

# PREDICT SET.BK

ราคานิ่ตยาดหลักทรัพย์แห่งประเทศไทย



Nattapong  
Pakdeeviboon



# CONTENT

**01**

## Motivation

- concept

**02**

## Approach

- data
- feature engineering
- normalization
- test train split
- model

**03**

## Experiment

- feature selection
- mse
- trading simulation

**04**

## Result

- result

**05**

## Conclusion

- summary
- limitations
- future project

# MOTIVATION

predict ค่า close ของ set.bk  
กำหนดค่า Close ของวันนี้ กับของวันพรุ่งนี้

Run Program - AI



อนาคต - วันถัดไป

ลงทุนซื้อ - ขาย

# DATA



$^{\wedge}$ SET.BK

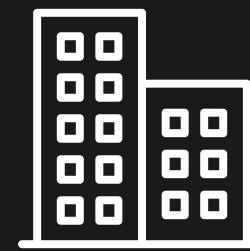
ดัชนีราคาหุ้นตลาดหลักทรัพย์แห่งประเทศไทย



4 กลุ่มนักลงทุน

- กลุ่มนักลงทุนร่วมกับสถาบันภายในประเทศ
- กลุ่มพ่อตีโبر์เกอร์
- กลุ่มนักลงทุนต่างประเทศ
- กลุ่มนักลงทุนรายย่อย

