Using AI for Stock Market Pattern Detection and Trading

Group Member Information

Student Name: Yuen Chun Ho

Student ID: 24113812G

Contribution: 100% (Individual Project)

Abstract

This report explores the application of artificial intelligence techniques for automated stock market pattern detection and trading. By combining object detection models (YOLOv12) with classification networks, we create a system capable of identifying profitable trading opportunities based on the "Cup and Handle" chart pattern. Our approach uses a two-stage pipeline: first detecting pattern formations in candlestick charts, then predicting profitability using two different decision model architectures. We compare decision models using full candlestick data versus technical indicators as inputs. Testing on S&P 500 and Tesla stock data demonstrates success rates of approximately 76% using full-sequence models and 68-71% using technical indicator-based models, suggesting that AI can effectively identify non-random price movements in financial markets. This approach transforms traditional technical analysis into a data-driven, objective methodology that could potentially create positive-expectation trading scenarios.

1. Introduction

1.1 Motivation

Stock market trading is often compared to a coin flip, where short-term price movements appear random with a 50-50 probability of success. However, if we could identify scenarios where the probability shifts in our favor (e.g., from 50% to 70%), we would create a positive expectation game worth playing. Technical analysis suggests such non-random moments exist, but traditional approaches rely heavily on subjective interpretation and experience.

This project aims to use artificial intelligence to:

1. Objectively identify non-random time points in stock data

2. Predict the probability of profitable trades at these points
3. Create a system with positive expectation value

1.2 Approach Overview

Our methodology employs a two-stage AI pipeline:

1. **Pattern Detection Stage**: We use YOLOv12 (You Only Look Once), a state-of-the-art object detection model, to identify specific chart patterns in stock data.
2. **Trading Decision Stage**: Once patterns are detected, we employ two different types of classification approaches:
   - **Full Sequence Models**: Using all 400 candlesticks' OHLCV data to make binary trading decisions
   - **Technical Indicator Models**: Using 10 technical indicators extracted from the last candlestick to make multi-label trading decisions

By converting visual chart patterns into quantifiable trading opportunities, we create a system that potentially outperforms random chance.

2. Background and Related Work

2.1 Object Detection in Financial Analysis

Recent advancements in computer vision, particularly object detection, have opened new possibilities for automating technical analysis. The evolution of YOLO algorithms has progressively improved speed and accuracy, with YOLOv12 introducing Area Attention (A2) modules and Residual Efficient Layer Aggregation Networks (R-ELAN) that enhance performance on complex patterns.

2.2 Stock Market Technical Analysis

Technical analysis uses historical price data to predict future price movements. The candlestick chart, developed in 18th century Japan, remains one of the most popular visualization methods. Each candlestick represents four critical price points:

- Open Price: The starting price at the beginning of the period
- Close Price: The final price at the end of the period
- High Price: The highest price reached during the period
- Low Price: The lowest price reached during the period

2.3 Dow Theory and Chart Patterns

The Dow Theory, one of the foundations of technical analysis, suggests that stock markets alternate between primary trends and secondary corrections, with random fluctuations occurring most of the time. However, it also proposes that certain periods demonstrate non-random behavior.

One such non-random pattern is the "Cup and Handle" formation, which resembles a teacup on a price chart followed by a small downward drift. Research has shown that this pattern can indicate potential upward price movements with above-average reliability.

2.4 Machine Learning in Stock Trading

Previous research has primarily focused on price prediction rather than pattern detection. Our approach differs by treating chart pattern identification as an object detection problem, followed by a classification task to determine trade profitability.

3. Methodology

3.1 Data Collection and Preprocessing

We utilized S&P 500 data from January 2000 to February 2025, comprising approximately 358,750 15-minute OHLC (Open, High, Low, Close) datapoints. The dataset was split as follows:

- 80% for training (further split into 80% for training and 20% for validation)
- 20% for testing

For additional validation, we also used Tesla stock data (5-minute intervals) from September 2021 to June 2024.

3.2 Training Process Overview

The training process consisted of several interconnected steps:

1. Preliminary manual annotation of Cup and Handle patterns
2. Initial training of the YOLOv12 model
3. Model-assisted additional data annotation
4. Iterative model refinement
5. Building classifier datasets using detected patterns
6. Training and evaluating classification models (both full-sequence and technical indicator based)

3.3 Object Detection with YOLOv12

The object detection component was designed to identify Cup and Handle patterns in candlestick charts. We implemented a sliding window approach where each window contained 400 candlesticks.

3.3.1 Annotation and Model Training

The training process followed a bootstrap approach:

1. Manual annotation of 76 images (66 for training, 10 for validation)
2. Initial YOLOv12 model training
3. Using the trained model to detect patterns in new data
4. Manual verification of new detections
5. Expanding the training dataset with verified detections
6. Retraining the model with the expanded dataset

This process was repeated several times, eventually building a dataset of 345 training and 55 validation samples.

3.4 Decision Model Development

We developed two different approaches for the decision models:

3.4.1 Full Sequence Models (400 Candlesticks)

These models take the full sequence of 400 candlesticks' OHLCV data as input and produce a binary classification outcome (profitable or not).

