GRU Action Forecast

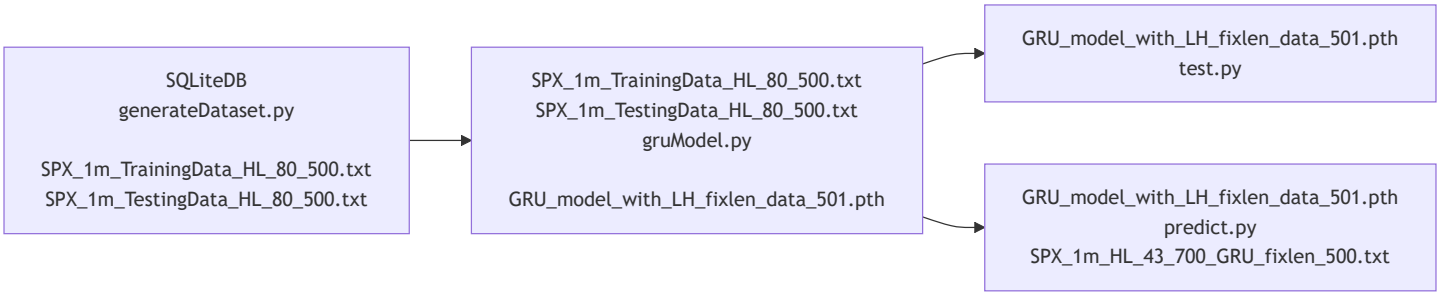


Table of Contents

Todo

1. ~~change trainning data format~~
2. ~~all global variables should read from a configuration file~~
3. ~~optimize Debug~~
4. ~~optimize logging~~
5. ~~clean code make all definitions at begining~~
6. ~~separate plot function from data process code~~
7. ~~🛠️🎯 use class~~

8. ~~send Test output to a file for future reference~~
9. train and test data should be the same other than start/end date
10. read output prediction data, find out accuracy
11. get rid of zigzagplus1.py
12. read any line of dataset, plot it on screen
13. write unit test for all functions and classes

Generate Dataset

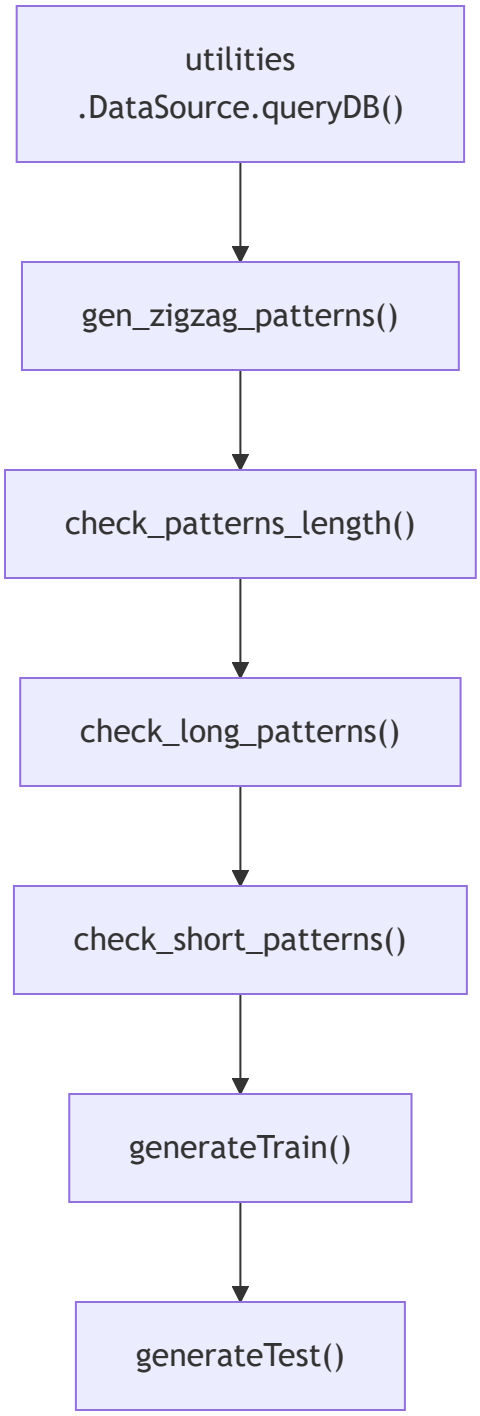
- Define Logger class for whole project
- Define global variables in cofig.ini
- load global variables from cofig.ini
- Generate dataset Source Code

```
def gen_zigzag_patterns(query_start, query_end):
  ... ...
  return ohlc_df, patterns_df
```