SET THAILAND



GOLD



$^{\wedge}$ DJI

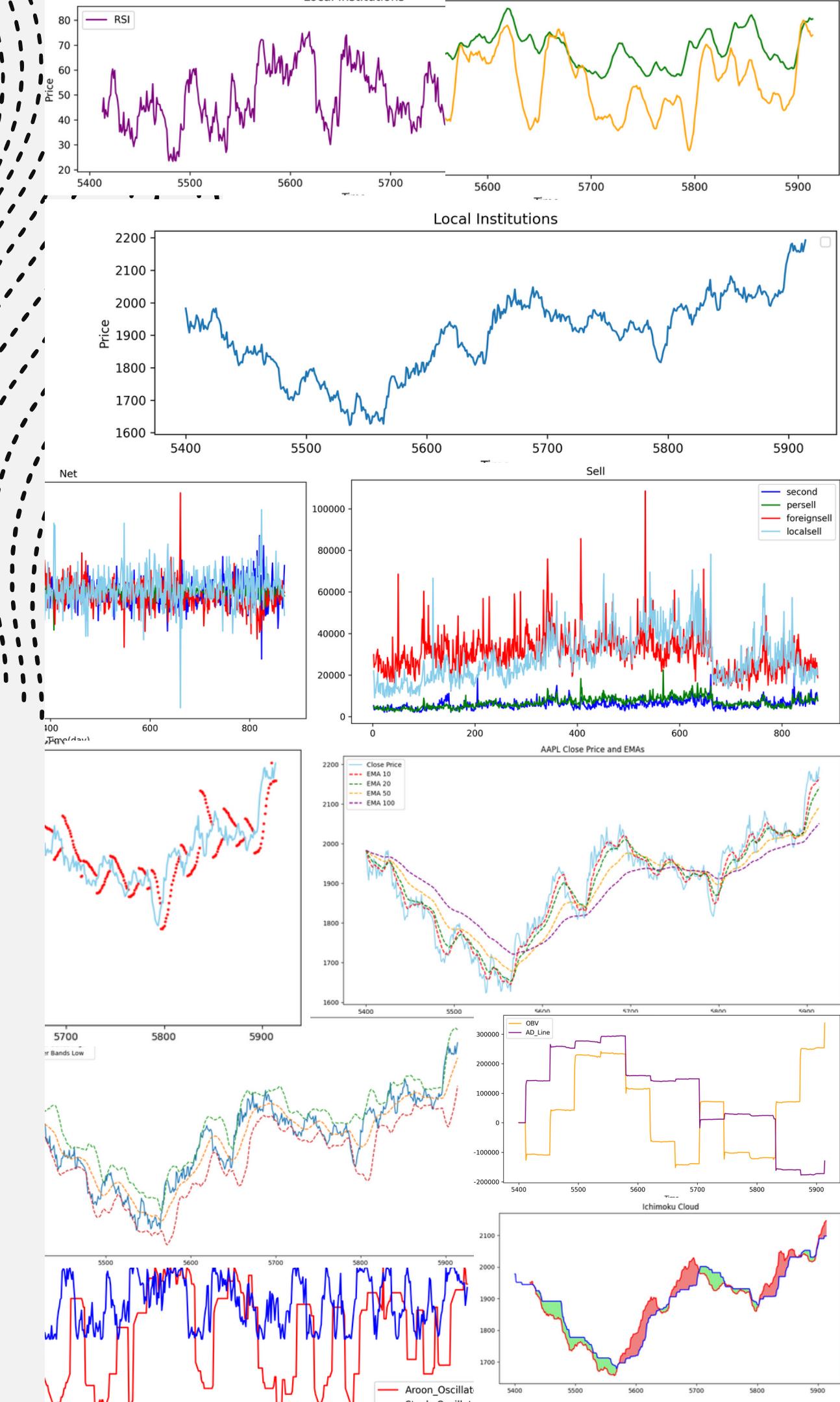
Dow Jones Industrial Average

DOWNLOAD IN YFINANCE LIBRARY



Technical Indicator

- EMA
- RSI
- Bollinger band
- Parabolic SAR
- Ichimoku cloud
- MACD
- Average Directional Movement
- On Balance Volume
- Aroon
- stochastic



# FEATURE ENGINEERING

การสกัดคุณลักษณะ คือ กระบวนการแปลง  
ข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปใช้งาน กับ  
Machine Learning

## 47 FEATURE

# NORMALIZING BY MINMAXSCALE

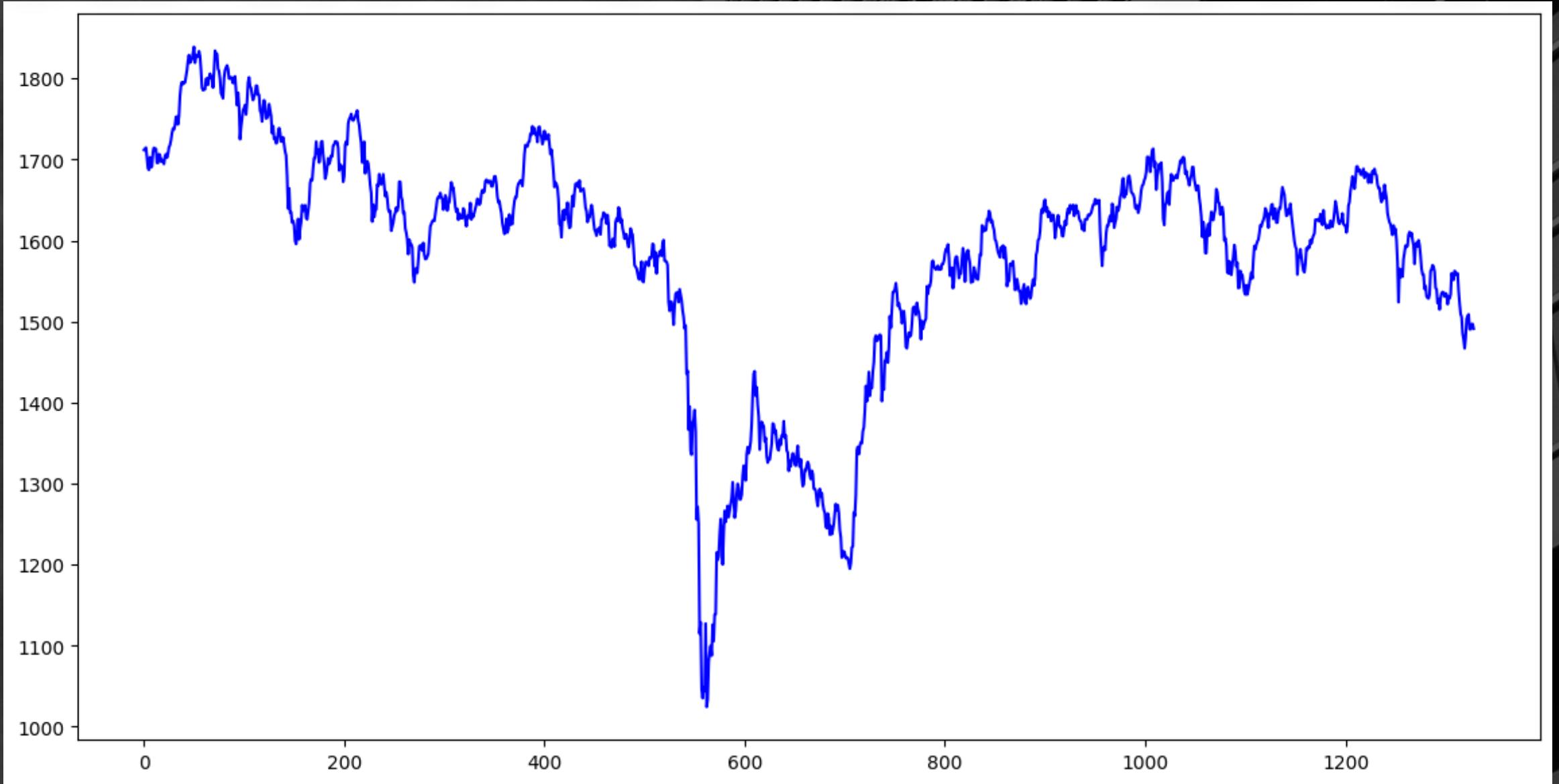
การปรับช่วงขอบเขตของข้อมูลตัวเลข  
แต่ละ Feature ให้อยู่ในช่วง [ 0, 1 ]

