



Time Series Modeling: Crude oil price prediction

Evgeniya Dontsova

October 24, 2019



Dataset:

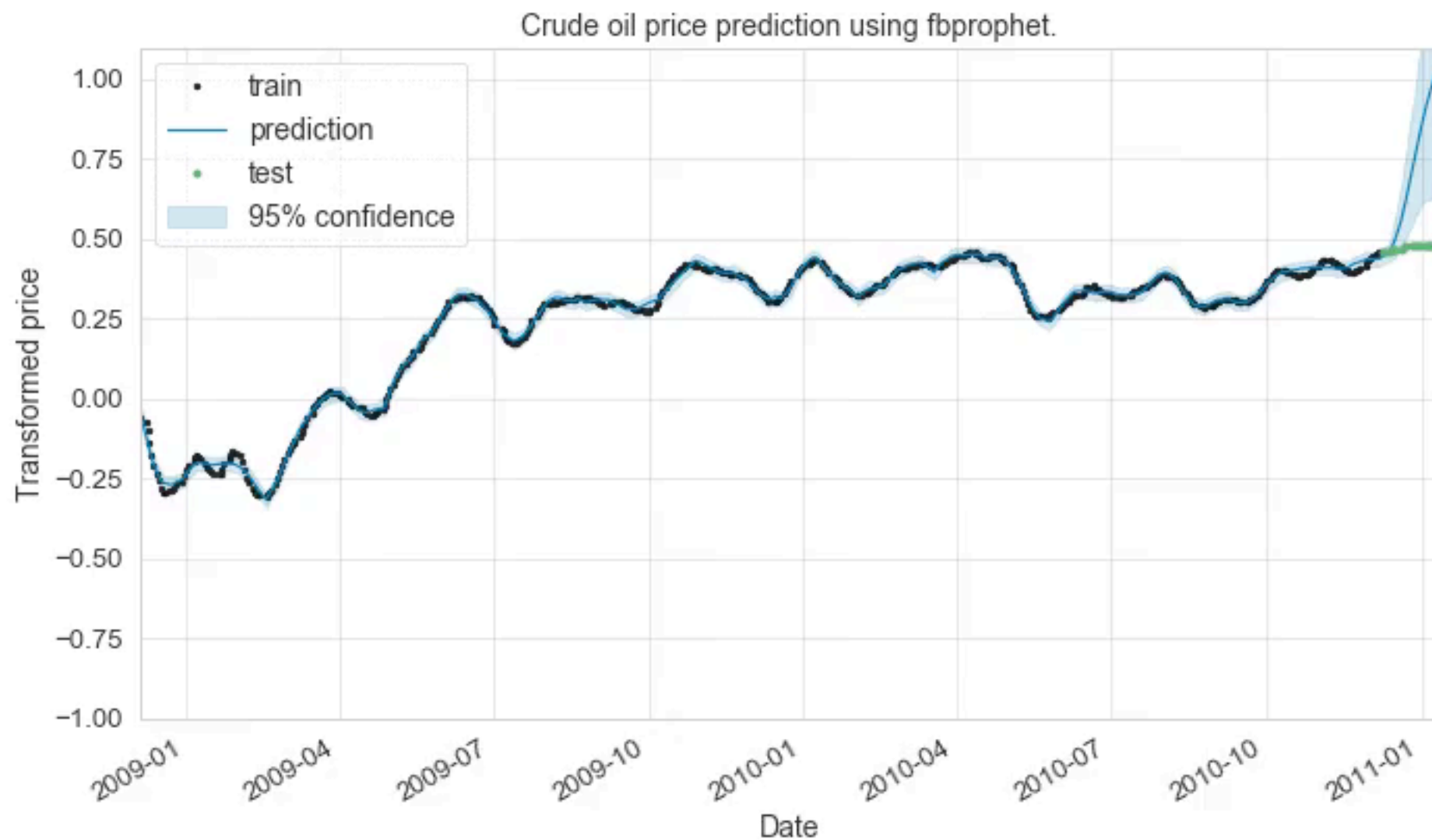
WTI (West Texas Intermediate) crude oil price



