Deep Learning-Based Technical Analysis System for Stock Market Pattern Recognition

In today's dynamic financial markets, the ability to rapidly identify and act upon trading opportunities is crucial for investment success. While traditional technical analysis relies heavily on human interpretation of chart patterns and indicators, advances in deep learning present an opportunity to automate and potentially improve this process. This project aims to develop a sophisticated stock scanning system that leverages deep learning techniques to identify technical patterns and generate trading signals with greater accuracy than conventional methods.

Data Description and Context

This study will utilize comprehensive stock market data from the S&P 500 companies, spanning from January 2020 to December 2024. The dataset will be sourced through the Yahoo Finance API, providing a robust foundation for analysis with the following characteristics:

Primary Variables:

* Temporal data: Trading dates and timestamps
* Price data: Daily open, high, low, and close prices (OHLC)
* Volume data: Daily trading volume
* Derived technical indicators:
  + Momentum indicators (RSI, MACD, Momentum)
  + Trend indicators (Various moving averages, Bollinger Bands)
  + Volatility indicators (Average True Range, Standard Deviation)

The dataset encompasses approximately 1,250 trading days per stock across 500 companies, resulting in roughly 625,000 daily observations. This substantial dataset provides adequate depth for both training deep learning models and conducting out-of-sample testing.

Data Collection Methodology: The data acquisition process will employ Python's yfinance library for automated daily harvesting of historical data. This approach ensures:

1. Consistent data quality and format
2. Automated handling of stock splits and dividends
3. Reliable access to real-time updates for live testing

Industry Context: The U.S. equity market represents over $7 trillion in market capitalization, with algorithmic trading accounting for approximately 60-70% of daily trading volume. Within this context, sophisticated pattern recognition tools have become increasingly crucial for maintaining competitive advantage in systematic trading operations.

Research Questions and Hypotheses

Primary Research Question: Can deep learning models identify and predict technical trading patterns with higher accuracy than traditional rule-based methods?

Specific Hypotheses:

H1: CNN-Based Pattern Recognition Accuracy

* Null Hypothesis (H0): A convolutional neural network (CNN) cannot identify technical chart patterns with accuracy exceeding 75%
* Alternative Hypothesis (H1): A CNN can identify technical chart patterns with accuracy exceeding 75%

H2: RNN Prediction Performance

* Null Hypothesis (H0): RNN/LSTM models do not provide better short-term (1-3 days) price movement predictions compared to traditional moving average crossover strategies
* Alternative Hypothesis (H2): RNN/LSTM models provide statistically significant improvements in prediction accuracy compared to moving average crossover strategies

H3: Combined Indicator Effectiveness

* Null Hypothesis (H0): The combination of multiple technical indicators processed through deep learning models does not provide better trading signals than individual indicators
* Alternative Hypothesis (H3): The combination of multiple technical indicators processed through deep learning models provides statistically significant improvements in trading signal accuracy

These hypotheses will be tested using a combination of:

1. Accuracy metrics for pattern recognition (precision, recall, F1-score)
2. Prediction performance metrics (RMSE, MAE, directional accuracy)
3. Trading performance metrics (risk-adjusted returns, Sharpe ratio, maximum drawdown)

The project will leverage key concepts and skills acquired throughout the MS in Advanced Data Analytics program, particularly drawing from:

* ADTA 5550 Deep Learning with Big Data
* ADTA 5560 Recurrent Neural Networks for Sequence Data
* ADTA 5240 Harvesting, Storing, and Retrieving Data

Expected outcomes include a fully functional stock scanning system capable of:

1. Real-time pattern recognition using CNN architectures
2. Price movement prediction using RNN/LSTM models
3. Automated trading signal generation based on model outputs
4. Performance evaluation and comparison with traditional methods

This research will contribute to the field by demonstrating the practical application of deep learning techniques to technical analysis, potentially offering improved accuracy and reliability compared to conventional methods. The findings will be valuable for both academic understanding of market pattern recognition and practical implementation in trading systems.