Deep Reinforcement Learning to Design Trading Strategies for Organizations listed in NEPSE

M.Sc. Final Defense

Presented By:
Ravi Prajapati (THA079MSISE011)

Supervised By:
Devendra Kathayat

Department of Electronics and Computer Engineering
Institute of Engineering, Thapathali Campus

August 25, 2024

Presentation Outlines

- Motivation
- Background
- Problem Statement
- Objectives of Project
- Scope of Project
- Originality of Project
- Potential Applications
- Literature Review
- Methodology
- Results
- Discussion and Analysis
- Future Enhancement
- Conclusion
- Project Schedule
- References

Motivation

- Utilizing deep reinforcement learning (DRL), specifically the Proximal Policy Optimization (PPO) algorithm, to enhance trading strategies in Nepal's hydro power sector within the NEPSE.
- The project seeks to develop adaptive strategies that respond dynamically to market changes, potentially leading to improved decision-making and market efficiency.

Background

- The Nepalese Stock Exchange (NEPSE) index serves as a pivotal indicator of the performance of Nepal's financial market.
- The Nepalese market presents unique challenges, including limited liquidity, sparse historical data, and susceptibility to external shocks.
- This project leverages deep reinforced learning to enhance forecasting accuracy of price of Hydropower sector stocks, enabling more informed decision-making and strategic investment planning.

Problem Statement

- Non Existence of techniques and trading strategies in accurately predicting trend of stocks in NEPSE exploring methodologies like PPO algorithm incorporating technical indicators for better price trend prediction.

Objectives of Project

- To develop a Proximal Policy Optimization (PPO) based trading strategy for organizations in the hydro power sector listed in NEPSE.
- To validate the effectiveness of the developed trading strategy through rigorous backtesting and simulation.

Scope of Project

- Models ability to Create a Trading Strategy for Organization in NEPSE to predict potential future price movements and trends and make trades.
- Model performance may be constrained by market dynamics and sector specificity.

Originality of Project

- Proximal Policy Optimization (PPO) based trading strategy specifically tailored for the hydro power sector within the Nepal Stock Exchange (NEPSE).

Potential Applications

- Implement PPO-based trading strategies in algorithmic trading systems.
- Offer tailored investment advice in the hydro power sector based on research findings, enhancing advisory services.
- Incorporate PPO algorithm into trading bots for autonomous decision-making, adaptable to market changes.
- Integrate PPO-based strategies into portfolio optimization platforms for enhanced portfolio performance and risk management.

Literature Review [1]

Research	Design of Stock Trading Agent Using Deep Reinforcement Learning	Deep reinforcement learning approach for trading automation in the stock market	Adaptive Stock Trading Strategies with Deep Reinforcement Learning Methods	Stock price forecast based on combined model of ARI-MA-LS-SVM	Performance of Deep Learning in Prediction of Stock Market Volatality
Author	Janak Kumar Lal	Kabbani, Taylan and Duman, Ekrem	Xing Wu, Haolei Chen, Jianjia Wang, Luigi Troiano, Vincenzo Loia, Hamido Fujita	Chenglin Xiao, Weili Xia, and Jijiao Jiang.	Moon Kyoung-Sook and KIM Hongjoong.
Date	September 2022	September 2022	May 2020	2020	2019
Duration	2012-01-01 to 2022-07-15	2009 to 2021	January 1st, 2008 to December 20th, 2018.	102 trading days of 2015.	7-year period from 2010 to 2016.

Literature Review [2]

Research	Design of Stock Trading Agent Using Deep Reinforcement Learning	Deep reinforcement learning approach for trading automation in the stock market	Adaptive Stock Trading Strategies with Deep Reinforcement Learning Methods	Stock price forecast based on combined model of ARI-MA-LS-SVM	Performance of Deep Learning in Prediction of Stock Market Volatality
Dataset	Data of four commercial Banks (ADBL, CZBIL, LBL and NABIL)	Stocks from Yahoo Finance	UK, US and Chinese Stock Market	selects various indicators of Auto-desk (002227) every day	tested 5 stock market indices,S&P500, NASDAQ, German DAX, Korean KOSPI200 and Mexico IPC
Data Used	HIstorical Data	Historical Data and News	Historical Data and Technical Indicators	historical data, but focusing on technical analysis	historical data and technical analysis

Literature Review [3]

Research	**Design of Stock Trading Agent Using Deep Reinforcemen t Learning**	**Deep reinforcement learning approach for trading automation in the stock market**	**Adaptive Stock Trading Strategies with Deep Reinforcement Learning Methods**	**Stock price forecast based on combined model of ARI-MA-LS-SVM**	**Performance of Deep Learning in Prediction of Stock Market Volatality**
Model Used	Double Deep Q learning agent	Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm	GRU, GDQN (Gated Deep Q-learning trading strategy) and GDPG (Gated Deterministic Policy Gradient trading strategy)	a cumulative auto-regressive moving average is proposed, which combines the least squares support vector machine synthesis model (ARI-MA-LS-SVM)	LSTM
What The Model Predict	Predicts when to buy and sell stocks	Buying and Selling Stocks	Buy and sell Stocks	predict the trend of stock prices	predictions of both market index and volatility

Methodology

- **Dataset**

  Includes data of stock price of 91 hydro power companies listed in Nepse.

  The table below shows the sample of data for Arun Valley Hydropower Development Company Limited (AHPC).

