTRATIO: pupil-teacher ratio by town 12. B: 1000(Bk−0.63)2 where Bk is the proportion of blacks by town 13. LSTAT: % lower status of the population
11. MEDV: Median value of owner-occupied homes in $1000s  
    We can see that the input attributes have a mixture of units.

Task

1. Please use any ML model (such as regression) to predict the housing value( MEDV ) for the test dataset. Please include all the procedures and code as well as the results.
2. (Optional) If you have time left, please set up an endpoint api for predicting the result. So your endpoint should take the input data in test dataset( single row of data or a batch) and return the predictions. Please include your code and steps to use them.