Daily data: 8501 observations, > 33 years, starting from 1986-03-01

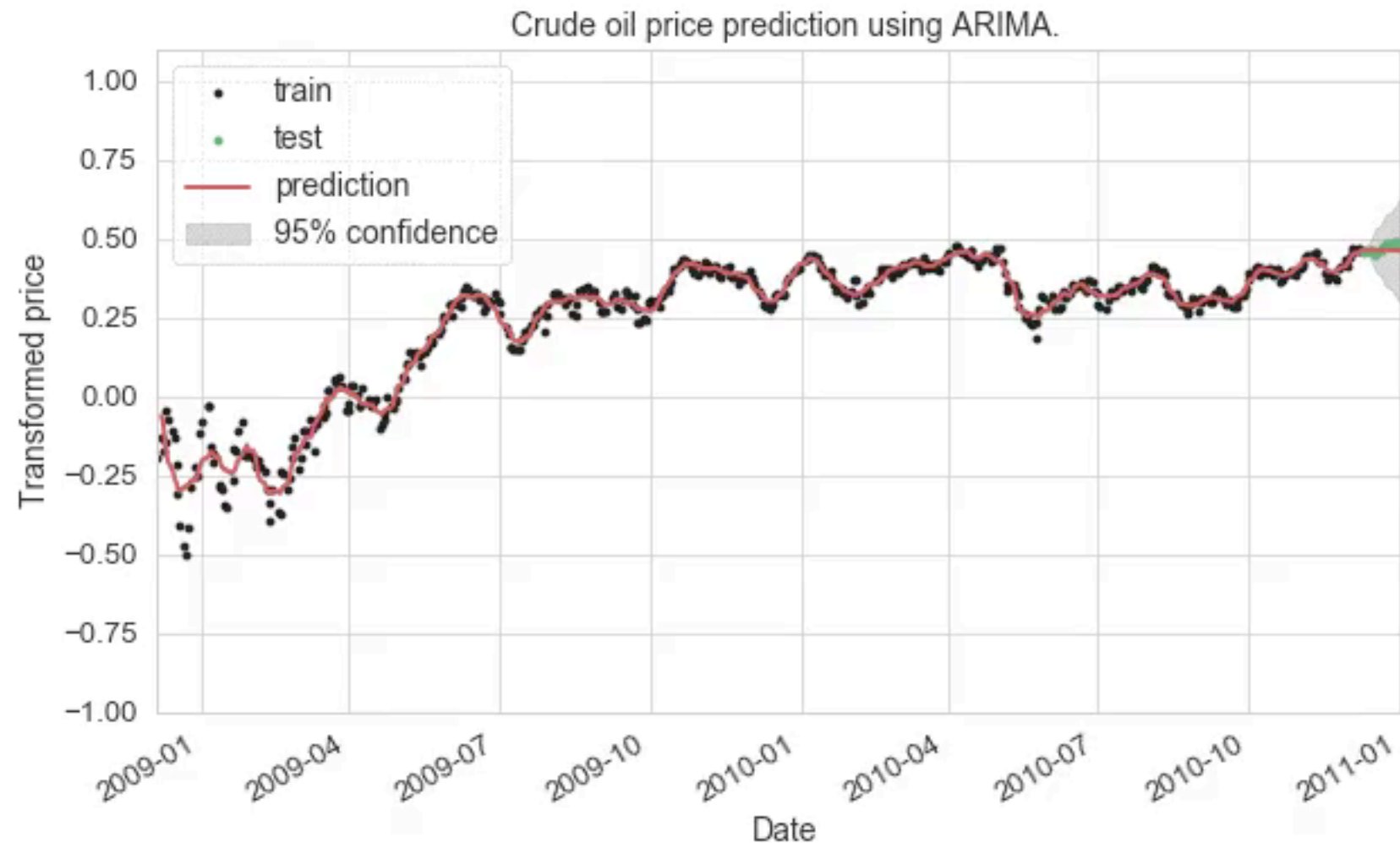
Weekly data: 1762 observations

Using Facebook Prophet Model: Train size = 2 years, Test size = 1 month



Mean Average Percentage Error (MAPE): 31.8 % +/- 28.6 %

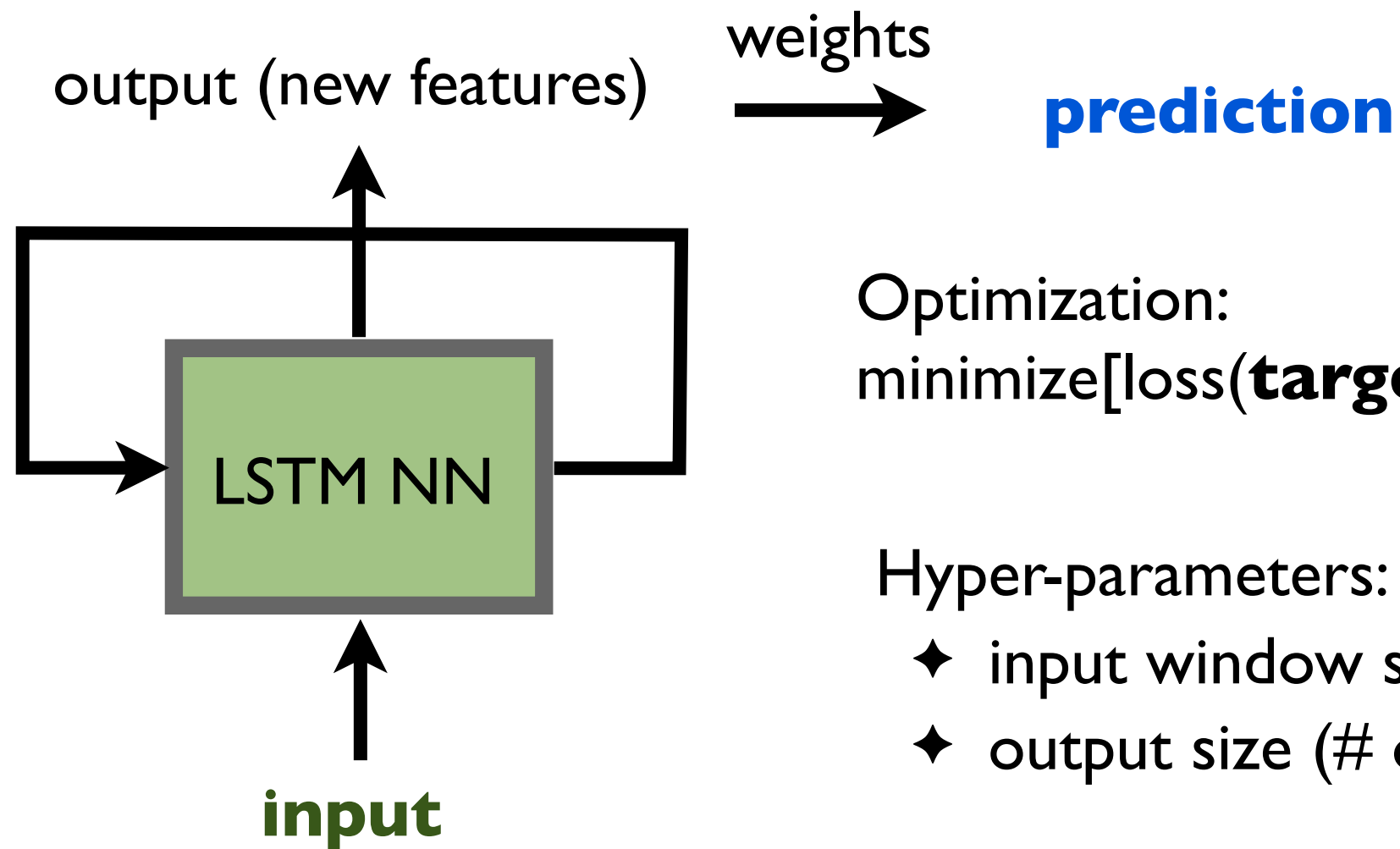
Using ARIMA Model: Train size = 2 years, Test size = 1 month



Mean Average Percentage Error (MAPE): 11.5 % +/- 16.0 %

A = Auto, R = Regressive, I = Integrated, M = Moving, A = Average

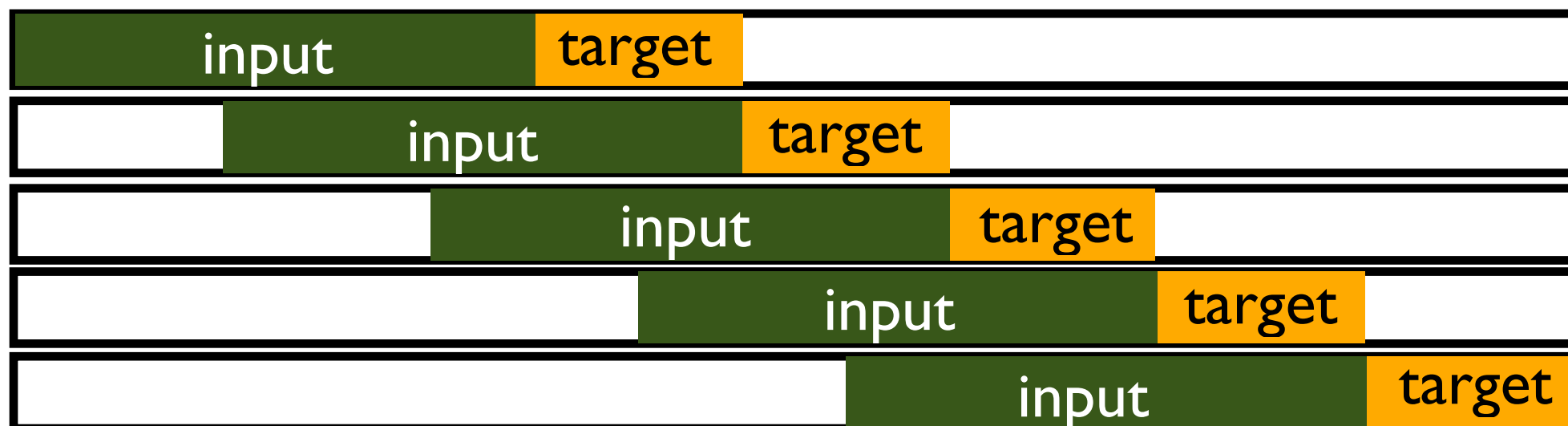
Long Short-Term Memory (LSTM) Neural Network (NN)