Date	Open	High	Low	LTP	% Change	Quantity	Turnover
5/30/2024	161.3	163	158	158.9	-1.49	172,008.00	27,470,999.70
5/29/2024	165	166	161.1	161.3	-0.8	200,870.00	32,775,237.30
5/27/2024	165.3	167	162.1	162.6	-1.33	191,749.00	31,276,705.00
5/26/2024	166	166	161.1	164.8	0.61	272,417.00	44,369,040.00
5/22/2024	163	166	161.2	163.8	1.74	323,763.00	53,071,154.50

Methodology [2]

- **Dataset Preprocessing**
    - **Converting price data into Technical Indicator Values**

      Mathematical calculations is done to calculate Indicators like SMA, EMA, MACD, Bollinger Band, Stochastic Oscillator, On Balance Volume, RSI and Average Directional Index(ADX).

Methodology [3]

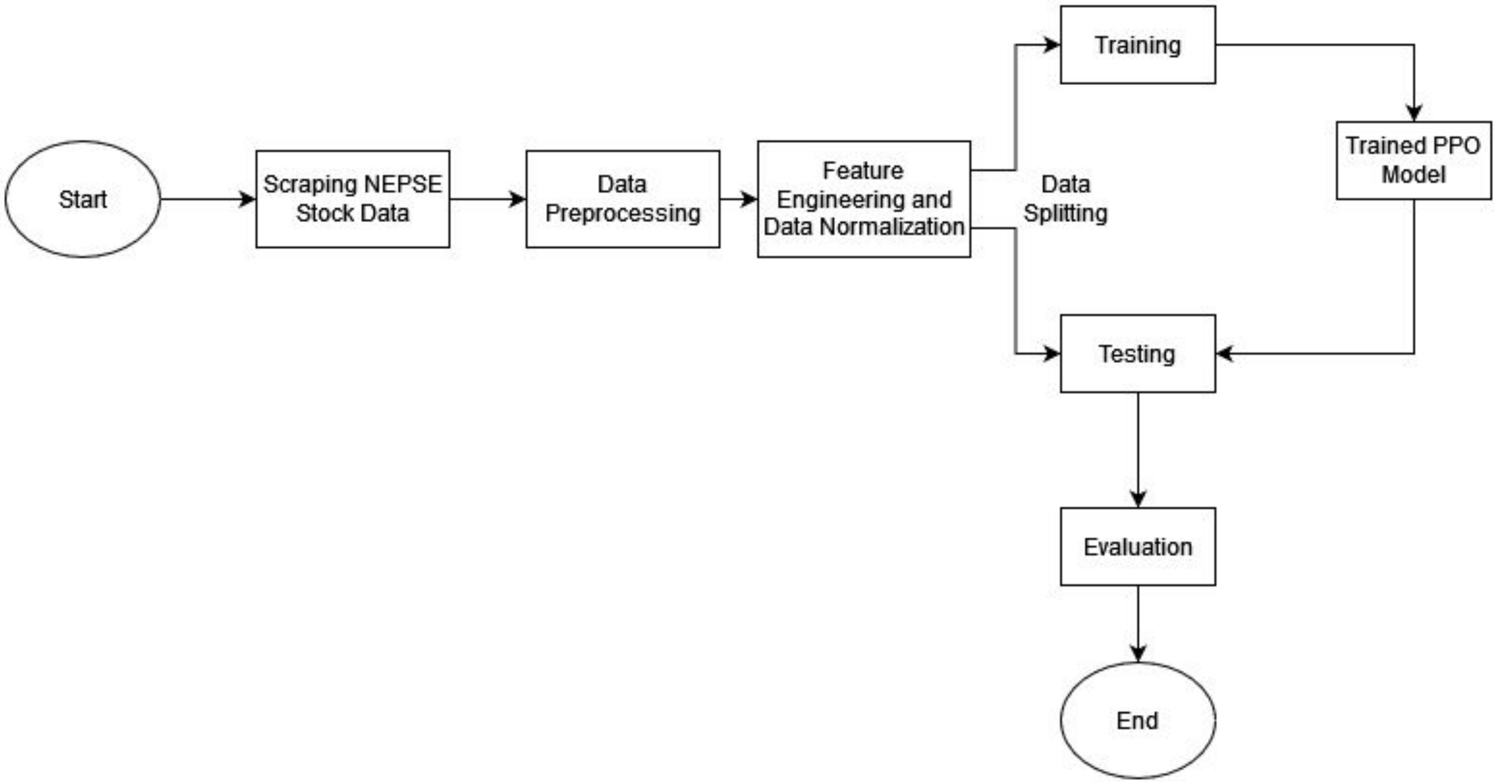


Fig: System block diagram

Methodology [4]

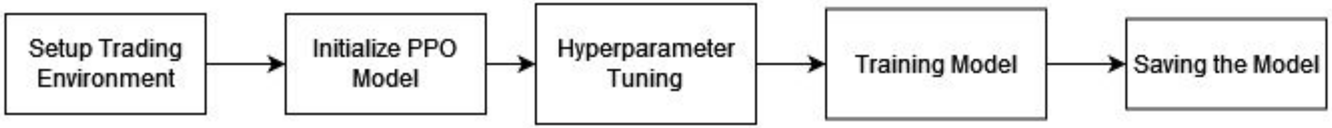


Fig: System block diagram Training

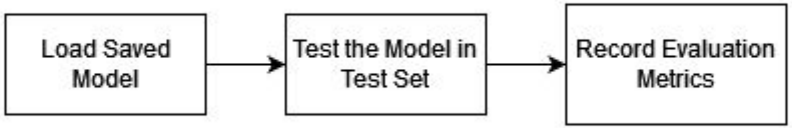


Fig: System block diagram Testing

Methodology [5]

PPO Algorithm

For the designing the trading model we will use Proximal Policy Optimization(PPO) algorithm.

PPO aims to efficiently optimize the policy to maximize cumulative rewards while ensuring stable and efficient updates.