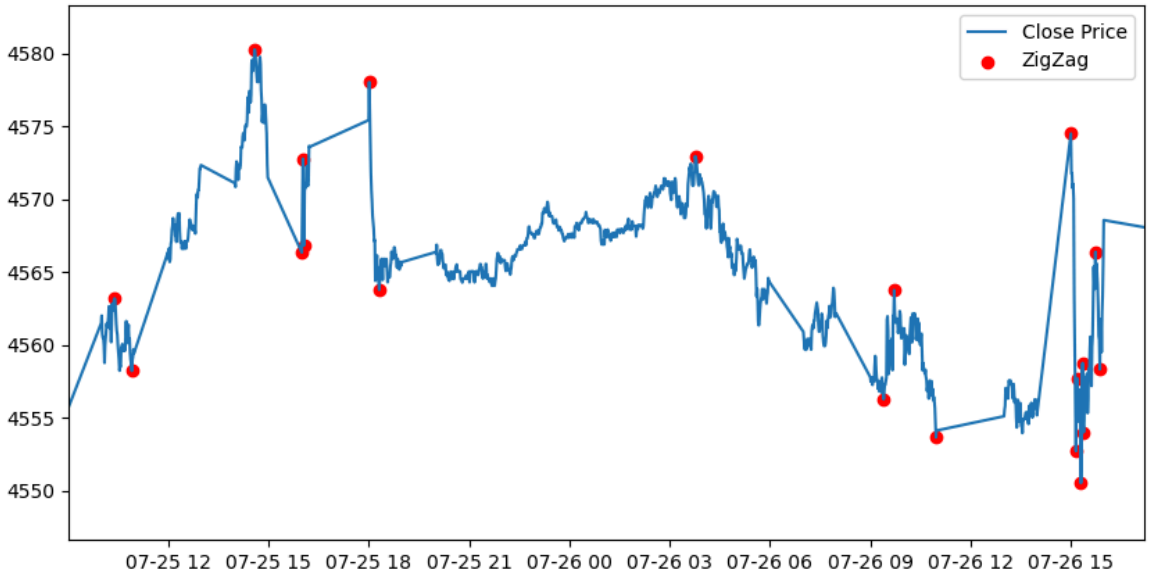
```
ohlc_df
                        Open      High       Low     Close  Volume
Datetime
2023-01-02 18:00:00  3872.998  3877.176  3863.860  3865.983       0
2023-01-02 18:01:00  3866.128  3867.372  3865.378  3865.980       0
2023-01-02 18:02:00  3865.878  3866.878  3865.360  3865.881       0
2023-01-02 18:03:00  3865.742  3865.742  3862.860  3863.613       0
2023-01-02 18:04:00  3863.363  3863.363  3860.742  3860.878       0
...                       ...       ...       ...       ...     ...
2023-12-29 16:08:00  4768.076  4768.076  4767.564  4767.817       0
2023-12-29 16:09:00  4767.564  4768.079  4767.564  4768.079       0
2023-12-29 16:10:00  4767.826  4768.127  4767.817  4767.817       0
2023-12-29 16:11:00  4767.690  4767.820  4767.687  4767.817       0
2023-12-29 16:12:00  4767.567  4767.877  4767.070  4767.070       0

[291380 rows x 5 columns]
```

