Financial Reinforcement Learning using Transformers

Andy Zhu, Wenjie Chen, Yun Zhe Chen, David Chong, Hongwei Li, Derrick Lin, Rensselaer Polytechnic Institute, Department of Computer Science
Guided by Prof. Yanglet Xiao-Yang Liu
Rensselaer Center for Open Source

INTRODUCTION

- Focus on transformer models for sequence modeling tasks.
- Aim to gain insights into the use of transformers for financial time series.
- Analyzing the training performance of transformers for stock prediction.
- Utilizing Yahoo Finance data for predictions.

1	date	DJI	GSPC	IXIC
2	2001-01-02	10646.150390625	1283.27001953125	2291.860107421875
3	2001-01-03	10945.75	1347.56005859375	2616.68994140625
4	2001-01-04	10912.41015625	1333.3399658203125	2566.830078125
5	2001-01-05	10662.009765625	1298.3499755859375	2407.64990234375
6	2001-01-08	10621.349609375	1295.8599853515625	2395.919921875

- Investigating the impact of various loss functions on model optimization.
- Loss functions in focus: Mean Squared Error (MSE), Mean Absolute Error (MAE), Cross-Entropy Loss.

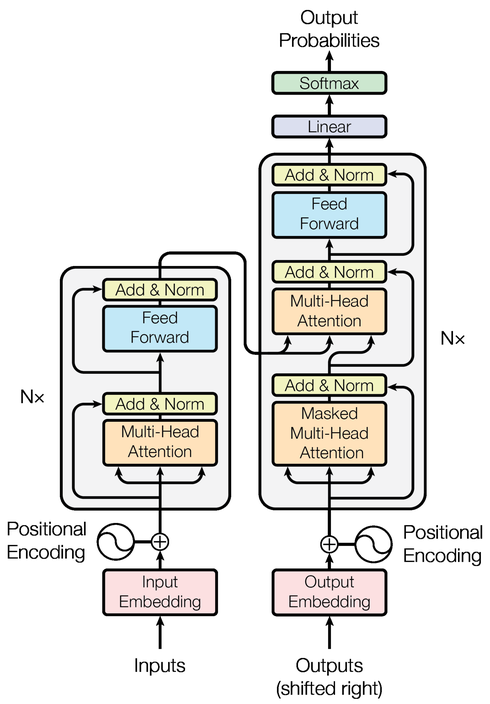
OBJECTIVE

The objective of this project is to conduct an empirical study into the training performance of transformer models in the context of different machine learning loss functions starting with a stock prediction model from Yahoo Finance data:

- Assess the effectiveness of MSE and MAE in transformer models for time series prediction.
- Evaluate the impact of Cross-Entropy Loss on transformer model performance.
- Compare performance metrics between Long Short-Term Memory (LSTM) models and transformers.
- Establish a robust baseline model for future financial applications.

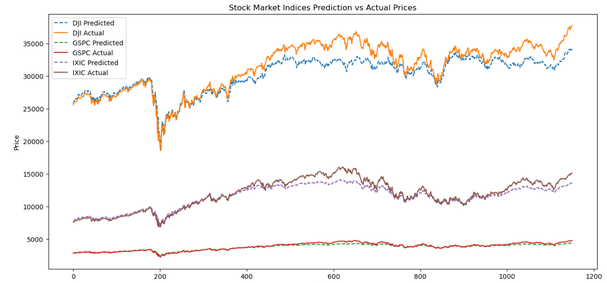
METHODOLOGY

1. Data Collection and Preparation
   - Extract and preprocess financial time series data from Yahoo Finance.
   - Construct a suitable time series dataset for model training and evaluation.
2. Implement baseline LSTM and transformer models using PyTorch.
   - Employ MSE and MAE as loss functions for initial training phases.
   - Test the applicability of Cross-Entropy Loss in prediction accuracy.
3. Performance Evaluation
   - Measure and analyze model performance using standard metrics
   - Compare results across different models and loss functions.

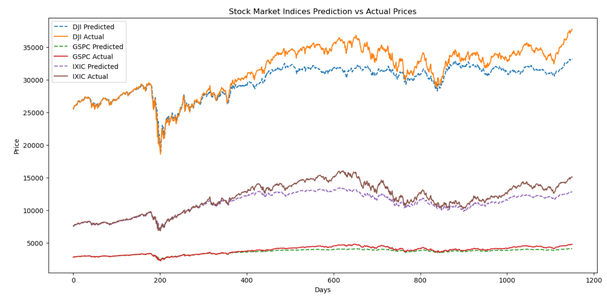


RESULT

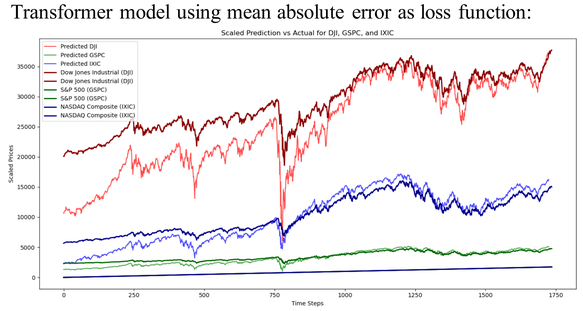
Transformer model using mean square error as loss function:



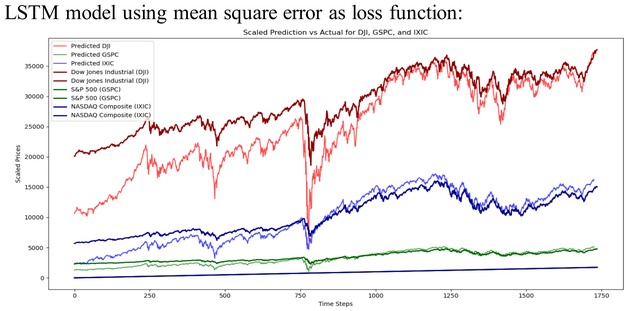
Transformer model using mean absolute error as loss function:



Transformer model using mean absolute error as loss function:



LSTM model using mean square error as loss function:



ANALYSIS / SUMMARY

- MSE and MAE as loss function for transformer model to predict stock indices have similar efficacy
- MAE improves the total loss by 3% compared to using MSE.
- Cross entropy is not suitable for stock price prediction because it is designed for classification tasks involving probabilities, not for regression tasks involving continuous numerical outputs like prices.