Model Architectures:

1. **1D CNN-LSTM Model**: Combined convolutional layers to extract features with LSTM to capture temporal dependencies
2. **Transformer Model**: Utilized self-attention mechanisms to understand temporal relationships in the price sequence

3.4.2 Technical Indicator Models (10 Indicators)

Instead of using the full OHLCV sequence, these models use only 10 technical indicators calculated from the last candlestick.

Feature Engineering:
We engineered the following technical indicators:

- RSI (Relative Strength Index)
- MACD (Moving Average Convergence Divergence)
- MACD Signal Line
- MACD Histogram
- Stochastic Oscillator K%
- Stochastic Oscillator D%
- Money Flow Index (MFI)
- Bollinger Band Width
- Average True Range (ATR)
- On-Balance Volume (OBV)

These features were log-transformed and standardized before training.

Model Architectures:

1. **1D CNN Model**: For processing the technical indicators
2. **Transformer Model**: Alternative approach using self-attention mechanisms

3.4.3 Multi-label Classification

For the technical indicator models, we implemented a multi-position classification approach with 5 different price targets, creating a more nuanced trading strategy beyond simple win/loss outcomes.

3.5 Trading Strategy Implementation

For each detected pattern, we established:

- Entry price: Based on the pattern's detection point
- Stop loss: Based on the pattern's lowest price point
- Take profit: Calculated based on risk-to-reward ratio

The trading rule was simple:

- If price hits stop loss first: Trade is a loss (0)
- If price hits take profit first: Trade is a win (1)

4. Experimental Results

4.1 YOLO Detection Performance

The final YOLOv12 model successfully identified Cup and Handle patterns with high precision. Visual evaluation showed good alignment between the detected patterns and their actual appearance in the charts.

4.2 Full Sequence Model Performance (400 Candlesticks)

4.2.1 S&P 500 Results

Training Set (285 samples):

- Both models: 80.70% success rate

Validation Set (72 samples):

- Both models: 76.39% success rate

Test Set (47 samples):

- Both models: 76.60% success rate

4.2.2 Tesla Stock Results (29 samples)

- Both models: 62.07% success rate

4.3 Technical Indicator Model Performance (10 Indicators)

Validation Results:

- CNN model: 68.06% accuracy
- Transformer model: 71.11% accuracy

Multi-label Classification Results:
The label distribution across the 5 positions showed decreasing success rates for more aggressive price targets:

- Position 1: 79.8% positive cases
- Position 2: 63.6% positive cases
- Position 3: 47.9% positive cases
- Position 4: 31.9% positive cases
- Position 5: 17.6% positive cases

F1 scores for the models:

- Position 1: Both models achieved F1=0.8661
- Position 2: Transformer (F1=0.7692) slightly outperformed CNN (F1=0.7434)
- Position 3: CNN (F1=0.4810) outperformed Transformer (F1=0.2609)
- Positions 4 & 5: Both models struggled (F1=0)

5. Discussion

5.1 Significance of Results

The results demonstrate that AI can successfully identify profitable trading opportunities based on chart patterns with performance significantly above random chance. The consistent 76% success rate on test data for full sequence models suggests the system has captured meaningful patterns rather than simply overfitting to training data.

5.2 Comparison of Model Approaches

- **Full Sequence Models** achieved better overall performance (76.60% on test data) compared to the technical indicator models (68-71% on validation). This suggests that the temporal dynamics captured in the full sequence provide valuable information that may be lost when reducing to just technical indicators.
- **Technical Indicator Models** provide a more nuanced multi-position trading approach but performed less consistently across different positions.

5.3 Limitations

1. **Dataset Size**: With only 345 training samples, the models may not generalize well to all market conditions
2. **Overfitting Concerns**: The high training accuracy (80.70%) compared to test accuracy (76.60%) indicates some overfitting
3. **Limited Pattern Coverage**: Only Cup and Handle patterns were considered
4. **Performance Degradation**: Lower success rates on Tesla data (62.07%) suggest challenges in generalizing across different stocks
5. **Position Imbalance**: For multi-label classification, later positions had very few positive cases, making it difficult to train effective models

6. Conclusion and Future Work

6.1 Conclusion

This project demonstrates the feasibility of using AI for automated chart pattern detection and trading. By combining object detection with classification models, we created a system that can identify profitable trading opportunities with success rates significantly above random chance. The full sequence models achieved approximately 76% success rate on S&P 500 test data, while technical indicator-based models achieved 68-71% accuracy with a more nuanced multi-position approach. This work transforms traditional technical analysis from a subjective art into a more objective, data-driven methodology.

6.2 Future Work

1. **Expanded Pattern Library**: Train the system to recognize additional chart patterns
2. **Larger Dataset**: Collect more samples to improve generalization
3. **Alternative Features**: Explore different technical indicators and feature combinations
4. **Market Regime Awareness**: Incorporate market regime detection to adapt to different market conditions
5. **Hybrid Models**: Combine the strengths of full sequence models with technical indicator approaches
6. **Real-time Deployment**: Develop a production system for real-time trading signals

References

1. Research on Cup and Handle Price Pattern Based on SVM Method. (2024). ResearchGate. https://www.researchgate.net/publication/379345807_Research_on_Cup_and_Handle_Price_Pattern_Based_on_Svm_Method
2. Train YOLOv12 Model. (2024). Roboflow Blog. https://blog.roboflow.com/train-yolov12-model/

`