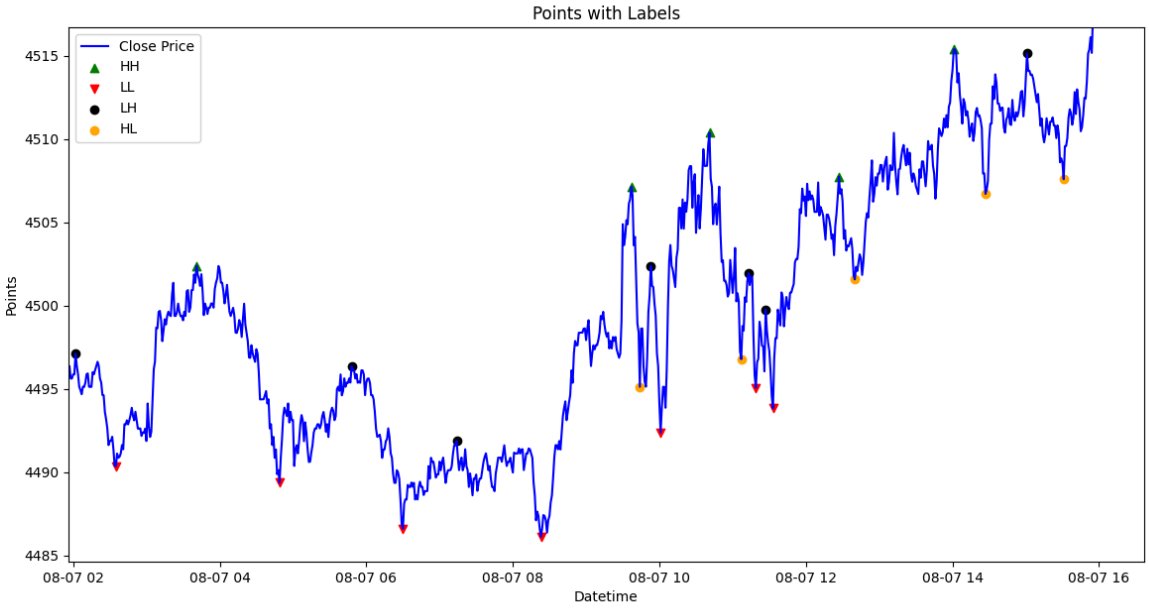
patterns_df

Datetime	Price	Label
2023-01-02 18:07:00	3867.372	HH
2023-01-02 20:27:00	3820.613	LL
2023-01-03 00:45:00	3850.378	LH
2023-01-03 01:06:00	3846.426	HL
2023-01-03 01:31:00	3851.110	HH
...	...	...
2023-12-29 15:20:00	4770.626	HL
2023-12-29 15:28:00	4776.177	HH
2023-12-29 15:54:00	4765.623	LL
2023-12-29 15:59:00	4772.629	LH
2023-12-29 16:12:00	4767.070	HL



- generate plots





🖍️ ☝️Need explaination of above image, ❓How to generate buy/sell points based on the image above ☝️ ❓ Better to have plot to support.

Input

SQLite database file: [data/stock_bigdata_2019-2023.db]

Output files

1. traning dataset

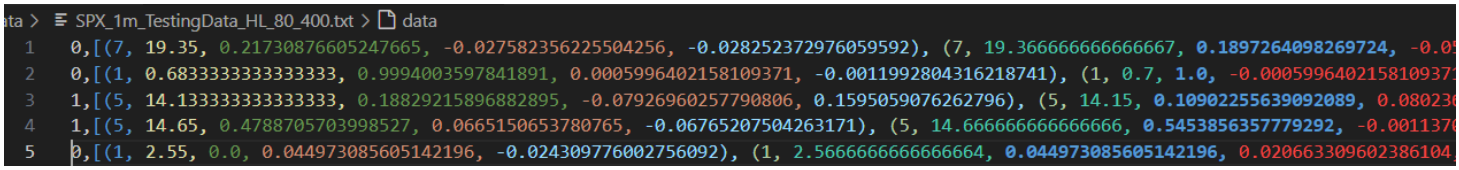
2. testing dataset

- 5 column data group

1. day of weeek
2. time of day
3. close price
4. velocity
5. accelerat

- first column
  1=long
  -1=short

total 60 points end by long/short point for each row which will be total of 5X60=300 numbers



Create GRU Model

- Generate GRU Action Forecast model

Input

- Trainning Dataset
- Testing Dataset

Output

- /GRU_model_with_LH_fixlen_data_501.pth

Test the model

- Test model get R-Square and MSE

Input

- /GRU_model_with_LH_fixlen_data_501.pth

Output

```
2024-09-24 11:13:37,788 - gru - INFO - 1. Load testing data from data/SPX_1m_TestingData_HL_80_5
2024-09-24 11:13:39,398 - gru - INFO - Data shape: (1684, 80, 5)
2024-09-24 11:13:39,398 - gru - INFO - Targets shape: (1684, 1)
2024-09-24 11:13:39,398 - gru - INFO - 2. Define dataset and dataloader
2024-09-24 11:13:39,399 - gru - INFO - 3. Instantiate the model, define the loss function and th
2024-09-24 11:13:39,399 - gru - INFO - Number of layers: 5
2024-09-24 11:13:39,400 - gru - INFO - 4. Load trained model from models/GRU_model_with_LH_fixle
2024-09-24 11:13:39,405 - gru - INFO - 5. Start testing loop
2024-09-24 11:13:41,524 - gru - INFO - Test Loss (MSE): 0.00309951
2024-09-24 11:13:41,526 - gru - INFO - Mean Absolute Error (MAE): 0.01776317
2024-09-24 11:13:41,526 - gru - INFO - R-squared (R2): 0.99687991
2024-09-24 11:13:41,629 - gru - INFO - Saved categorized signals to file : data/SPX_1m_HL_43_700
2024-09-24 11:13:41,629 - gru - INFO - Execution time of test(): 2.2243 seconds
2024-09-24 11:13:41,629 - gru - INFO - ================================= Done
```