Methodology [6]

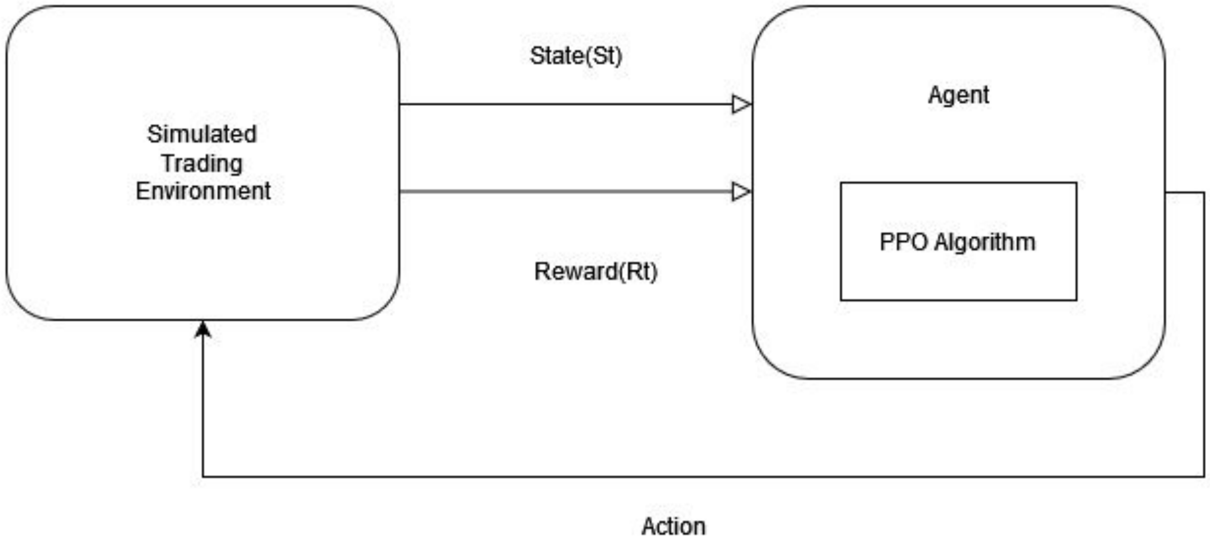


Fig: Working of the Model

Methodology [7]

Training Process:

1. **Data Collection:** Collects trajectories by interacting with the environment using the current policy.
2. **Policy Update:** Adjusts the policy using the clipped objective function.
3. **Iteration:** Repeats the process of data collection and policy update to improve performance.

Methodology [8]

Core Mechanism:

- **Clipped Surrogate Objective:** PPO uses a clipped objective function to prevent large updates that could destabilize training.
    - **Formula:**

$$L^{CLIP}(\theta) = \mathbb{E}_t \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t \right) \right]$$

    - It limits how much the new policy can differ from the old policy to avoid drastic changes.

Methodology [9]

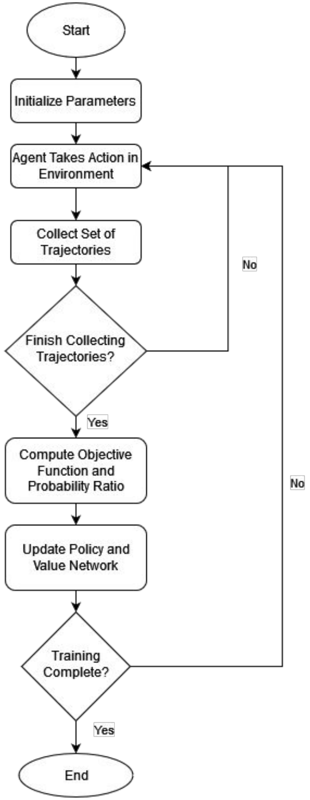


Fig: PPO Algorithm Working

Methodology [10]

- **Validation Techniques**
  - We evaluate the model's performance using metrics such as :
    - *Annualized Return or Cumulative Return*
    - *Sharpe Ratio*
    - *Maximum Drawdown*
    - *Final Portfolio Value*

Methodology [11]

- **Two Cases**
  - We evaluate the model's performance for two Case :
    - ***Case I :*** Includes data of stock price of 91 hydro power companies listed in Nepse and the evaluation period spanned from May 9, 2024, to July 29, 2024.
    - ***Case II:*** Includes data of stock price of 21 hydro power companies listed in Nepse and the evaluation period spanned from January 2, 2022, to July 29, 2024.

Results [1]

