# Using AI to Develop Stock Trading Strategy

Members: Ryan Kuo, Lee Po-Ting, Cheng-Jih Chang, Johnson Chang

Mentor: Jerry Liao

## Introduction

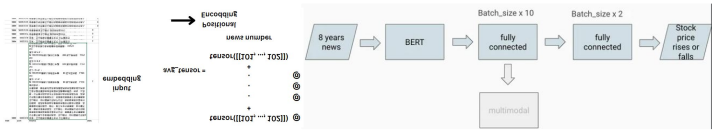
- Traditional data analysis methods face significant challenges. Effectively utilizing multiple data sources to improve stock price prediction accuracy has become an important research topic.
- Analyzing a single data source often fails to comprehensively reflect market dynamics. Existing prediction models typically rely on single-modal data, resulting in insufficient prediction accuracy. Hence the initiative for combining multimodal data to improve accuracy
- we aim to to explore multimodal training methods, combining text processing and graphical analysis, and using Graph Neural Networks (GNN) to transform features, thereby enhancing portfolio profit prediction performance.

## Dataset

- Stock price data and macroeconomic data are fetched from yahoo finance api and FRED api, and the images are generated by those data.
- News text data are crawled from popular news websites in Taiwan.
- Stocks relationships in GNN are found on stock market analysis software, XQ.

## Methodology and Results

### Text Processing




### Graphical Processing

- Used stock prices to perform 60-day normalization and generate grayscale stock candlestick charts.
- Trained a CNN with a filter size of 5x3 to capture more price trend ranges. The output consists of two neurons for softmax, trained using a loss function, and another output with 10 neurons to extract vectors for GNN training and testing.
- Also Experimented with ResNet, VGGNet, and U-net.

### Macroeconomic indicators

- Created a macroeconomic indicators vector to provide the model with a broader and more diverse set of data.
- Performed 60-day normalization on this data to ensure consistency and minimize the impact of different time points.

### GNN Training

- Create Total Network, Homogeneous, and Heterogeneous types to illustrate relationship among all companies for GAT training.
- Build final model, test, and squeeze vectors into 1 for each company in each trading day
- Sort companies by their vectors

## Discussion and Conclusion

- Testing accuracy: 49.5% ~ 52%
- Backtesting: negative return rate
- 4 reasons for failure:
  1. Too Many Dimensions, but too Little Data
  2. Not Total Market Analysis for Factor Investing
  3. Building Model by Closing Price
  4. Only Consider Return Rate, not Stock Price

## Future Work

- Conduct Total Market Analysis
- Collect More Data
  1. shorten time intervals
  2. enlarge sources of market information
- Invest in Hardware of Higher Level

## References

- https://economics.yale.edu/research/re-imagining-price-trends

stanCode
標準程式教育機構