Stock prediction using transformer

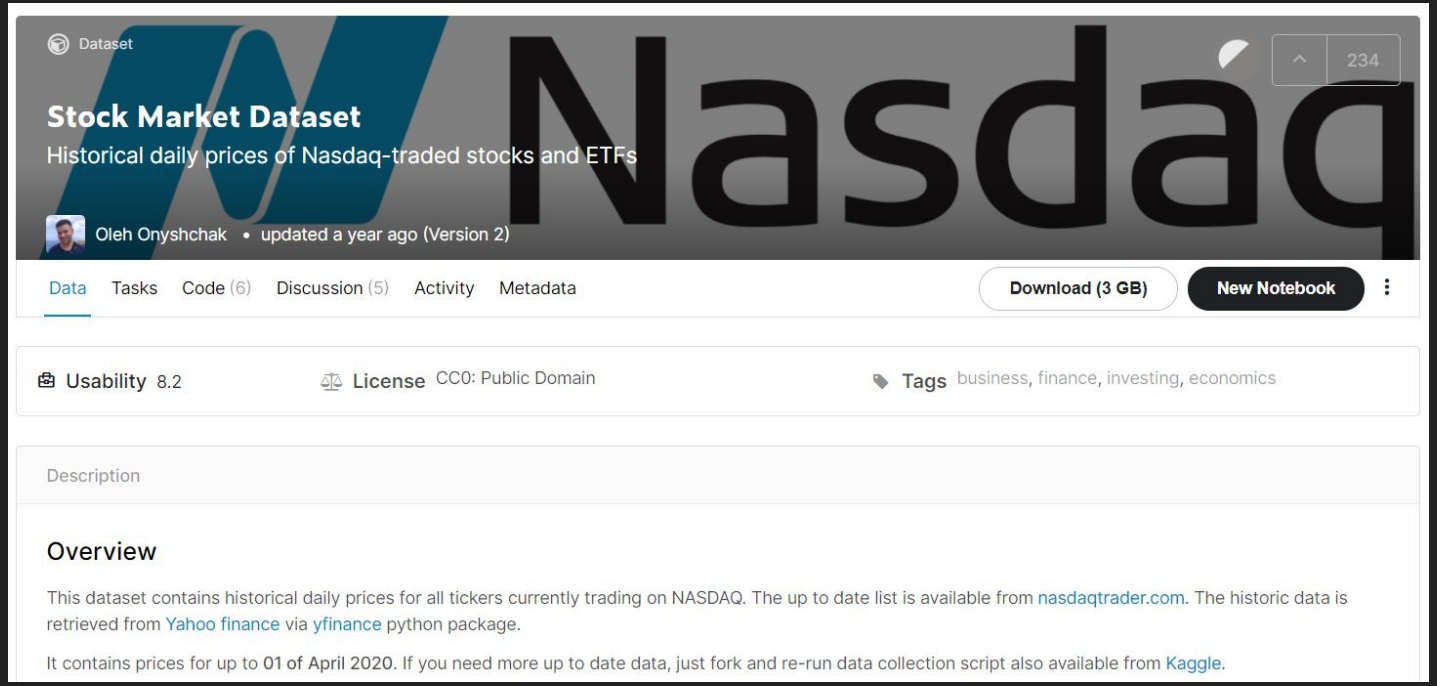
group 32
309552007 袁鈺勛
0716039 謝仁翔

Introduction

- We want to predict companies' closing returns given stock features.
- There are 8 features used for the prediction:
  - Open
  - High
  - Low
  - Closing
  - Volume
  - Periodic time
  - Non-periodic time
  - Symbol

Dataset

The dataset is from Kaggle.



Features

- Open, high, low, closing, and volume
  - These 5 features are converted to percentage changes.
- Periodic time and non-periodic time
  - Periodic time is the sine function of the mean value of the 5 features above.
  - Non-periodic time is the linear function of the mean value of the 5 features above.
- Symbol
  - The sum of the embeddings of each character in the symbol.