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

# TRAIN TEST SPLIT

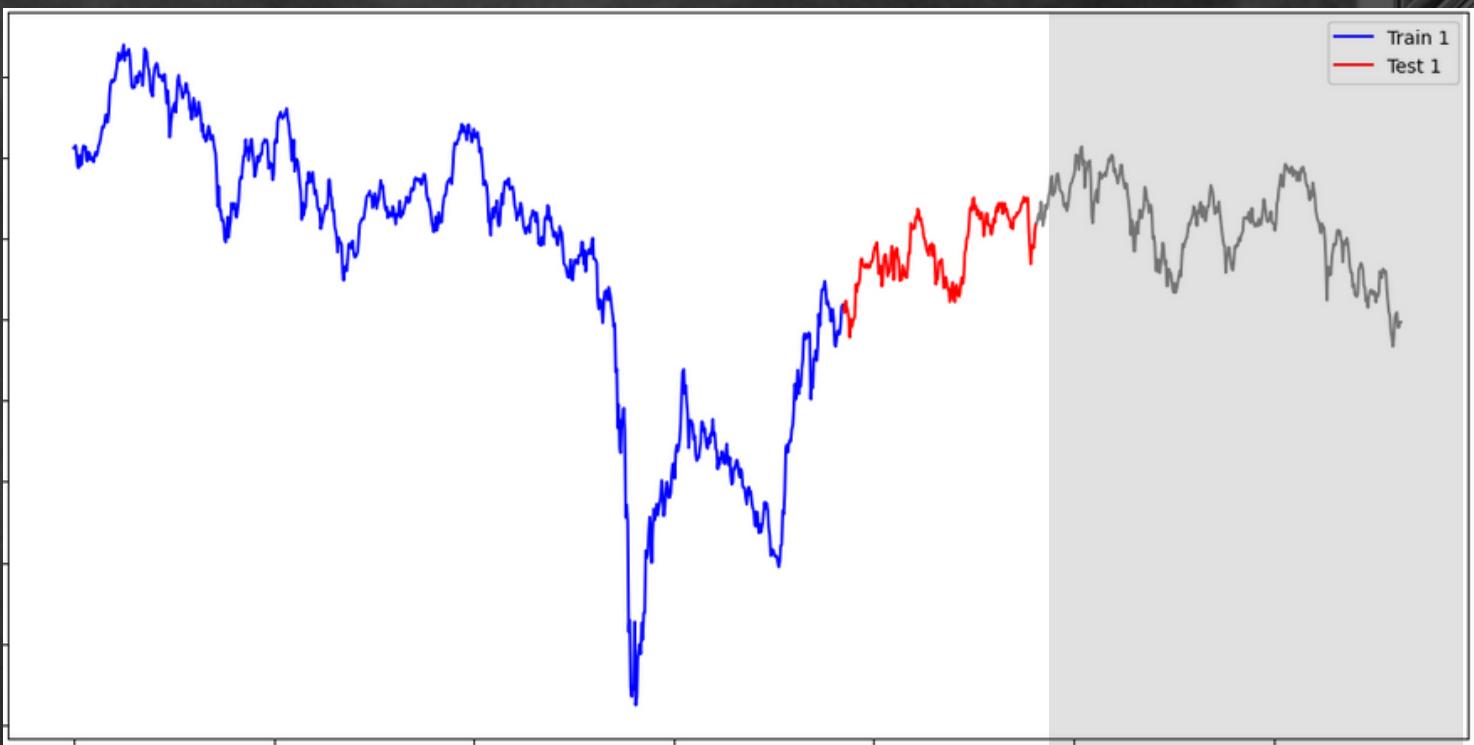