Performance of Model

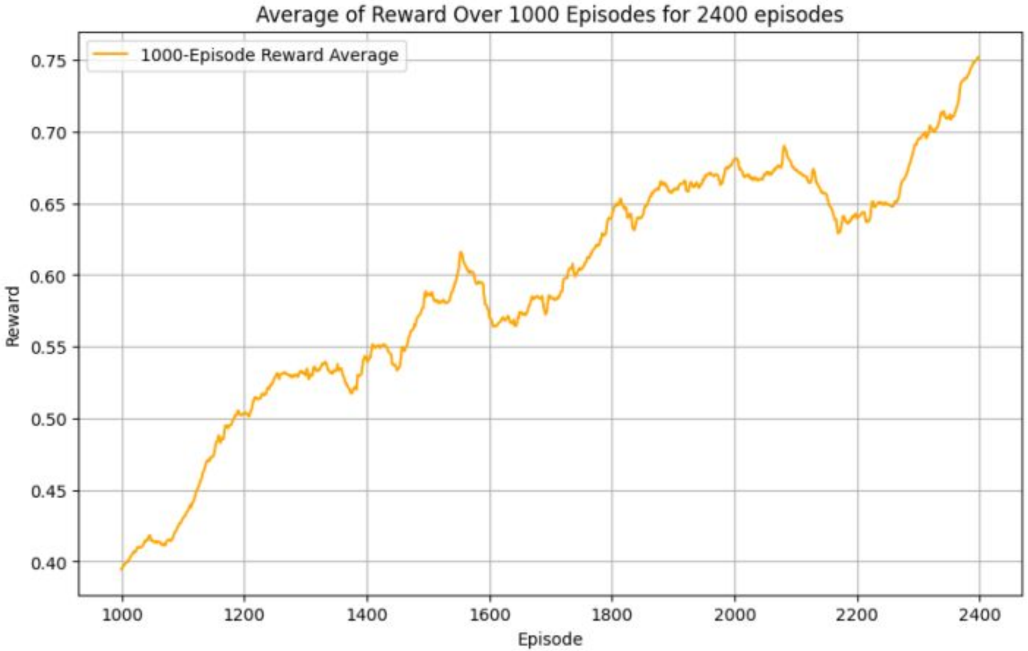


Fig: Average Reward of the Model

Results [2]

Performance of Model

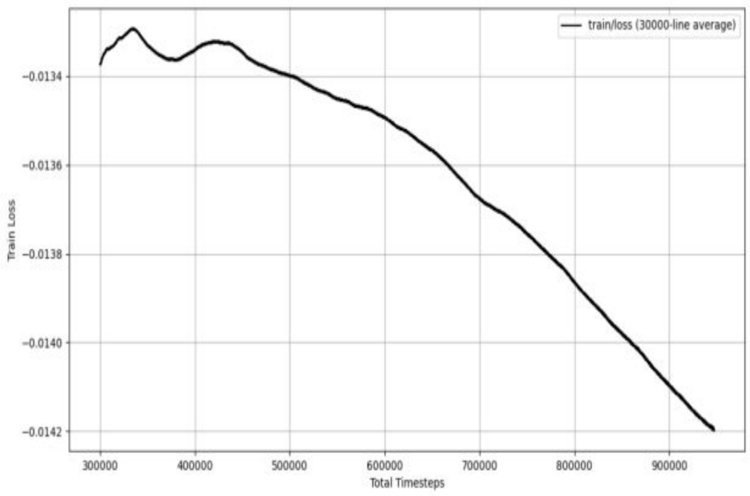


Fig: Train Loss of the Model

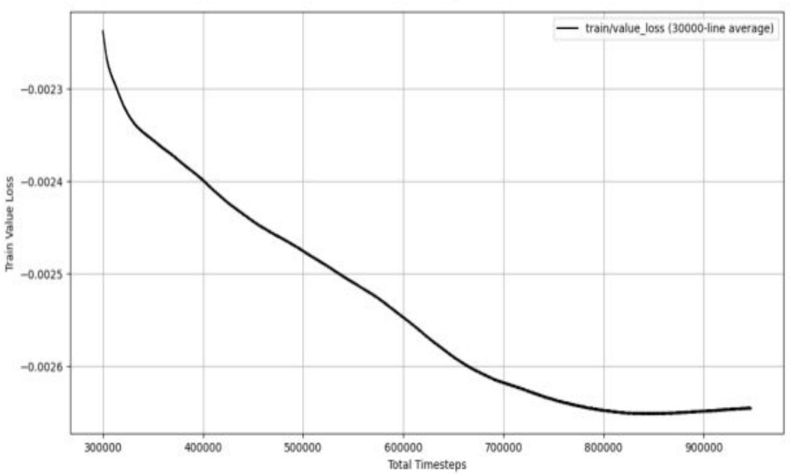


Fig: Train Value Loss of the Model

Results [3]

Performance of Model

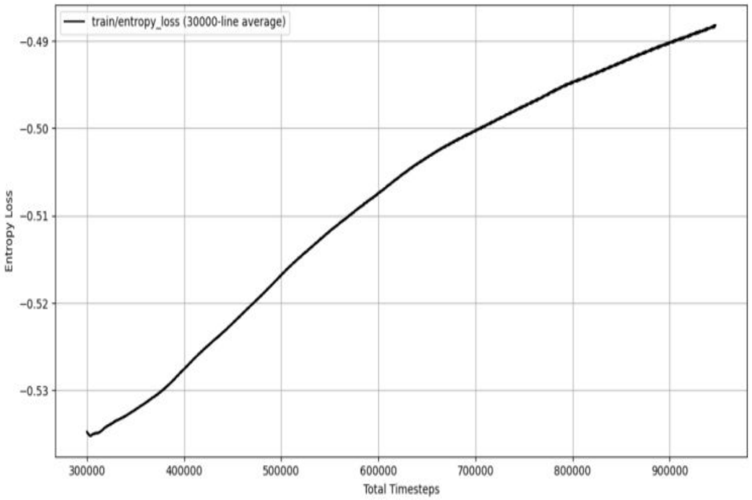


Fig: Entropy Loss of the Model

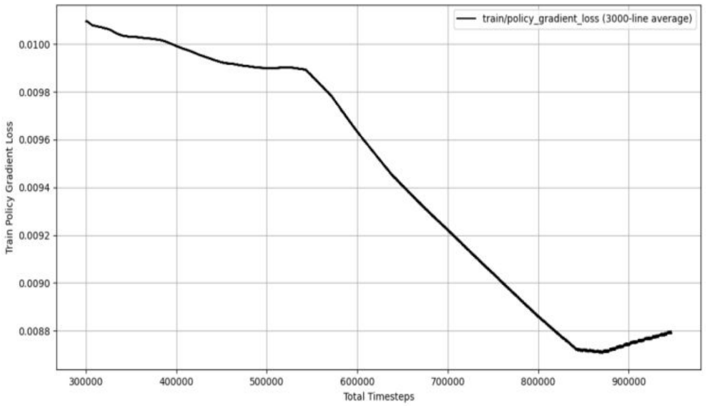


Fig: Train Policy Gradient Loss of the Model

Results [4]