Optimization:
 $\text{minimize}[\text{loss}(\mathbf{target}, \mathbf{prediction})]$

Hyper-parameters:

- ♦ input window size
- ♦ output size (# of new features)



Regression Problem: Predict 5 days

1 LSTM model:



Hyper-parameter:
input window
optimal = 25

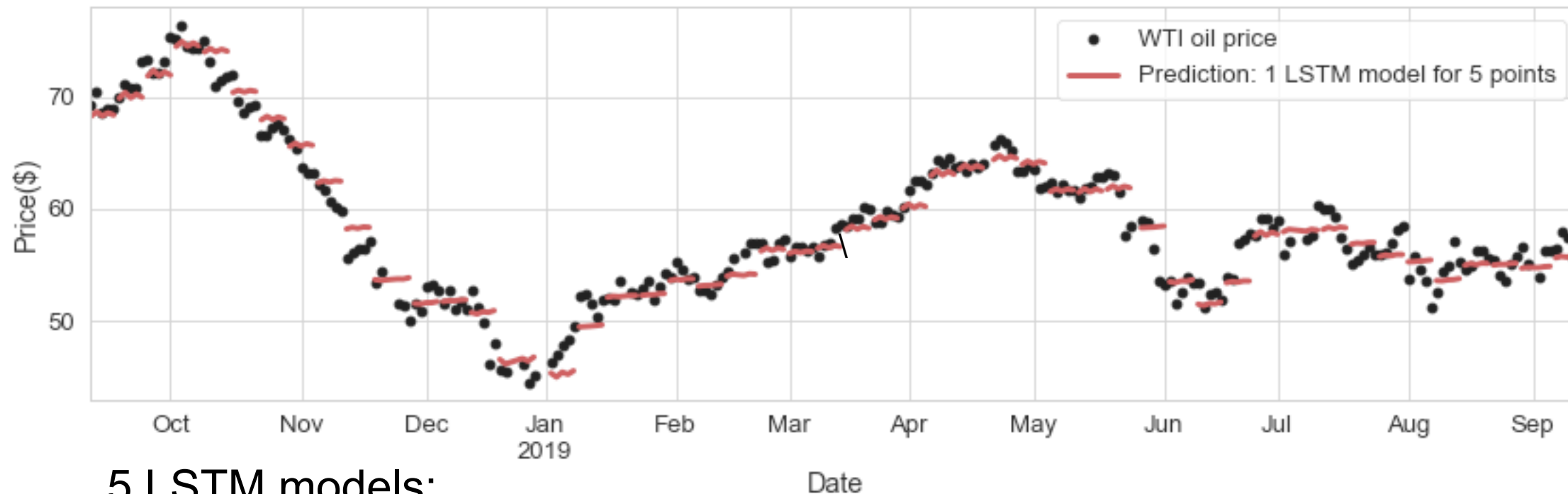
5 LSTM models:



Hyper-parameters:
input windows
optimal =
[29, 20, 17, 18, 14]

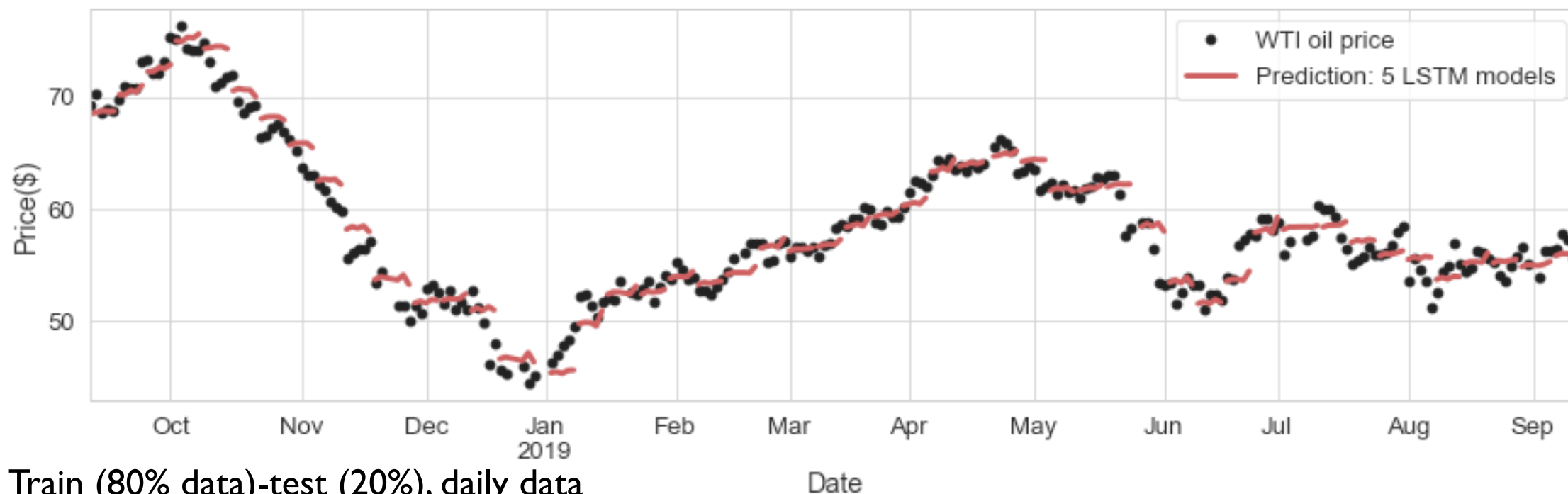
Regression Problem: Predict 5 days (test data)

