



# OPTIVER-TRADING AT THE CLOSE

DIVYAMAN SINGH RAWAT

Stock ID	Input sequence for target prediction at time ID K					Target prediction at time K
<b>0</b>	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 0 at K</i>
<b>1</b>	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 1 at K</i>
...	...	...	...	...	...	...
<b>199</b>	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 199 at K</i>

Stock ID	Input sequence for target prediction at time ID K					Target prediction at time K
0	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 0 at K</i>
1	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 1 at K</i>
...	...	...	...	...	...	...
199	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 199 at K</i>
Stock ID	Input sequence for target prediction at time ID (K+1)					Target prediction at time (K + 1)
0	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b><i>Target at K</i></b>	<i>Target prediction for stock ID 0 at (K+1)</i>
1	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b><i>Target at K</i></b>	<i>Target prediction for stock ID 1 at (K+1)</i>
...	...	...	...	...	...	...
199	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b><i>Target at K</i></b>	<i>Target prediction for stock ID 199 at (K+1)</i>

Stock ID	Input sequence for target prediction at time ID K					Target prediction at time K
0	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 0 at K</i>
1	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 1 at K</i>
...	...	...	...	...	...	...
199	<i>Target at (K - N)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at (K - 2)</i>	<i>Target at (K - 1)</i>	<i>Target prediction for stock ID 199 at K</i>
Stock ID	Input sequence for target prediction at time ID (K+1)					Target prediction at time (K + 1)
0	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b>Target at K</b>	<i>Target prediction for stock ID 0 at (K+1)</i>
1	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b>Target at K</b>	<i>Target prediction for stock ID 1 at (K+1)</i>
...	...	...	...	...	...	...
199	<i>Target at (K - N + 1)</i>	<i>Target at (K - N + 2)</i>	...	<i>Target at (K - 1)</i>	<b>Target at K</b>	<i>Target prediction for stock ID 199 at (K+1)</i>
Stock ID	Input sequence for target prediction at time ID (K+2)					Target prediction at time (K + 2)
0	<i>Target at (K - N + 2)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at K</i>	<b>Target at (K+1)</b>	<i>Target prediction for stock ID 0 at (K+2)</i>
1	<i>Target at (K - N + 2)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at K</i>	<b>Target at (K+1)</b>	<i>Target prediction for stock ID 1 at (K+2)</i>
...	...	...	...	...	...	...
199	<i>Target at (K - N + 2)</i>	<i>Target at (K - N + 1)</i>	...	<i>Target at K</i>	<b>Target at (K+1)</b>	<i>Target prediction for stock ID 199 at (K+2)</i>

# METHODS OF SEQUENCE CONSTRUCTION

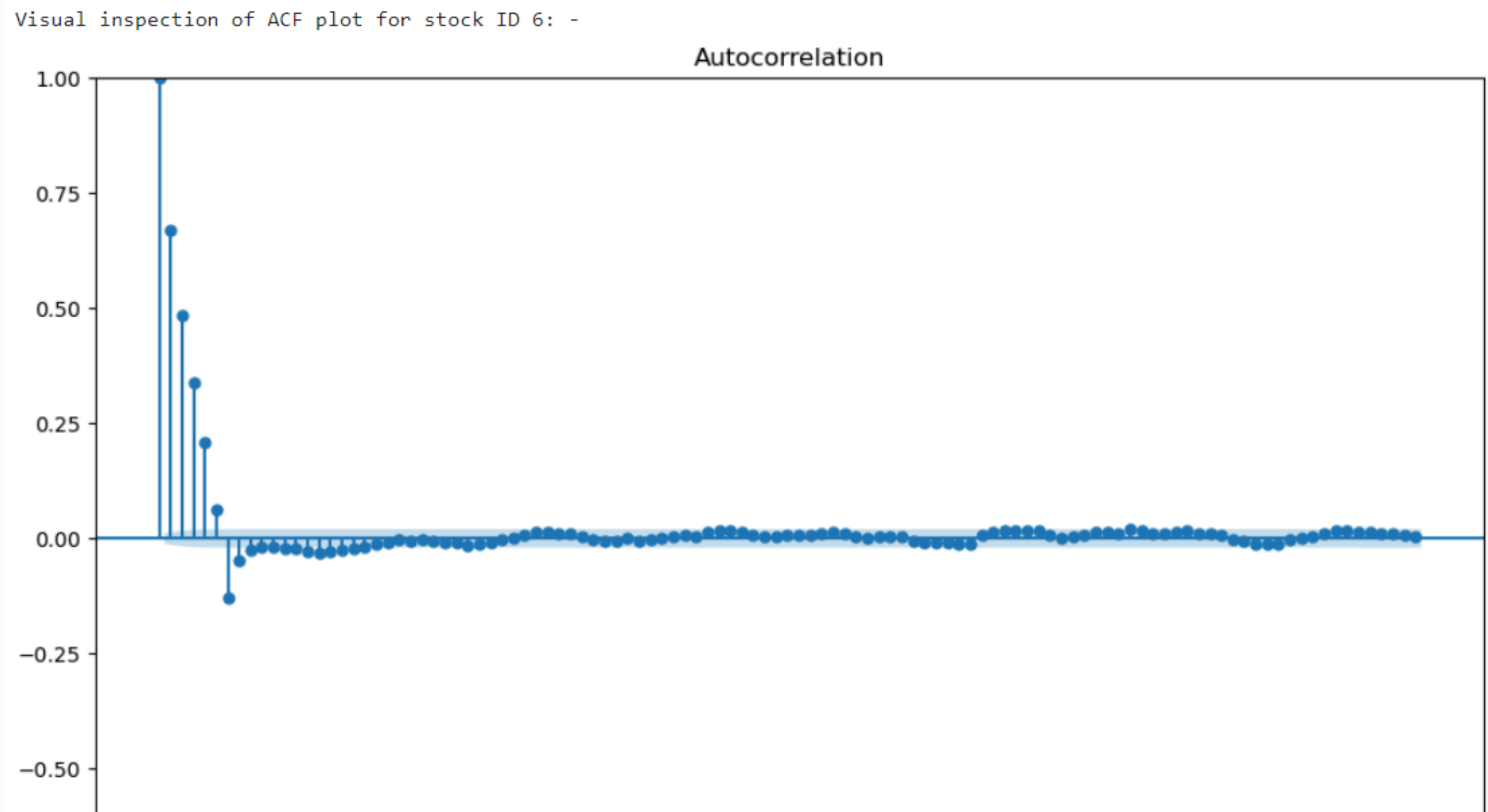
- **Method 1:** - Set  $N = 10$  i.e., choose previous 10 lagged targets.