Output For Best Case And Worst Case Scenario(Case I)

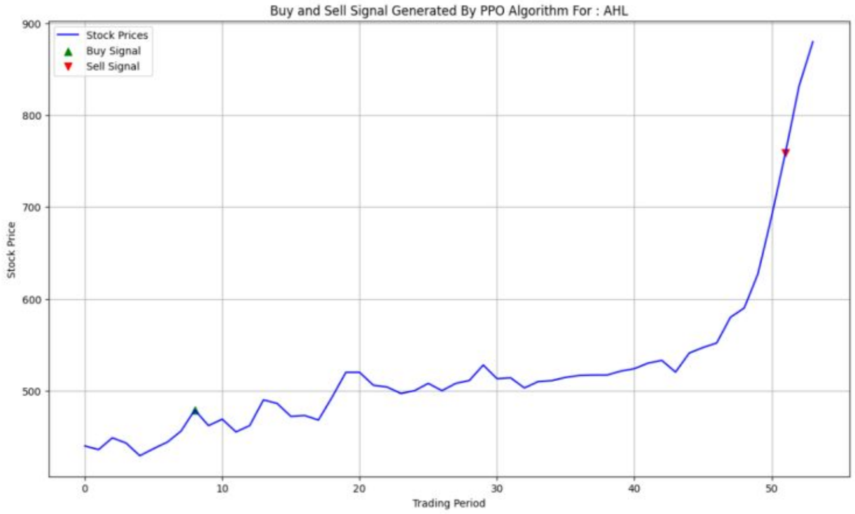


Fig : Best Case Scenario Asian Hydropower Limited(AHL) Buy and Sell Signal

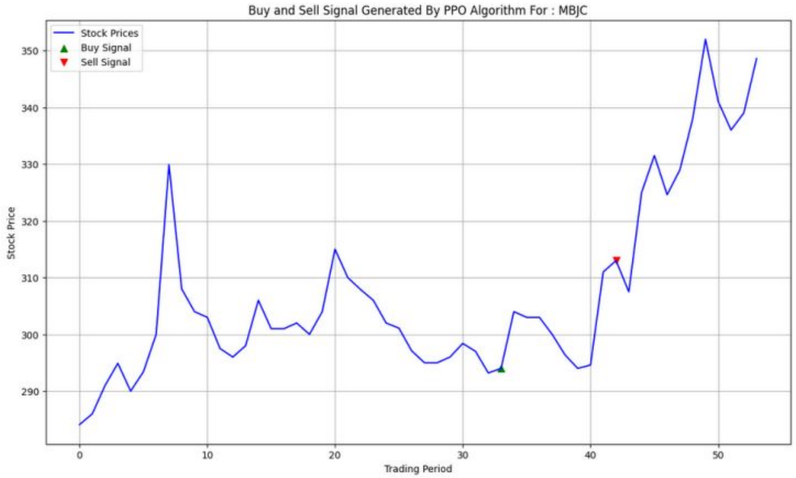


Fig: Worst Case Scenario Madhya Bhotekoshi Jalavidyut Company Limited(MBJC) Buy Sell Signal

Results [5]

Output For Best Case Scenario(Case II)

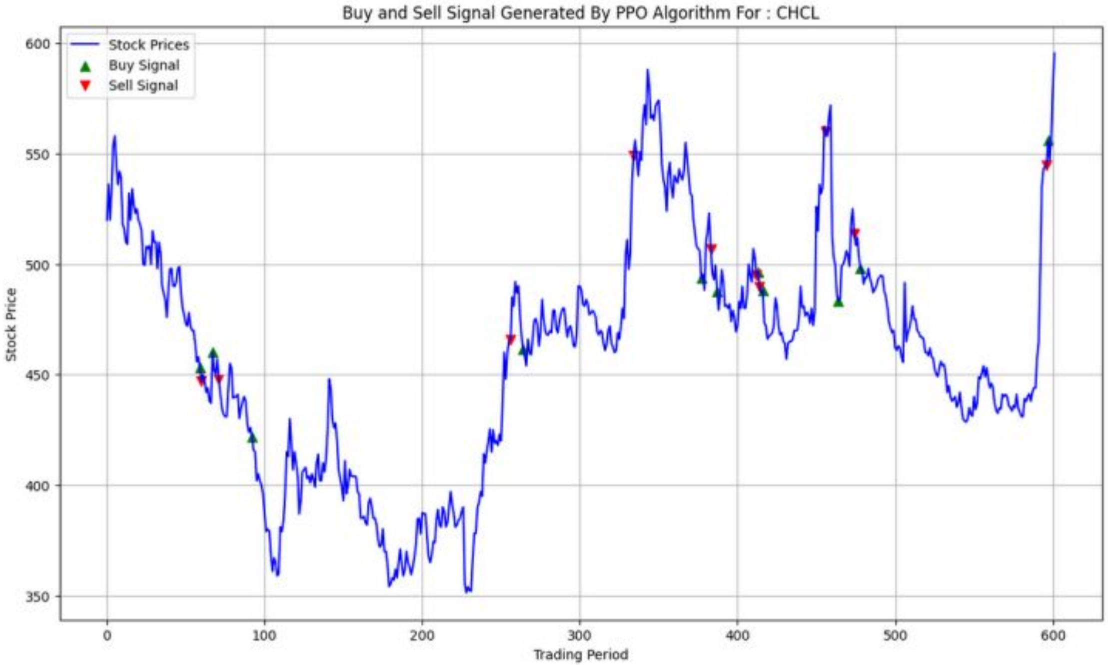


Fig: Best Case Scenario Chilime Hydro power Company Limited(CHCL) Case II

Results [5]