1 LSTM model:



MAPE (test data) :
1.98 +/- 1.96

5 LSTM models:



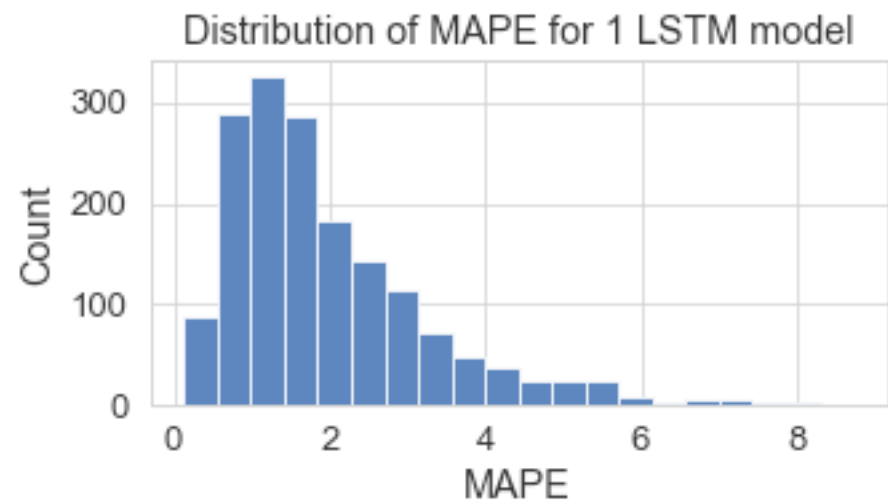
MAPE (test data) :
1.91 +/- 1.91

For comparison
ARIMA MAPE:
4.80 +/- 3.09

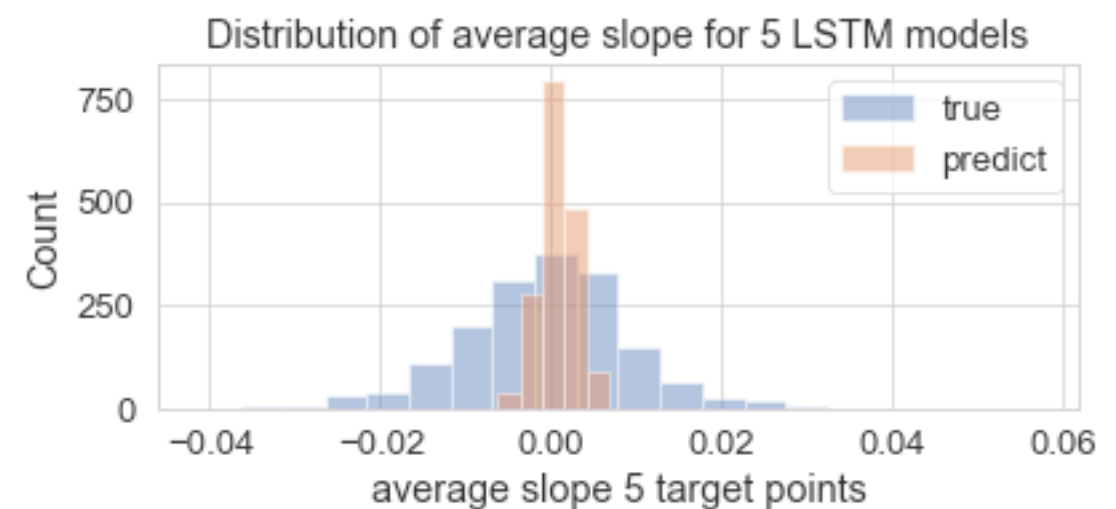
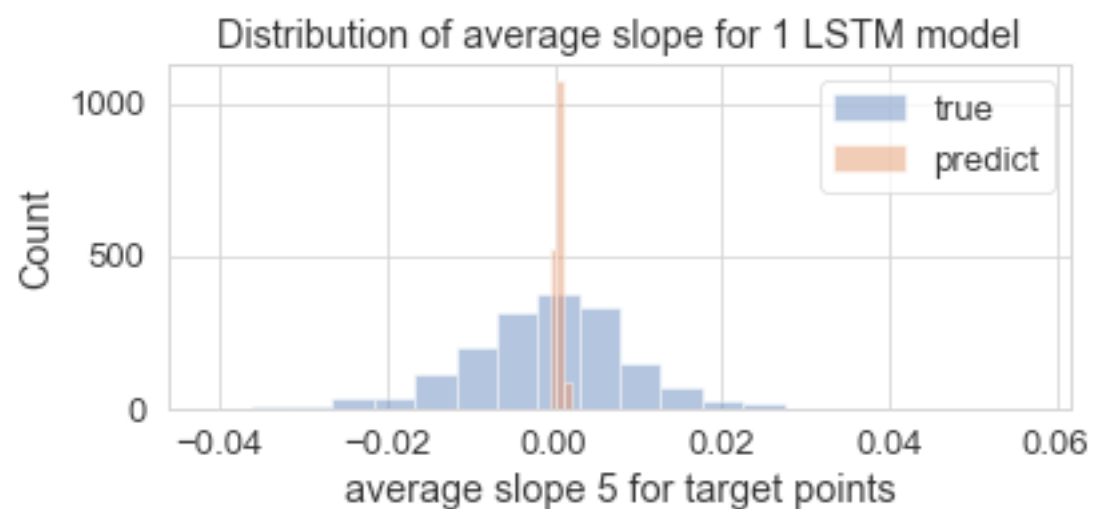
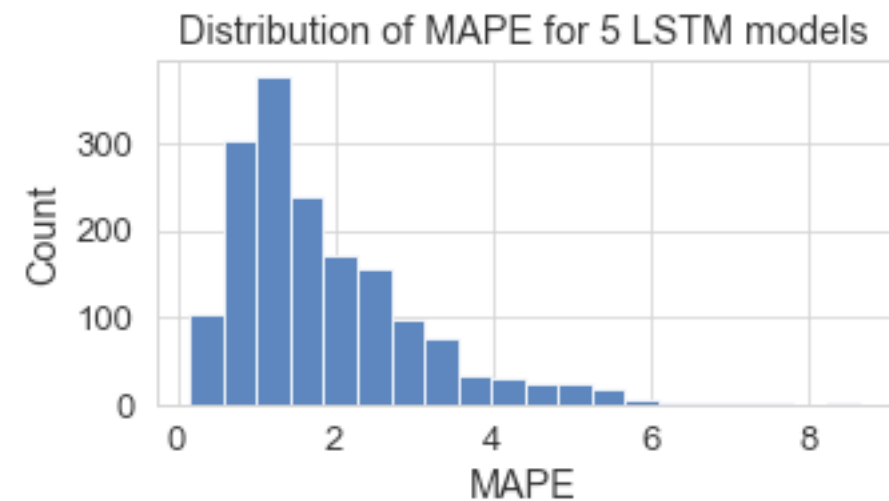
Train (80% data)-test (20%), daily data