START  
**11/6/2017**

END  
**7/12/2023**

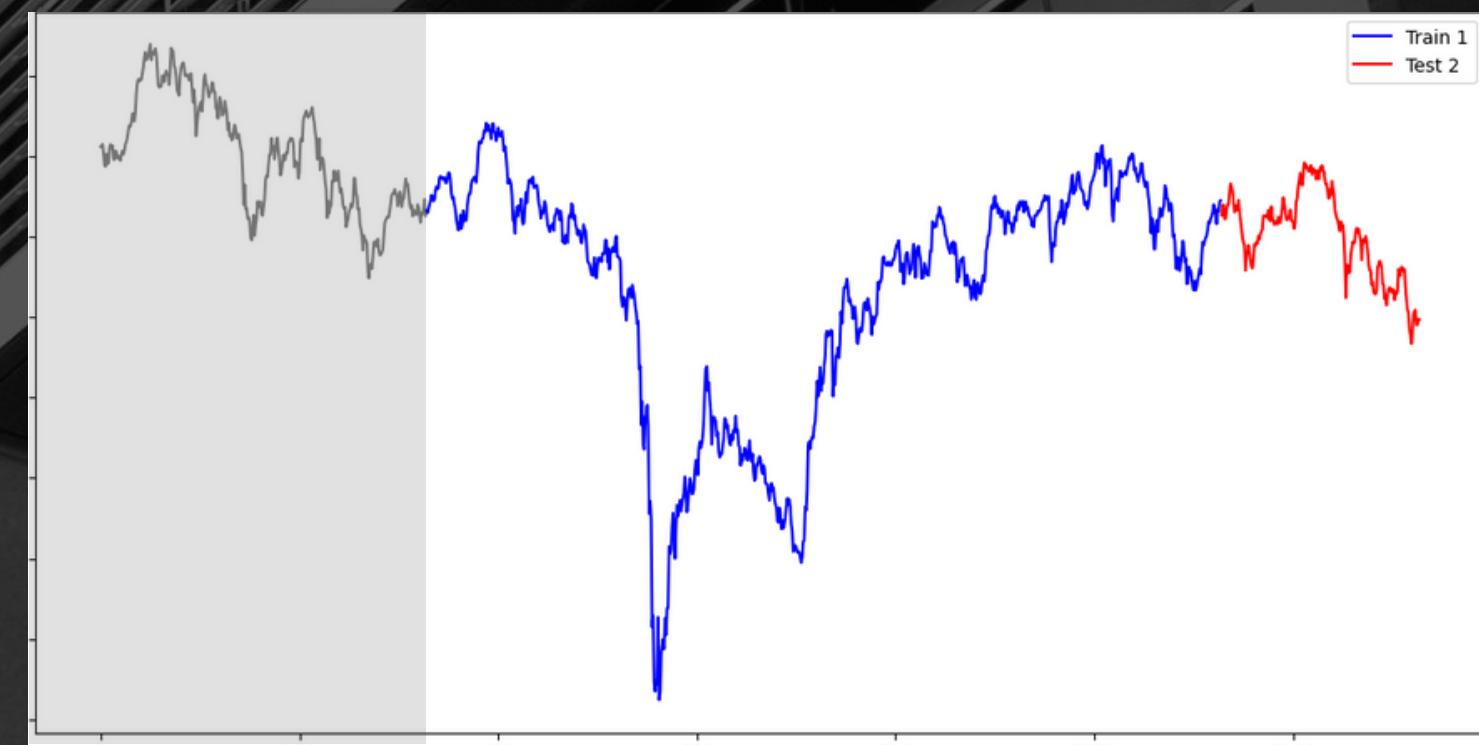


| 1300 ข้อมูล