Model Structure

- We use a multihead transformer encoder and three linear layers to predict the percentage change during training process.
- The predicted percentage changes will be used to produce the closing returns in inference.
- Input of the model
  - (batch size, sequence length, 5)
- Input of the transformer encoder
  - (batch size, sequence legnth, 8)
- Output
  - Percentage changes

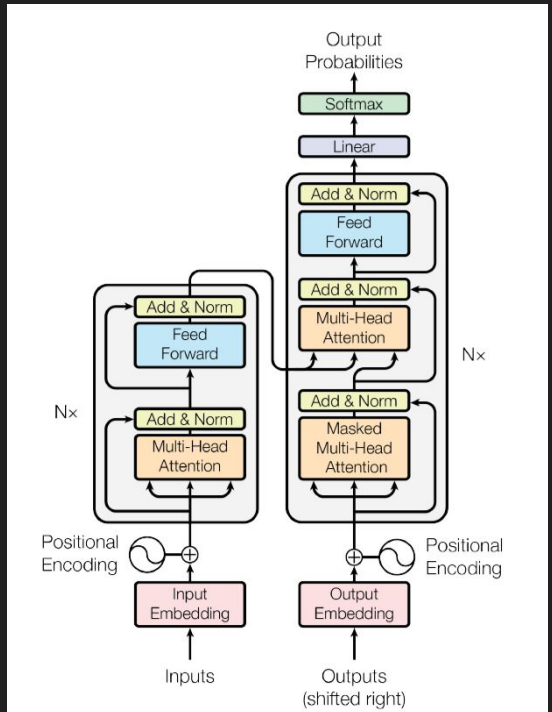


Figure 1: The Transformer - model architecture.