Regression Problem: 5 LSTM models (test data)

1 LSTM model:

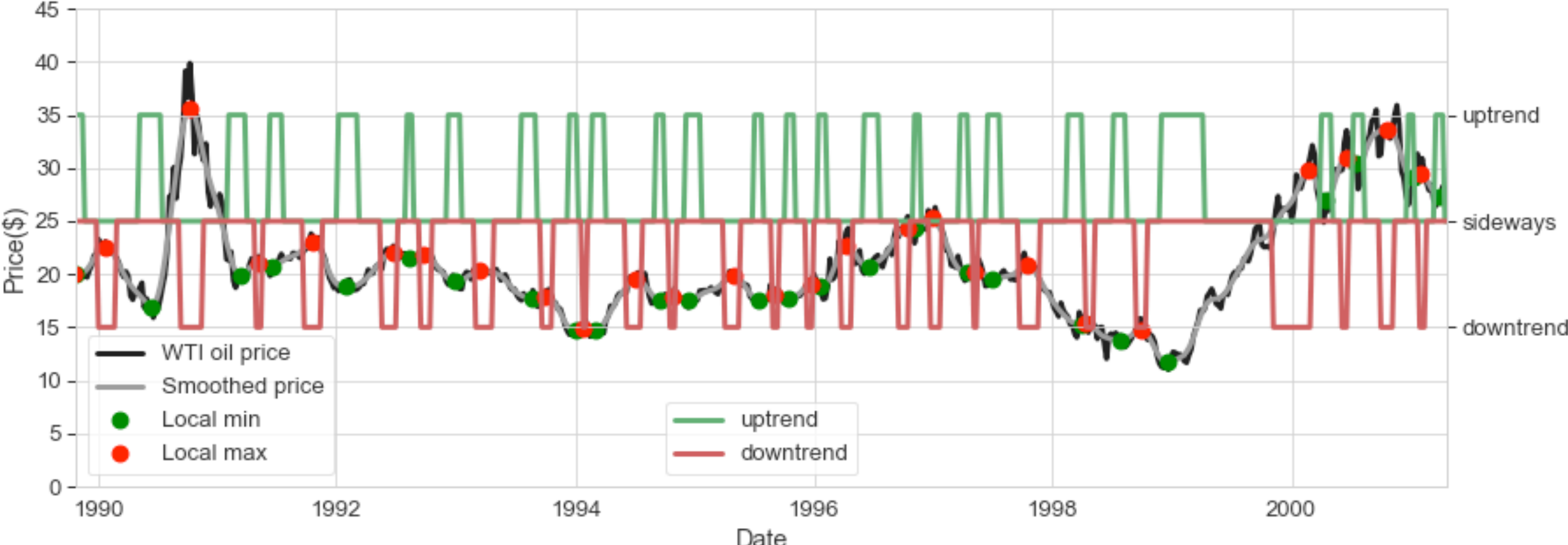


5 LSTM models:

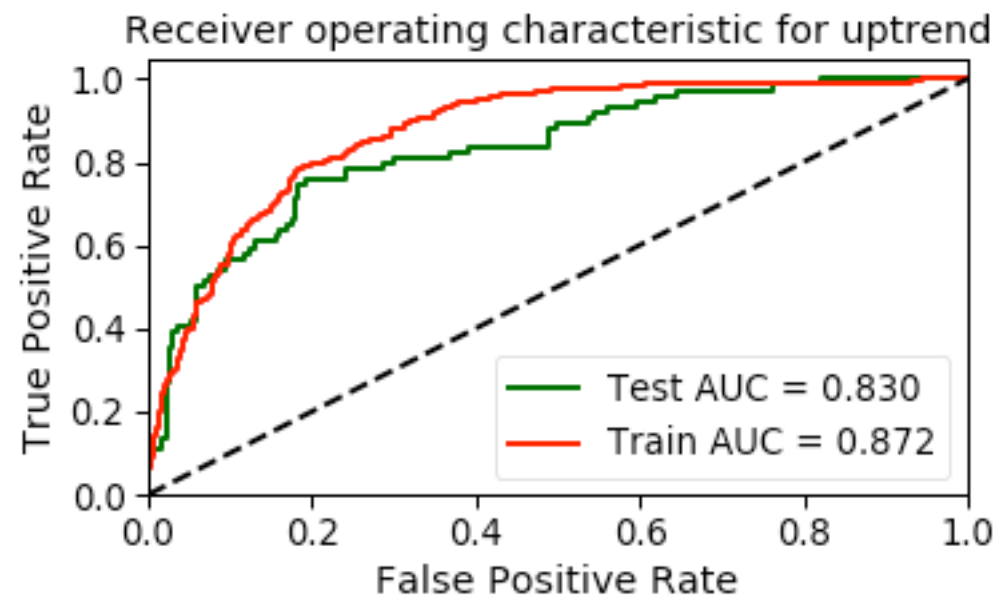


How to predict behavior/trend more accurately?

and downtrend target



Classification Problem: Predict uptrend



Model: LSTM NN

Input:

1. smoothed price
2. predicted probabilities (another model)