# TRAIN TEST SPLIT

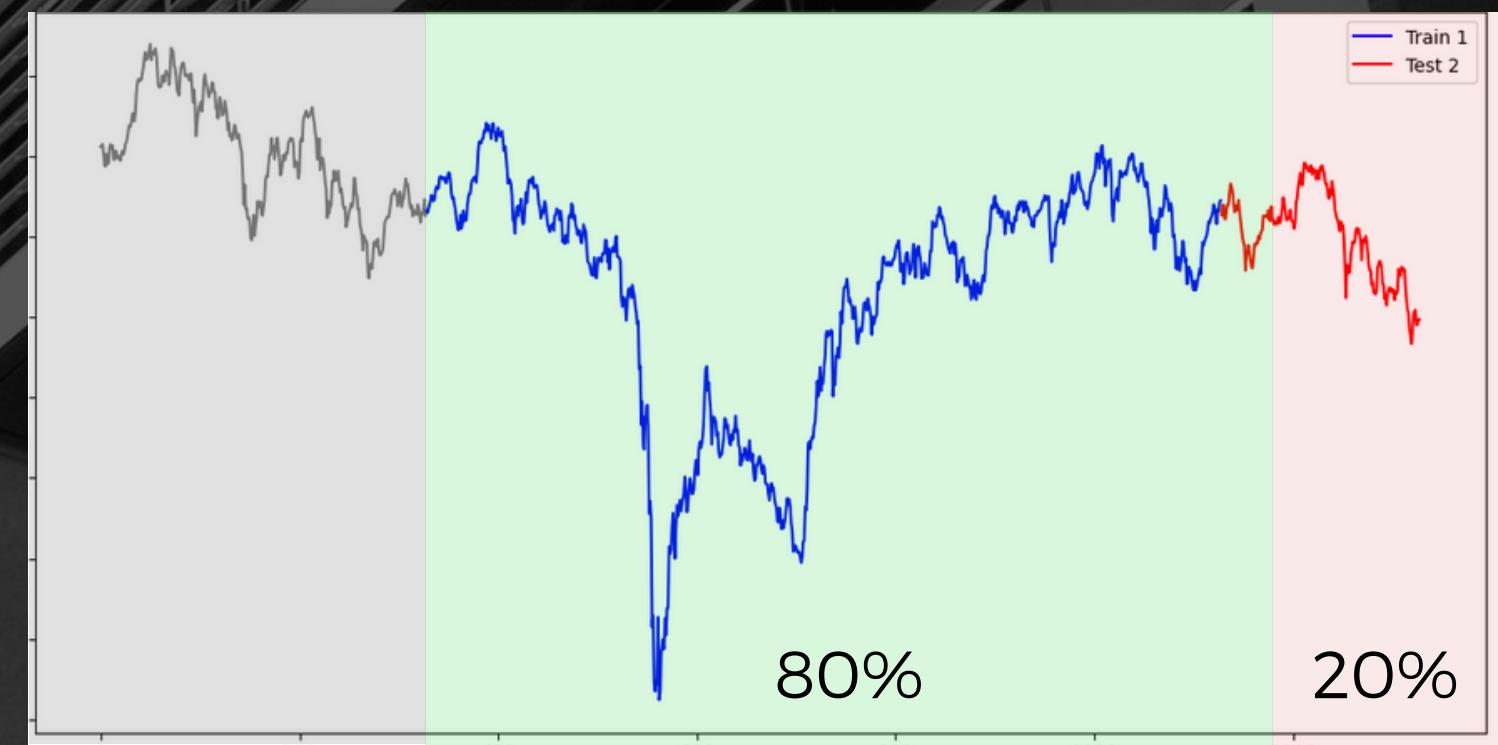
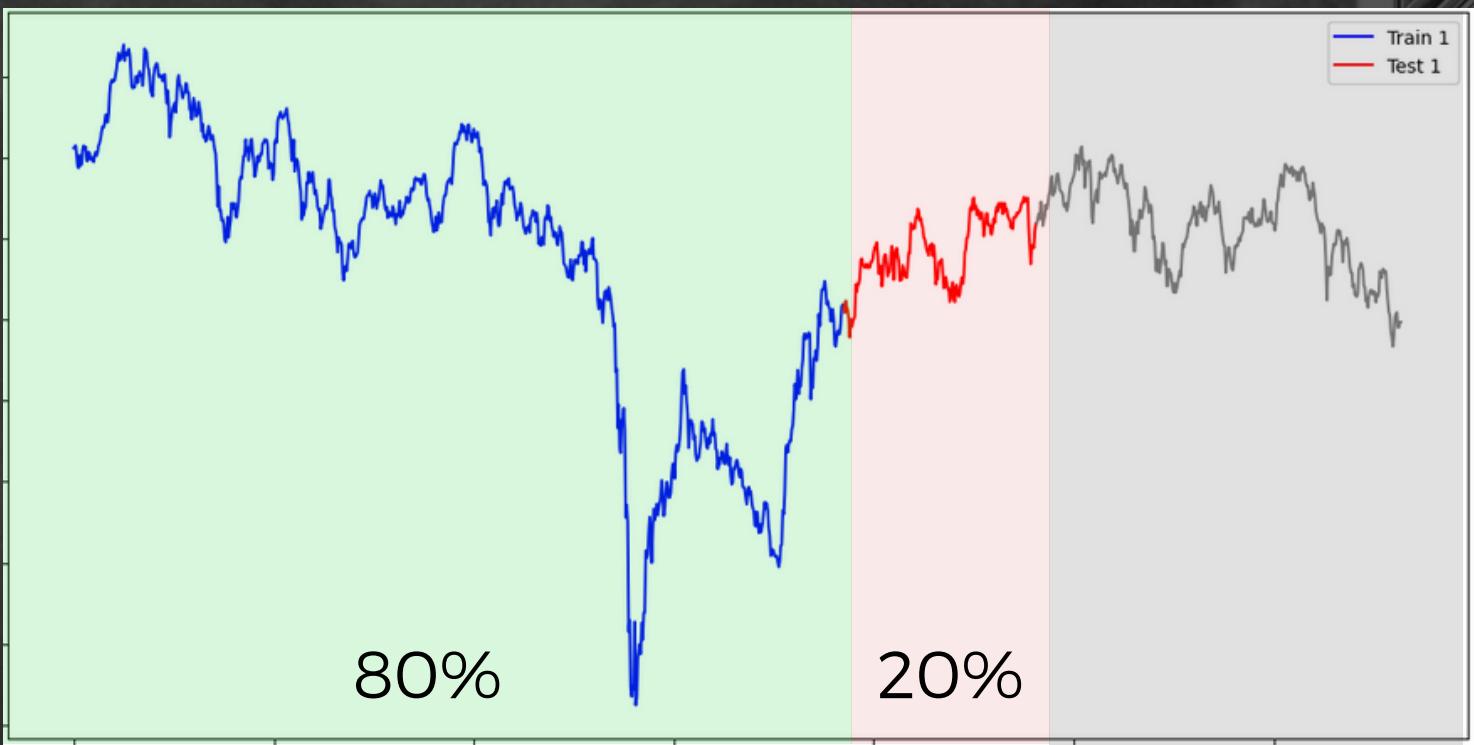