- data/SPX_1m_HL_80_500_GRU_fixlen_500.txt

Predict using the model

- predict from testing data by using previous generated model that saved in a file

input

- the model file name is defined in config.ini
- the test data file name is defined in config.ini

output

- the predict result file name is defined in config.ini

```
Target[1.] : Output[0.9852] -> Signal[1.0]
Target[1.] : Output[0.9828] -> Signal[1.0]
Target[1.] : Output[0.9788] -> Signal[1.0]
Target[1.] : Output[0.9798] -> Signal[1.0]
Target[1.] : Output[0.9942] -> Signal[1.0]
Target[1.] : Output[0.9789] -> Signal[1.0]
Target[1.] : Output[0.9650] -> Signal[1.0]
Target[1.] : Output[0.9837] -> Signal[1.0]
... ...


2024-09-24 10:31:19,875 - gru - INFO - 1. Load testing data from data/SPX_1m_TestingData_HL_80_5
2024-09-24 10:31:21,394 - gru - INFO - Data shape: (1684, 80, 5)
2024-09-24 10:31:21,394 - gru - INFO - Targets shape: (1684, 1)
2024-09-24 10:31:21,394 - gru - INFO - 2. Define dataset and dataloader
2024-09-24 10:31:21,394 - gru - INFO - 3. Instantiate the model, define the loss function and th
2024-09-24 10:31:21,394 - gru - INFO - Number of layers: 5
2024-09-24 10:31:21,394 - gru - INFO - 4. Load trained model from models/GRU_model_with_LH_fixle
2024-09-24 10:31:21,394 - gru - INFO - 5. Start testing loop
2024-09-24 10:31:21,394 - gru - INFO - Randomly selected 10 rows and their corresponding outputs
2024-09-24 10:31:21,418 - gru - INFO - Test Output:  1.0135 => Categorized Output:  1.0,
2024-09-24 10:31:21,421 - gru - INFO - Test Output: -1.0031 => Categorized Output: -1.0,
2024-09-24 10:31:21,435 - gru - INFO - Test Output: -1.0092 => Categorized Output: -1.0,
2024-09-24 10:31:21,449 - gru - INFO - Test Output:  1.0013 => Categorized Output:  1.0,
2024-09-24 10:31:21,466 - gru - INFO - Test Output: -0.9915 => Categorized Output: -1.0,
2024-09-24 10:31:21,477 - gru - INFO - Test Output: -1.0087 => Categorized Output: -1.0,
2024-09-24 10:31:21,483 - gru - INFO - Test Output: -1.0060 => Categorized Output: -1.0,
2024-09-24 10:31:21,499 - gru - INFO - Test Output: -0.9803 => Categorized Output: -1.0,
2024-09-24 10:31:21,501 - gru - INFO - Test Output: -1.0313 => Categorized Output: -1.0,
2024-09-24 10:31:21,516 - gru - INFO - Test Output:  1.0100 => Categorized Output:  1.0,
2024-09-24 10:31:21,534 - gru - INFO - Test Output: -0.9957 => Categorized Output: -1.0,
2024-09-24 10:31:21,538 - gru - INFO - Test Output:  0.9820 => Categorized Output:  1.0,
2024-09-24 10:31:21,551 - gru - INFO - Test Output: -1.0023 => Categorized Output: -1.0,
2024-09-24 10:31:21,566 - gru - INFO - Test Output:  0.9771 => Categorized Output:  1.0,
2024-09-24 10:31:21,583 - gru - INFO - Test Output:  1.0199 => Categorized Output:  1.0,
2024-09-24 10:31:21,583 - gru - INFO - Test Output: -1.0413 => Categorized Output: -1.0,
2024-09-24 10:31:21,603 - gru - INFO - Test Output:  0.9827 => Categorized Output:  1.0,
2024-09-24 10:31:21,617 - gru - INFO - Test Output:  0.9888 => Categorized Output:  1.0,
2024-09-24 10:31:21,632 - gru - INFO - Test Output:  1.0297 => Categorized Output:  1.0,
2024-09-24 10:31:21,637 - gru - INFO - Test Output: -1.0142 => Categorized Output: -1.0,
2024-09-24 10:31:21,637 - gru - INFO - ================================= Done
```