# METHODS OF SEQUENCE CONSTRUCTION

- **Method 1:** - Set  $N = 10$  i.e., choose previous 10 lagged targets.
- **Method 2:** - Construct sequence based on lagged targets with  $ACF > 0.2$

# METHODS OF SEQUENCE CONSTRUCTION

## METHOD 2 (ACF > 0.2)



# METHODS OF SEQUENCE CONSTRUCTION

## METHOD 2 (ACF > 0.2)...CONTINUED

Common set of lag values with ACF > 0.2 = Lags 1, 2, 3, and 4



# STOCK ID REPRESENTATION

- **Approach 1:** -One-hot vector of dimension 200.

# STOCK ID REPRESENTATION

- **Approach 1:** -One-hot vector of dimension 200.
- **Approach 2:** -8-bit vector of dimension 8.

# STOCK ID REPRESENTATION

- **Approach 1:** -One-hot vector of dimension 200.
- **Approach 2:** -8-bit vector of dimension 8.
- **Approach 3:** - Exclude stock ID from data.

# VARIATIONS IN LSTM MODEL

- **Method of sequence construction**

# VARIATIONS IN LSTM MODEL

- **Method of sequence construction**
- **Choice of Stock ID representation**

# VARIATIONS IN LSTM MODEL

- **Method of sequence construction**
- **Choice of Stock ID representation**
- **Number of hidden units [4, 8, 16, 32]**

# VARIATIONS IN LSTM MODEL

- **Method of sequence construction**
- **Choice of Stock ID representation**
- **Number of hidden units [4, 8, 16, 32]**
- **Number of layers [1]**

# TESTING ON VALIDATION SET

	Stock ID representation							
Model Number	One-hot representation of dimension 200	Binary representation of dimension 8	No representation	Hidden size	Input sequence	Input size	Number of Layers	Number of parameters
1	X	—	—	16	10	201	1	14033
2	X	—	—	16	Number of lagged targets with ACF > 0.2	201	1	14033
3	X	—	—	32	Number of lagged targets with ACF > 0.2	201	1	30113
4	—	X	—	4	Number of lagged targets with ACF > 0.2	9	1	245
5	—	X	—	8	Number of lagged targets with ACF > 0.2	9	1	617
6	—	X	—	32	Number of lagged targets with ACF > 0.2	9	1	5537
7	—	—	X	8	10	1	1	361
8	—	—	X	16	10	1	1	1233
9	—	—	X	32	10	1	1	4513
10	—	—	X	16	Number of lagged targets with ACF > 0.2	1	1	1233
11	—	—	X	32	Number of lagged targets with ACF > 0.2	1	1	4513

**Best performing model** = Model 4

**Best MAD** = 4.444



# FACTORS INFLUENCING MODEL PERFORMANCE

- **Stock ID representation**

# FACTORS INFLUENCING MODEL PERFORMANCE

- **Stock ID representation**
- **Non-linear relationship**

# FACTORS INFLUENCING MODEL PERFORMANCE

- **Stock ID representation**
- **Non-linear relationship**
- **Time-varying relationship**

# FACTORS INFLUENCING MODEL PERFORMANCE

- **Stock ID representation**
- **Non-linear relationship**
- **Time-varying relationship**
- **Small number of features**



## CONTACT ME

**E-mail** divyamans@gmail.com

**Linked-In** <https://www.linkedin.com/in/divyaman-rawat/>

**Project Github: -** <https://github.com/DivNewBeg/Optiver-trading-at-close>

THANK YOU!