| 1000 นับจากเริ่มต้น



| 1000 นับจากข้อมูลล่าสุด

# TRAIN TEST SPLIT



| Train 80%  
| Test 20%



## Decision Tree Regression



## Random Forest Regression

n\_estimators = 100  
randomstage = 100



## Support Vector Regression

epsilon=0.1  
C=15  
kernel='linear'  
gamma='scale'



## Long Short-Term Memory

01

02

03

04



# FEATURE EXTRACTION



**SelectFromModel**

scikit learn command



**Lasso**

model

**23 FEATURE**



# MSE

## MEAN SQUARE ERROR

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

ค่าเฉลี่ยของผลต่างสองยกกำลังสอง ระหว่าง  
ค่าที่เป็นเลขจำนวนจริงที่เป็นผลเฉลยกับค่าที่  
เกิดจากการคำนายนของ Model

# TRADING

PORT = 1M BATH

$$\text{ผลต่าง} = \text{Close}_{t+1} - \text{Close}_t$$

ผลต่างมากกว่า 0



ซื้อ

ผลต่างน้อยกว่า 0



ขาย

ค่าธรรมเนียม 0.15%

BUY AND HOLD

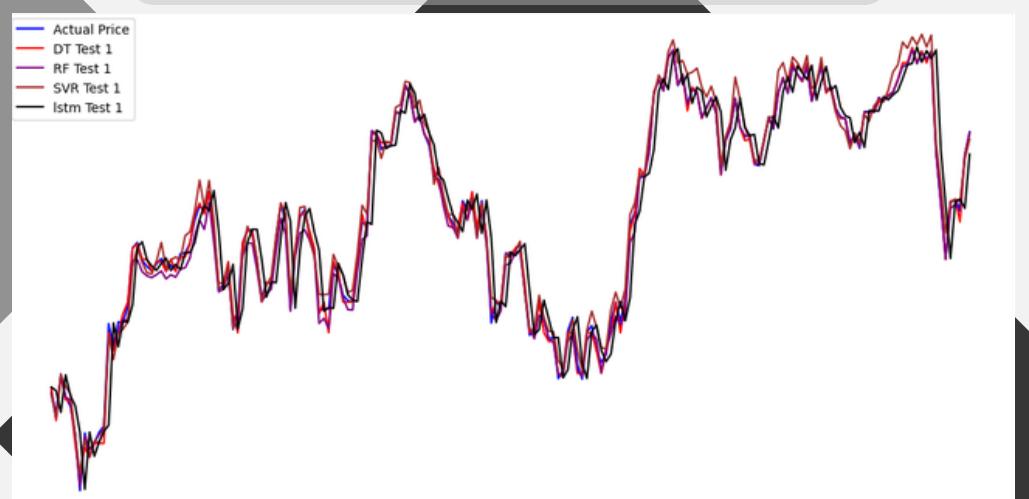
ACTUAL CLOSE

# RESULT

ប័ណ្ណលម្អិតនៅក្នុង

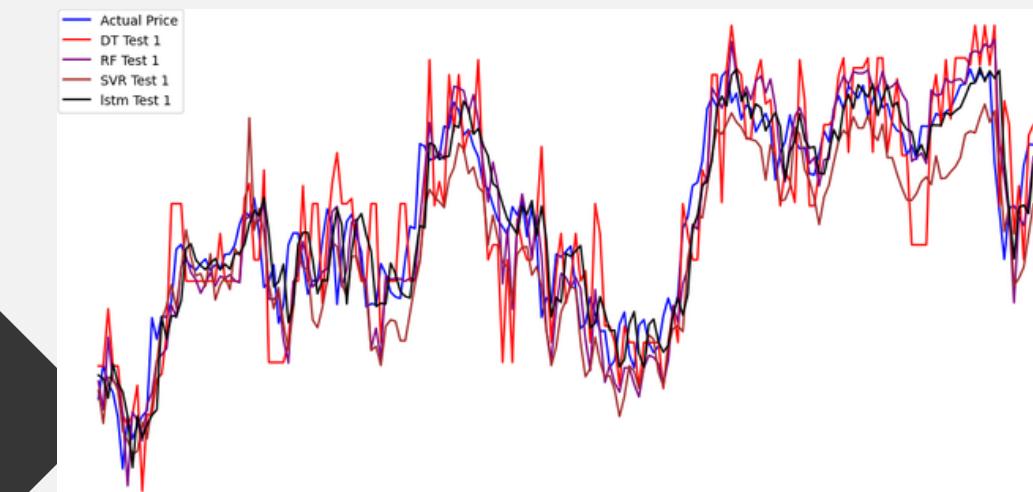
MSE

CLOSE <sub>T</sub>	TEST
DT	4.319
RF	11.481
SVR	18.584
LSTM	143.438



MSE

CLOSE <sub>T+1</sub>	TEST
DT	609.406
RF	291.979
SVR	426.219
LSTM	150.961



%PROFIT

RETURN FROM  
TRADING

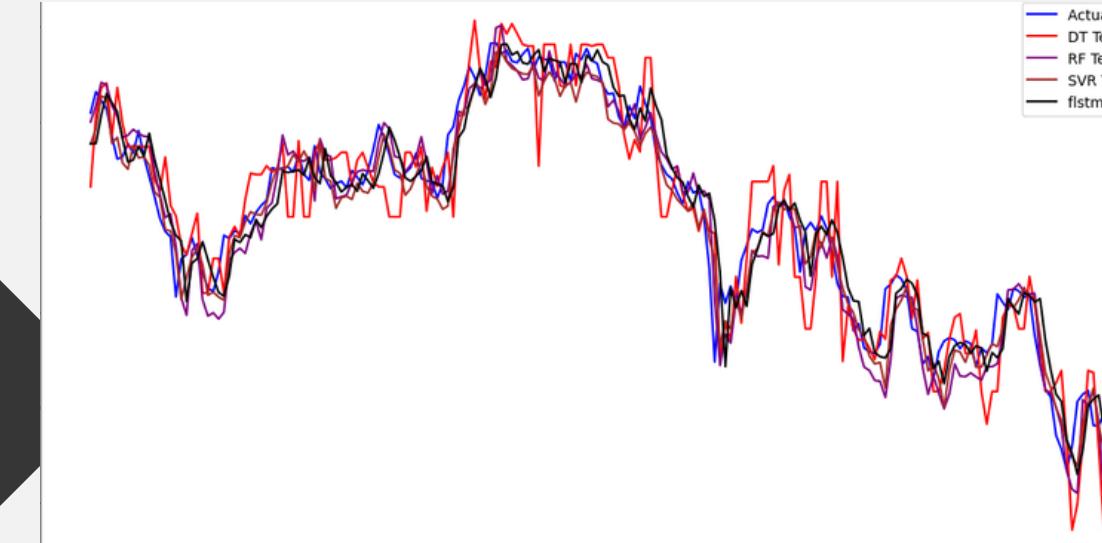
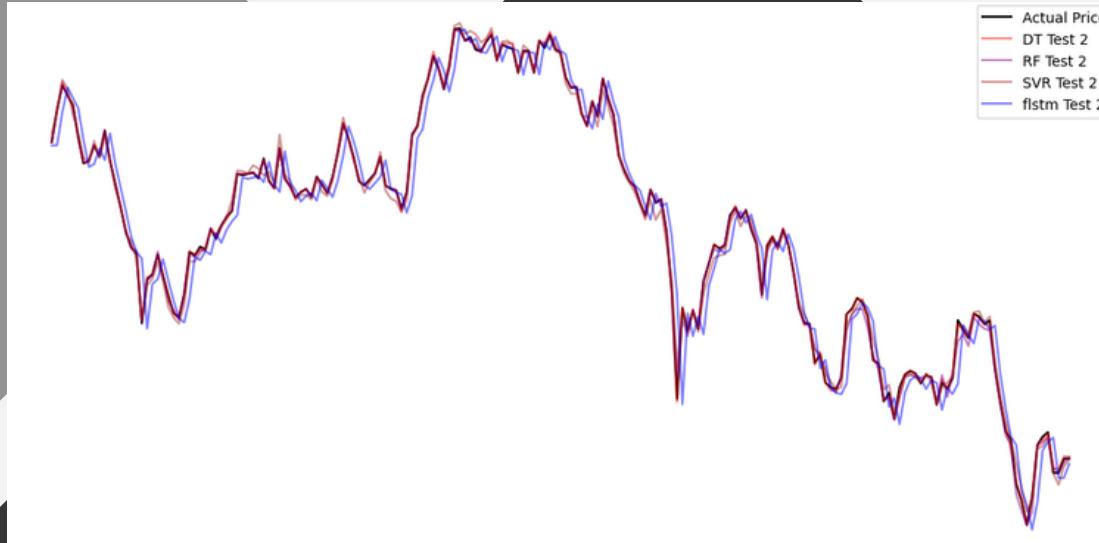
DT	-11.47%
RF	-11.70%
SVR	6.69%
LSTM	6.05%
ACTUAL	50.04%
B&H	6.36%