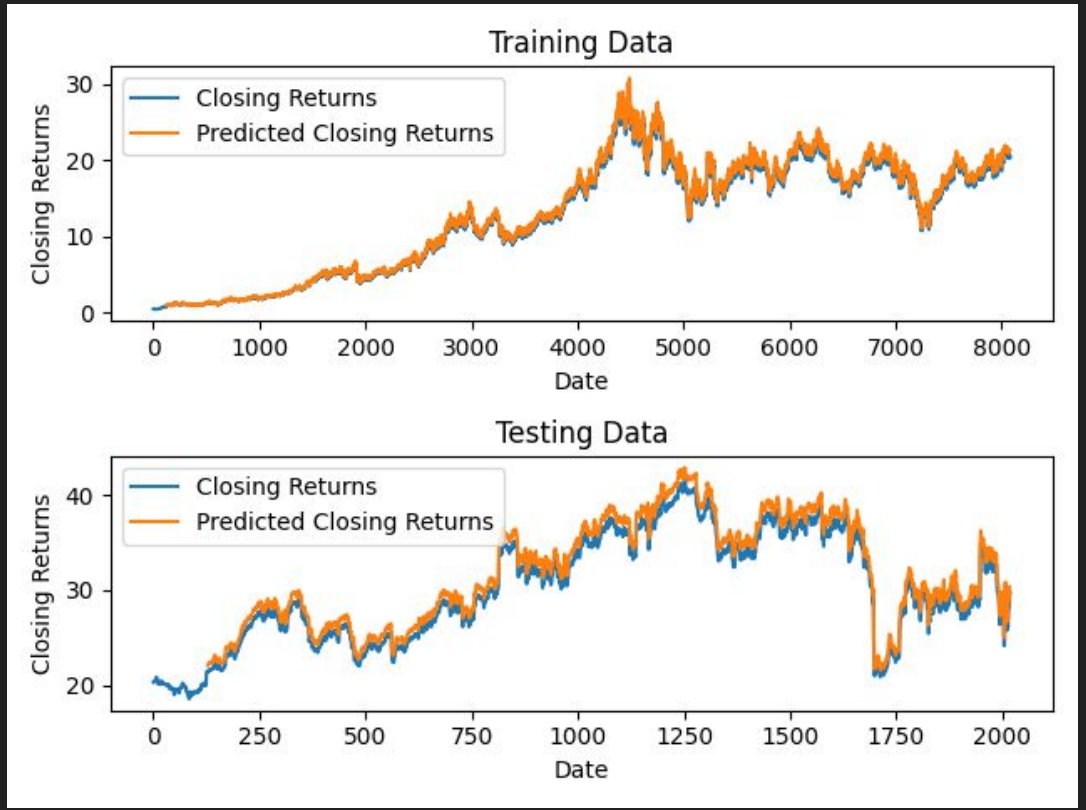
Loss Function

- We use L2 loss as loss function.
- $E_{x,y\sim pdata}[\ (y - y'(x))^2\ ]$

Hyperparameters

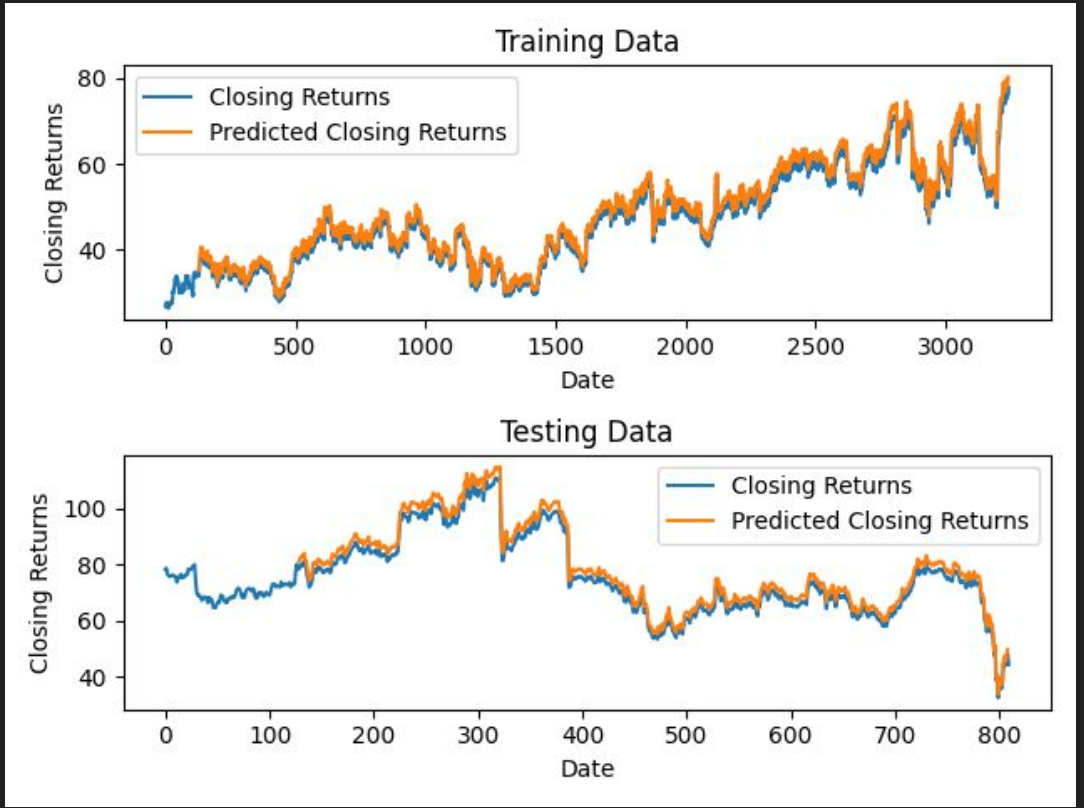
- Batch size: 256
- Learning rate: 0.001
- Sequence length : 128
- Number of encoders: 6
- Attention dimension: 128
- Hidden size: 512
- Dropout: 0.3
- Number of heads: 8

ConAgra Brands, Inc. (CAG)



Drive Shack Inc. (DS)

Magellan Health Inc. (MGLN)



Conclusion

- Predicted closing return is not accurate enough.
- Future works:
  - Use the date as time feature directly
  - Longer sequence length and deeper model
  - Use another way to embed the symbol

Q&A