Output For Worst Case Scenario(Case II)

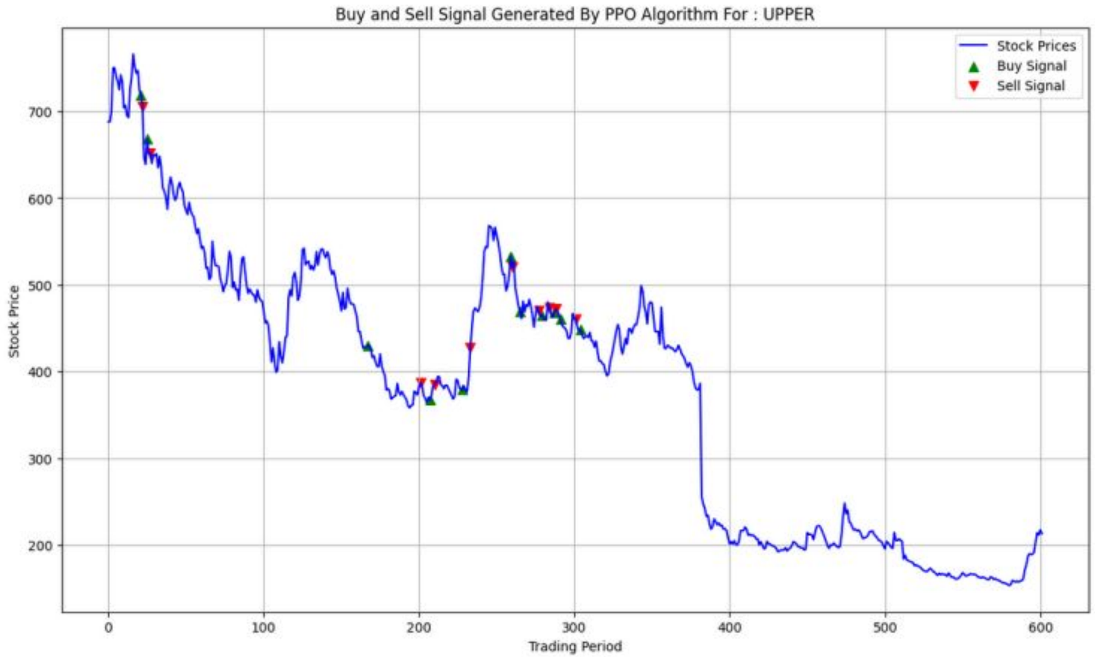


Fig: Worst Case Scenario Upper Tamakoshi Hydropower Limited(UPPER) Case II

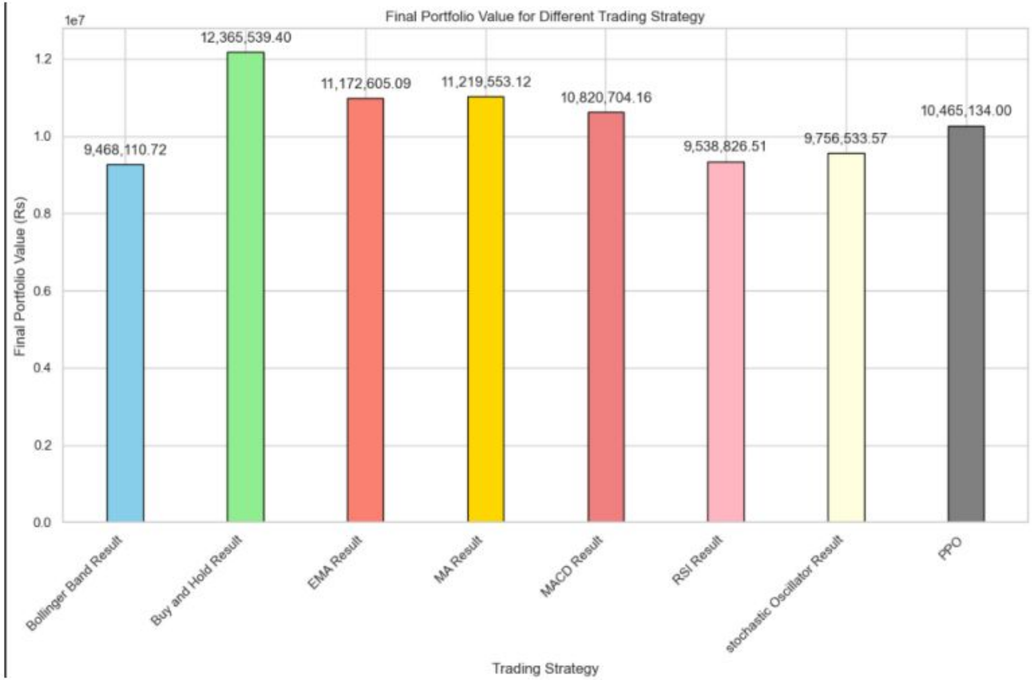
Discussion and Analysis[1]

Comparison Between PPO Algorithm with Traditional Benchmark Strategy

Benchmark Strategies Used:

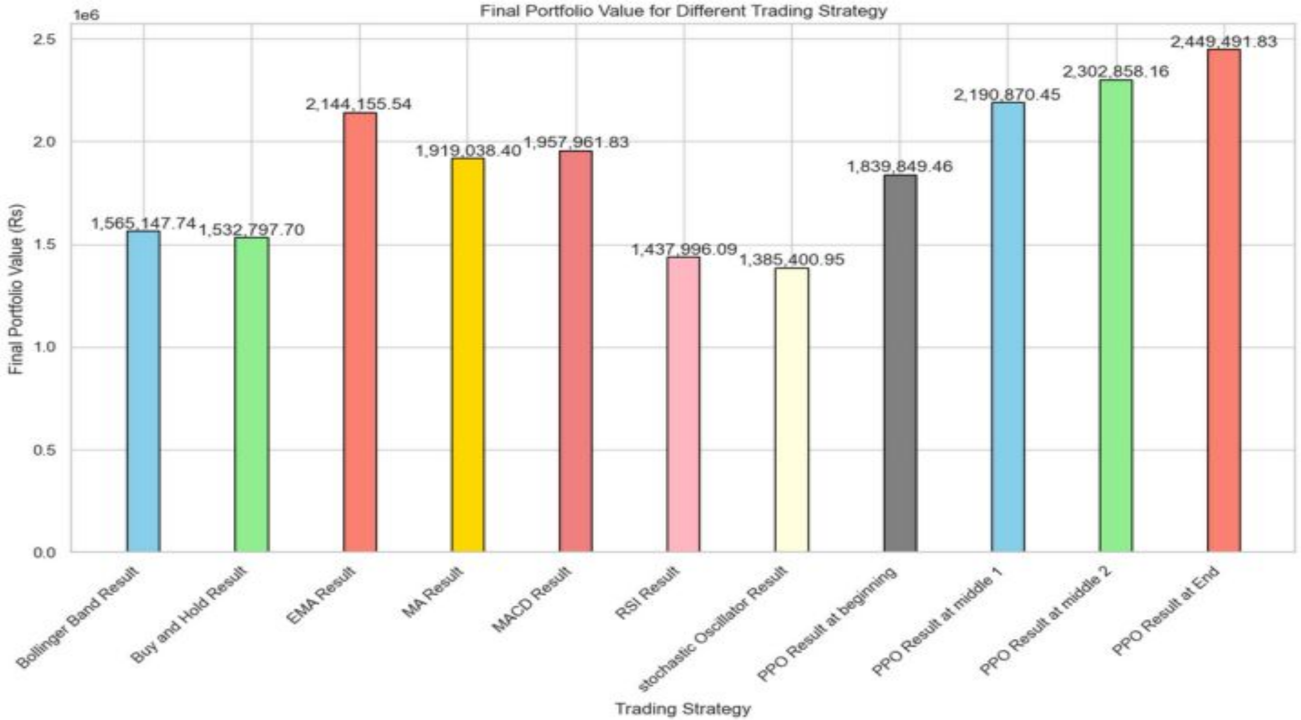
- **Bollinger Band**
- **Buy and Hold Strategy**
- **EMA Crossover**
- **MA Crossover**
- **MACD and Signal Line**
- **RSI Value**
- **SO Value**

Discussion and Analysis[2]



Bar Graph of Return of Different Strategy Case I

Discussion and Analysis[3]



Bar Graph of Return of Different Strategy Case II

Discussion and Analysis[4]

Comparison Between Theoretical and Simulated Outputs

Theoretical Expectation:

- Use trend-following indicator with volatility, which adjust to market volatility and highlight potential breakout or reversal points.
- So Theoretically during bullish trends, the strategy should aim to generate profits by executing buy orders as the market moves upward. Conversely, during bearish trends, it seeks to lock in gains or avoid losses by initiating sell orders.

Discussion and Analysis[5]

Comparison Between Theoretical and Simulated Outputs

Simulated Output:

- In **Case I** the agent's portfolio grew a lot because of a strong bullish market demonstrating that the agent effectively captured bullish market trends.
- However in **Case II**, because of the declining market, the agent was not able to grow the portfolio much in 3 years. In theory, the buying should not be done in declining market however the agent traded in declining market and could not grow the portfolio well.

Discussion and Analysis[6]

Comparison Between Theoretical and Simulated Outputs

Key area Discrepancy:

The discrepancy is seen in the bearish market trend where:

- In Case II, in the bearish market the agent's portfolio did not grow much over 3 years and traded during a declining market. This is inconsistent with the theoretical expectation, where the agent should ideally avoid trading in a declining market and and capitalize shorter bullish phases if possible.

Discussion and Analysis[7]

Comparison Between Theoretical and Simulated Outputs

Reason for Discrepancy:

Suboptimal Risk Management: The agent have not incorporated sufficient risk management techniques, such as stop-loss orders or position sizing rules, leading to significant losses in a declining market.

Ineffective Selling Strategy in Bearish Markets: The agent continued to buy even in a declining market, which led to underperformance and a failure to grow the portfolio.

Hyperparameters : Hyperparameters like learning rate, clip range, and batch size can significantly impact the agent's performance and contribute to discrepancies between theoretical expectations and simulated outputs.

Discussion and Analysis[8]

PPO Algorithm Strengths

- **Market Trend Identification:** Optimizes trading strategies through continuous market interaction.
- **Signal Generation:** Use reinforcement learning to generate reliable buy and sell signals.
- **Adaptive Learning:** Adjusts to changing market conditions by learning from new data.
- **Risk Management:** Incorporates risk-adjusted returns to minimize potential losses.

Discussion and Analysis[9]

PPO Algorithm Weaknesses

- **Overfitting Risk:** Complexity can lead to overfitting, reducing robustness in new market conditions.

- **Computational Intensity:** Requires significant computational resources, limiting real-time applicability.

Future Enhancement

- Integrate economic indicators, news sentiment, and social media data.
- Expand coverage to include all NEPSE stocks for better generalization.
- Adjust reward functions to account for market impact and real-world trading factors.
- Integrate Stop-loss actions to minimize loss.