# RESULT

ข้อมูลชุดสอง

MSE	
CLOSE <sub>T</sub>	TEST
DT	0.739
RF	4.285
SVR	10.696
LSTM	231.800

MSE	
CLOSE <sub>T+1</sub>	TEST
DT	453.122
RF	257.886
SVR	183.839
LSTM	193.136



%PROFIT

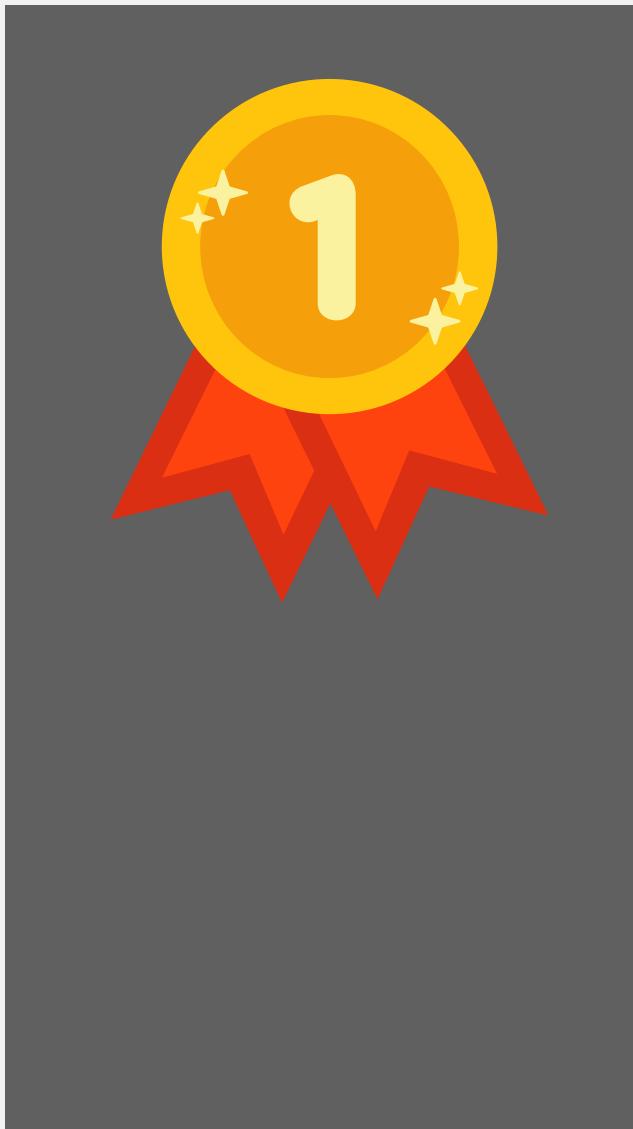
RETURN FROM  
TRADING

DT	-12.57%
RF	-15.58%
SVR	-2.60%
LSTM	-11.95%
ACTUAL	38.54%
B&H	-8.99%

# CONCLUSION



SUPPORT VECTOR  
REGRESSION



LONG SHORT-  
TERM MEMORY



DECISION TREE  
REGRESSION



RANDOM FOREST  
REGRESSION



# limitations

---

- เวลา
- ข้อมูล
- จำนวนครั้ง

**train test**

# Future

---

- Hyper parameter tuning
- LSTM parameter
- ensemble model
- TimeFrame Hour

# Q & A

Presentation by Nattapong Pakdeeviboon



2B Explore KMUTT

**THANK YOU  
FOR WATCHING**

