

# ASS3: ANN

6440126922 นิปุ่น อังควิชัย

จากการสร้างแบบจำลอง Artificial Neural Network ทำนายราคาบ้าน ตามขั้นตอนใน Google Colab notebook นี้ [🔗 ASS3: ANN\\_Nipun Angkavichai.ipynb](#)

## กระบวนการ Data preprocessing

ทำ Data normalization กับข้อมูลประเภท Numeric ได้แก่ 'crim', 'zn', 'indus', 'nox', 'rm', 'age', 'dis', 'rad', 'tax', 'ptratio', 'black', 'lstat' ทั้ง 12 columns (มีเพียง rad ที่เป็นข้อมูลประเภท Categorical)

## การพัฒนาโมเดล

เริ่มต้นด้วยการทำ hyperparameter fine-tuning กับโมเดล Artificial Neural Network โดยมีรายละเอียดดังนี้

```
def build_model(hp):  
    model = keras.Sequential()  
    model.add(layers.Dense(units=hp.Int('units', min_value=32, max_value=512, step=32),  
                           activation='relu', input_dim=X_train.shape[1]))  
    model.add(layers.Dense(1, activation='linear'))  
    model.compile(optimizer=keras.optimizers.Adam(learning_rate=hp.Choice('learning_rate', values=[1e-2, 1e-3, 1e-4, 1e-5])),  
                  loss='mean_squared_error', metrics=['mae'])  
    return model
```

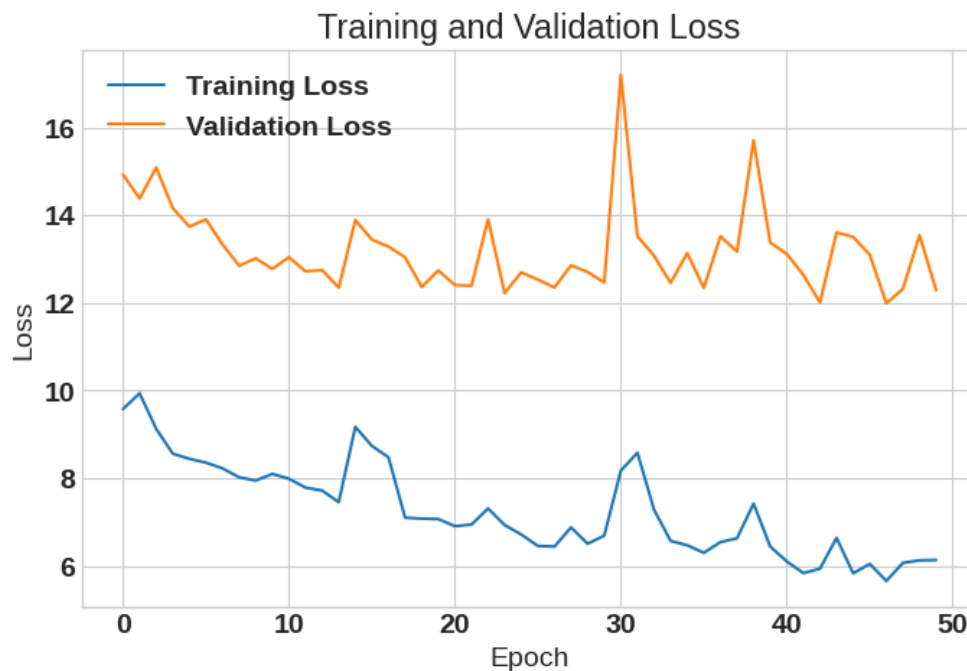
- สำรวจจำนวนของ hidden unit ตั้งแต่ 32 ไปจนถึง 512 โดยขยับขึ้นทีละ 32
- สำรวจขนาด learning rate [1e-2, 1e-3, 1e-4, 1e-5]

นอกจากนี้ก็จะใช้ MSE เป็น loss function และใช้ MAE เป็น metric ในการวัดประสิทธิภาพของโมเดลระหว่างการเทรนโมเดล

จากกระบวนการนี้ทำให้ได้ผลลัพธ์ hyperparameter ที่ดีที่สุดโดยมี val\_loss (MSE) อยู่ที่ 13.23 ดังนี้

- hidden unit = 64
- learning rate = 1e-2 หรือ 0.01

ขั้นตอนต่อไปจะเทรนโมเดลด้วย hyperparameter ที่ดีที่สุดกับข้อมูลทั้งหมด และ plot loss เป็นกราฟ ระหว่างการเทรนโมเดล จะได้กราฟดังนี้



พบว่าจำนวน epoch ที่มี val\_loss หรือ loss ใน validation set ต่ำที่สุดคือ 47 หลังจากนั้นเราจึงใช้ epoch นี้ในการเทรนโมเดลที่ดีที่สุดเพื่อนำไปทดสอบกับ Test set ของเราก็คือ

## ผลลัพธ์

เมื่อนำโมเดลมาทดสอบกับ Test set จะได้ผลลัพธ์ดังนี้

Mean Absolute Error (MAE)	2.16
Mean Squared Error (MSE)	8.50
Root Mean Squared Error (RMSE)	2.91

## อภิปรายผลลัพธ์

จากผลลัพธ์ เมื่อดูจาก RMSE ซึ่งเป็นมาตรวัดที่เหมาะสมในการตีความประสิทธิภาพโมเดลเนื่องจากอยู่ในหน่วยเดียวกันกับค่าที่ทำนายออกมา ได้ค่าเฉลี่ยของขนาดความคลาดเคลื่อนเท่ากับ 2.91 ทำให้เห็นว่าโมเดลทำนายออกมาได้ค่อนข้างใกล้เคียงกับราคาบ้านใน Test set นอกจากนี้โมเดลยังมีความคลาดเคลื่อนใน metric ต่างๆ ที่ไม่ต่างกัน Training set และ Validation set มาก ในระหว่างที่เทรนโมเดล ทำให้เห็นว่าโมเดลของเราไม่ overfitting หรือทำได้ดีเพียงแต่ข้อมูลที่ใช้เทรน