Conclusion

- The PPO algorithm was effectively applied to trade hydropower stocks on NEPSE.
- PPO Algorithm demonstrated its ability to learn and adapt to market conditions for profitable trading decisions.

Tentative Timeline

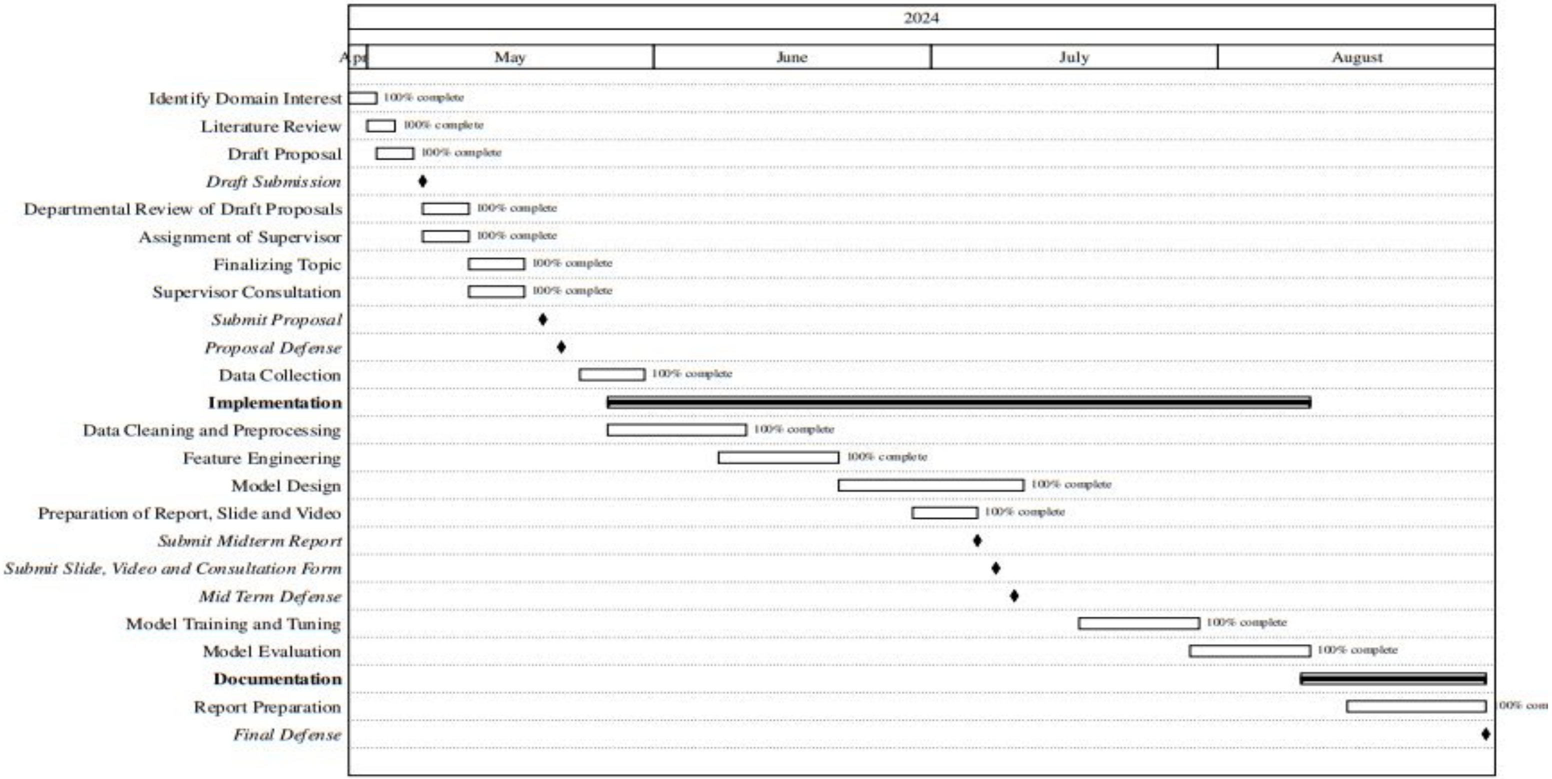


Fig: Gantt Chart

References [1]

[1] Taylan Kabbani and Ekrem Duman. Deep reinforcement learning approach for trading automation in the stock market. IEEE Access, 10:93564–93574, 2022.

[2] Moon Kyoung-Sook and KIM Hongjoong. Performance of deep learning in predicion of stock market volatility. Economic Computation & Economic Cybernetics Studies & Research, 53(2), 2019.

[3] Janak Kumar Lal. Design of Stock Trading Agent Using Deep Reinforcement Learning. PhD thesis, IOE Pulchowk Campus, 2022.

[4] Xu Wang, Sen Wang, Xingxing Liang, Dawei Zhao, Jincai Huang, Xin Xu, Bin Dai, and Qiguang Miao. Deep reinforcement learning: A survey. IEEE Transactions on Neural Networks and Learning Systems, 35(4):5064–5078, 2024.

References [2]

[5] Yuhui Wang, Hao He, and Xiaoyang Tan. Truly proximal policy optimization In Uncertainty in Artificial Intelligence, pages 113–122. PMLR, 2020.

[6] Xing Wu, Haolei Chen, Jianjia Wang, Luigi Troiano, Vincenzo Loia, and Hamido Fujita. Adaptive stock trading strategies with deep reinforcement learning methods. Information Sciences, 538:142–158, 2020.

[7] Chenglin Xiao, Weili Xia, and Jijiao Jiang. Stock price forecast based on combined model of ari-ma-ls-svm. Neural Computing and Applications, 32(10):5379–5388, 2020. 33