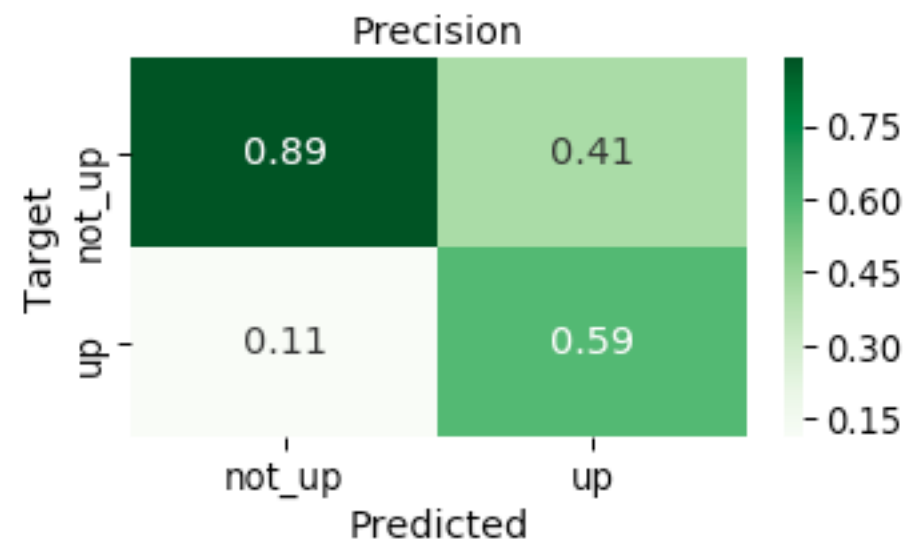
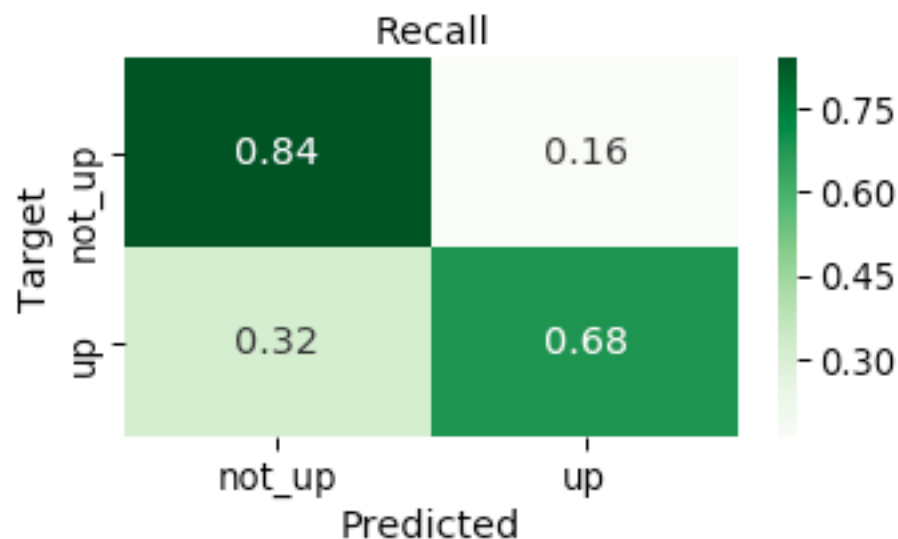
Hyper-parameters:

input window = **80**
hidden size = 4
patience = 30

Model Stability:

10-fold slide-window CV: Train AUC = 0.855 +/- 0.051
Test AUC = 0.832 +/- 0.084

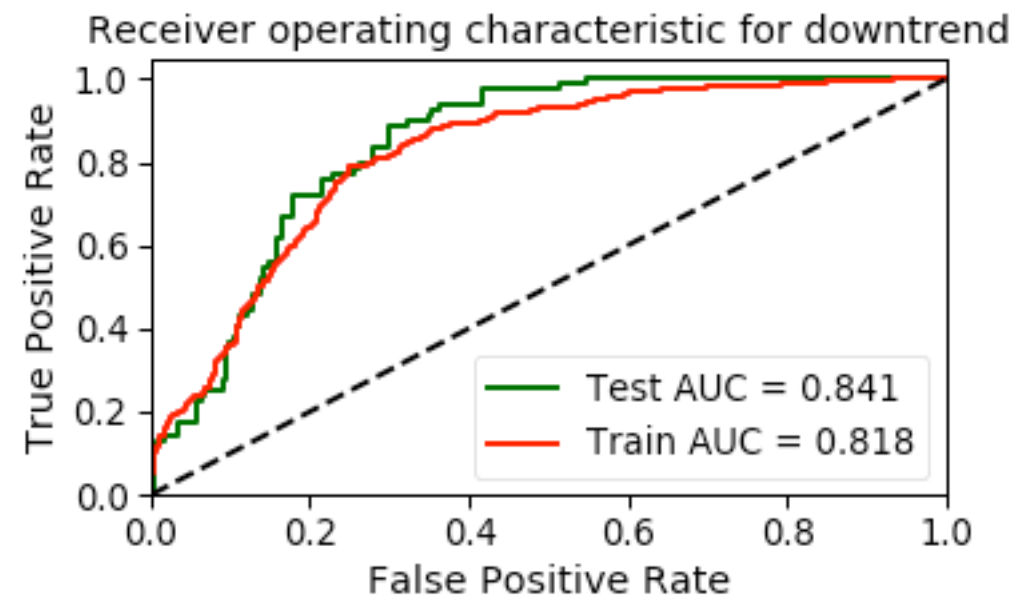
Rescaled confusion matrices:



Shift threshold probability by optimizing f1-score:

threshold = 0.56
accuracy = 0.80

Classification Problem: Predict downtrend



Model: LSTM NN

Input:

1. smoothed price
2. predicted probabilities (another model)

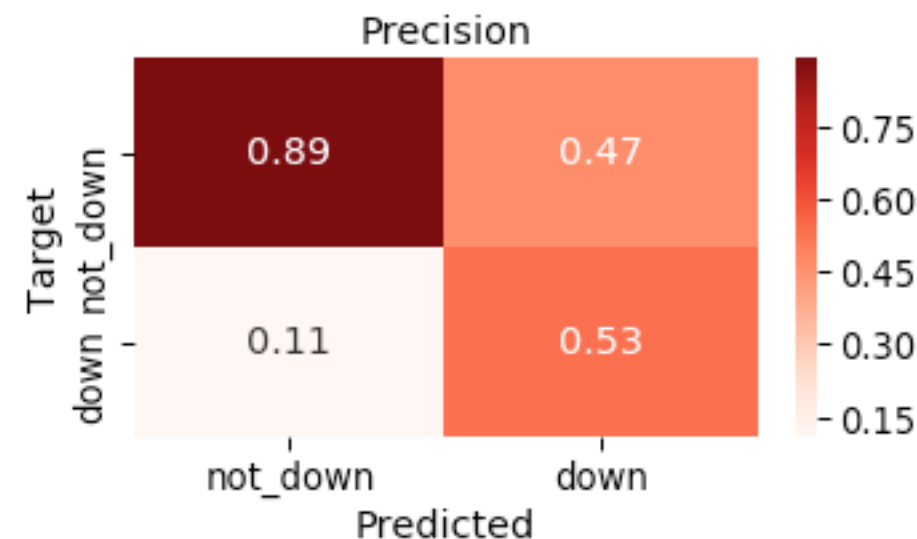
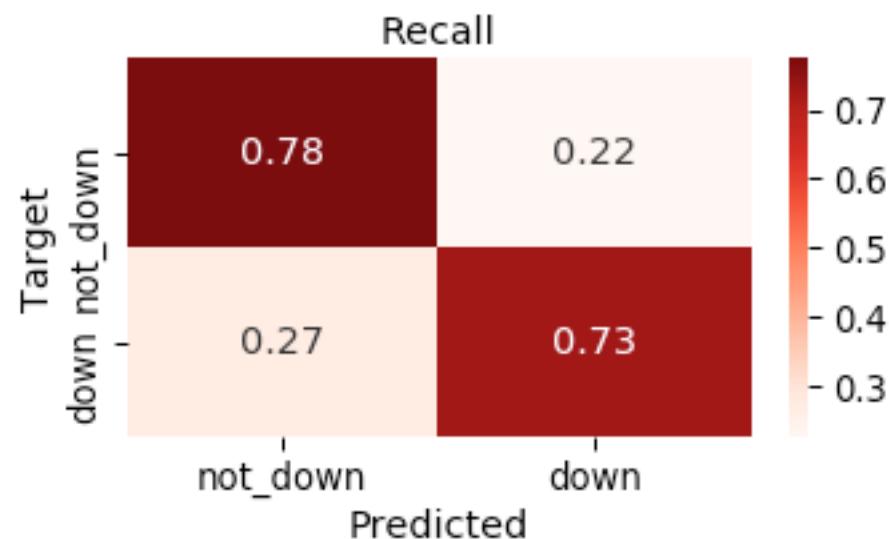
Hyper-parameters:

input window = **10**
hidden size = 4
patience = 30

Model Stability:

10-fold slide-window CV: Train AUC = 0.825 +/- 0.039
Test AUC = 0.825 +/- 0.084

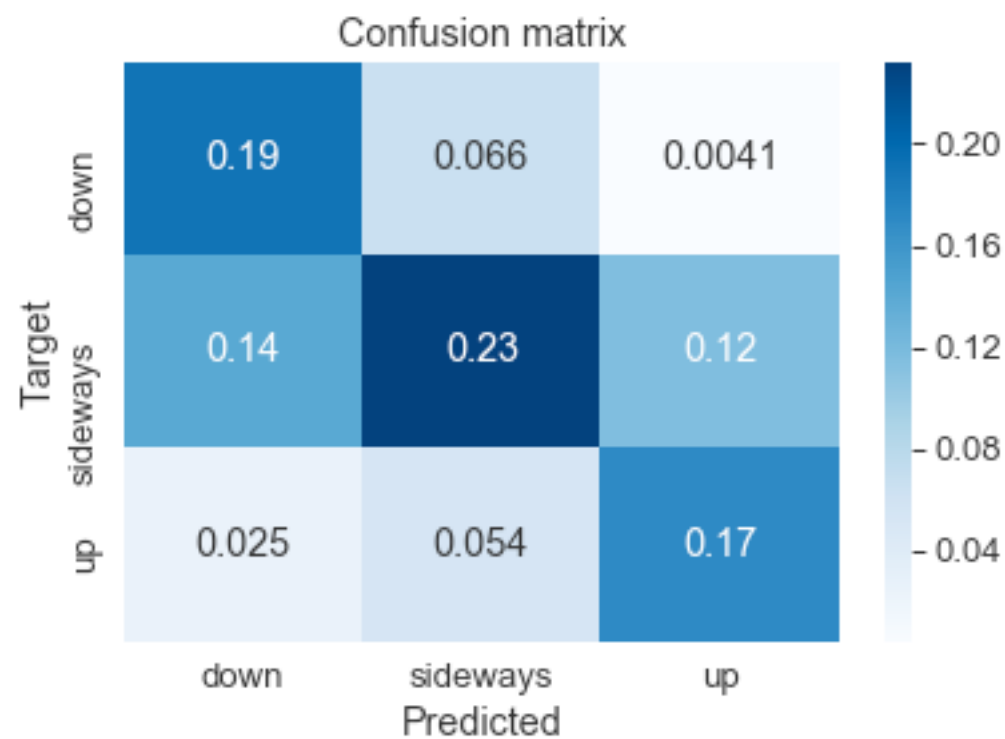
Rescaled confusion matrices:



Shift threshold probability
by optimizing f1-score:

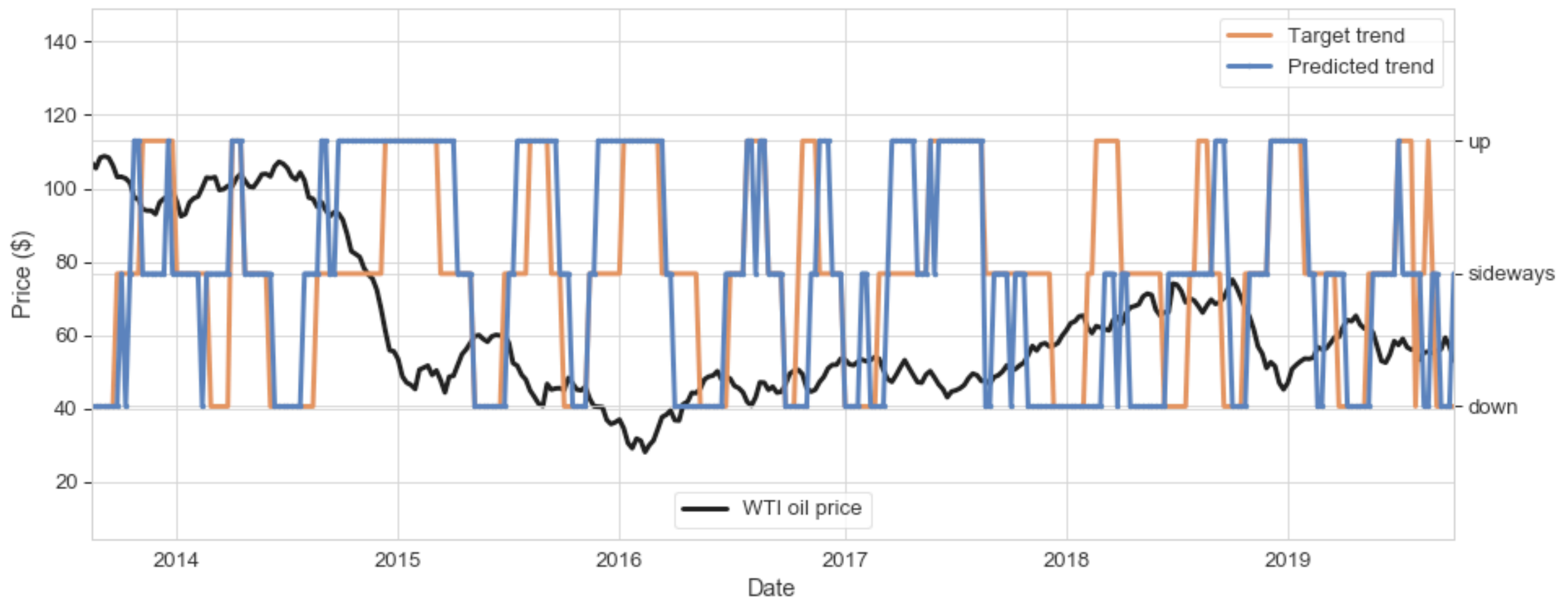
threshold = 0.24
accuracy = 0.76

Classification Problem: Combining uptrend and downtrend models



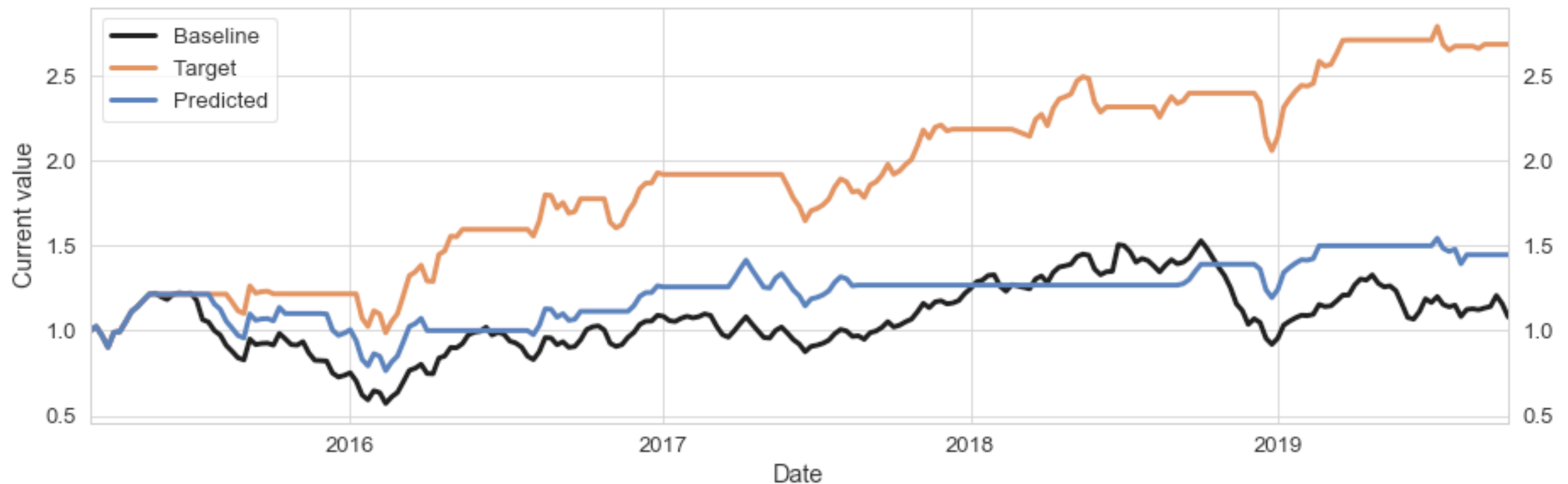
	precision	recall	f1-score	support
up	0.53	0.73	0.62	63
sideways	0.66	0.47	0.55	118
down	0.59	0.68	0.63	60
accuracy			0.59	241

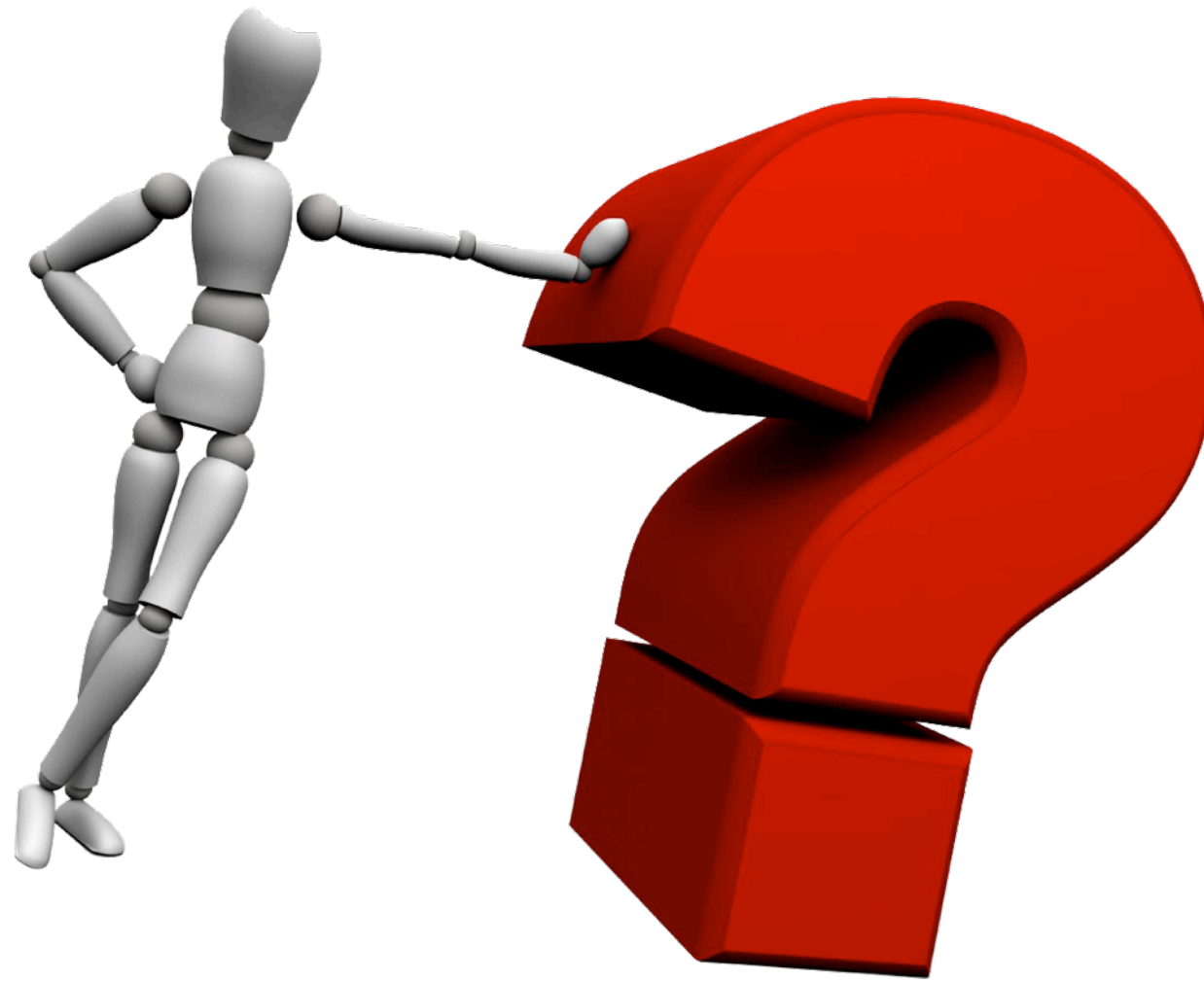
Classification Problem: Direct comparison for the test data



Model Assessment:

use uptrend/downtrend signals to buy/sell





Contact information:

Evgeniya Dontsova

Ph. D. in Mechanical Engineering



jennyakateva@gmail.com



<https://www.linkedin.com/in/evgeniya-dontsova-98432b195/>



<https://github.com/evgeniya1>