Proposal Report for the Degree of Master of Computer Science

**"A Proposed Hybrid Machine Learning Models for Predictive Analysis of Nepal’s Stock Prices.”**



**Bishal Maharjan**

**LC00015002348**

**Lincoln International College of Management & IT Computer Science and Multimedia Department**

**Lincoln University, Malaysia**

**Dec, 2024**

**Acknowledgements**

Initially, I would like to thank my university Lincoln International College of Management & IT for allowing me to join this program. I would also like to express my deep respect to all the module leaders who drove me throughout the semester and brought me here specially our Program Coordinator, Ms. Monica Regmi for her constant support and guidance.

I want to express my sincere gratitude to my supervisor, Prof Dr. Sudan Jha for his consistent encouragement, support and supervision which helped me in completion of my research. This research would not have been possible at all without his assistance.

I also want to express my gratitude to all the participants and professionals who supported me wholeheartedly. I want to express my gratitude to my friends for their assistance in the data collection for the survey.

**Table ofContents**

[Chapter 1: Introduction 1](#_Toc186400056)

[1.1 Background 1](#_Toc186400057)

[1.2 Problem Statement 2](#_Toc186400058)

[1.3 Research Objectives 2](#_Toc186400059)

[1.4 Significance of the study 2](#_Toc186400060)

[1.5 Scope and Limitations of the Study 3](#_Toc186400061)

[1.6.1 Scope 3](#_Toc186400062)

[1.6.2 Limitations 3](#_Toc186400063)

[Chapter 2: Literature Review 4](#_Toc186400064)

[2.1 ML Models in Stock Price Prediction 4](#_Toc186400065)

[2.2 Alternative Data Sources 4](#_Toc186400066)

[2.3 Focus on Emerging Markets 5](#_Toc186400067)

[2.4 Evaluation Metrics 5](#_Toc186400068)

[2.5 Gaps in Literature 5](#_Toc186400069)

[Chapter 3: Research Methodology 6](#_Toc186400070)

[3.1 Proposed Workflow 6](#_Toc186400071)

[3.2 Research approach 7](#_Toc186400072)

[3.3 Tools and Methods 8](#_Toc186400073)

[3.3.1 Tools 8](#_Toc186400074)

[3.3.2 Methods 9](#_Toc186400075)

[3.4 Data Collection 10](#_Toc186400076)

[Chapter 4: Epilogue 11](#_Toc186400077)

[4.1 Research Timeline 11](#_Toc186400078)

[4.2 Expected Output 12](#_Toc186400079)

[References 13](#_Toc186400080)

# Chapter 1: Introduction

## 1.1 Background

The integration of machine learning (ML) into stock price prediction has revolutionized financial analysis, offering enhanced accuracy and insights compared to traditional forecasting methods. Stock markets, as vital components of economic systems, reflect the underlying dynamics of demand, supply, investor sentiment, and macroeconomic variables. Accurate forecasting of stock prices is crucial for investors, policymakers, and financial institutions, enabling them to make informed decisions that can mitigate risks and maximize returns. However, in emerging markets like Nepal, the challenge of stock price prediction is amplified due to the unique economic environment, data scarcity, and higher market volatility [1][2].

Traditional models such as ARIMA and linear regression have limitations in capturing the complex, nonlinear behaviors exhibited in financial time series data. Advances in ML techniques, including Long Short-Term Memory (LSTM) networks, Gradient Boosting Machines (GBMs), and Support Vector Machines (SVMs), have demonstrated superior capabilities in this area. For example, LSTM models are very effective in preserving long-term dependencies and discovering hidden patterns in sequential data [3][5]. Meanwhile, hybrid approaches such as CNN-LSTM combine feature extraction with temporal prediction, further enhancing model robustness [7].

Despite the success of ML in developed markets, research on its application to Nepal's stock market, specifically the Nepal Stock Exchange (NEPSE), is scarce. NEPSE is characterized by unique challenges, including limited historical data, less diversification of market participants, and significant socio-political influences. These factors necessitate a localized approach to stock price forecasting that considers both structured and unstructured data sources, such as economic indicators and news sentiment [4][5]. This study aims to address these gaps by developing a comprehensive ML framework tailored to NEPSE, leveraging state-of-the-art techniques to improve predictive accuracy and reliability.

## 1.2 Problem Statement

Accurately predicting stock prices in emerging markets is a formidable challenge due to several key factors. First, the limited availability of historical financial data and macroeconomic indicators in markets like Nepal restricts the efficacy of traditional and modern predictive models [1][2]. Second, the Nepal Stock Exchange (NEPSE) exhibits high volatility, reflecting the country’s unique socio-economic and political conditions, which are often unaccounted for in conventional approaches [4]. Likewise, existing stock prediction frameworks primarily focus on developed economies, resulting in a lack of robust models tailored to the distinct characteristics of emerging markets like Nepal [5].

This research aims to address these challenges by leveraging advanced machine learning techniques and integrating unconventional data sources, such as sentiment analysis of local financial news, to enhance prediction accuracy and reliability. Stock price prediction in emerging markets is hindered by inadequate data, high volatility, and the lack of robust predictive models tailored to these economies. While existing studies have demonstrated the potential of machine learning for financial forecasting, most focus on developed markets. This creates a significant research gap, particularly in Nepal’s context, where unique market dynamics and data availability issues pose additional challenges [2][4].

## 1.3 Research Objectives

The research aims to meet the following objectives.

1. To compare a machine learning framework to predict stock prices in NEPSE.
2. To evaluate the performance of advanced ML models, such as LSTM, ARIMA, and SVM, in forecasting stock prices.
3. To incorporate sentiment analysis from local financial news and macroeconomic indicators to enhance prediction accuracy.
4. To provide actionable insights for investors, financial institutions, and policymakers in Nepal.

## 1.4 Significance of the study

This research’s significance lies in the fact that it has the potential to contribute valuable insights in the field of Nepal’s financial ecosystem. Some important points that demonstrate its significance include:

1. **Investor Decision-Making:** This study can help in developing a robust predictive framework in line with NEPSE enables investors to make informed decisions, reduce risk and improve returns by accurately predicting market movements.
2. **Financial Institutions:** This study’s predictive models can help financial institutions optimize portfolio management, develop innovative financial products, and improve risk assessment strategies.
3. **Policymakers:** By integrating alternative data sources such as news sentiment and macroeconomic indicators, this study provides policymakers with a deeper understanding of market dynamics, thereby facilitating better economic regulation and policy formulation.
4. **Academic Contributions:** The research fills a critical gap in the literature by focusing on Nepal’s underexplored financial market, advancing the discourse on ML applications in emerging economies.
5. **Technological Innovation:** Leveraging advanced ML techniques such as LSTM and CNN-LSTM models drives innovation in predictive analytics, providing a methodological framework that can be extended to other emerging markets.
6. **Economic Development:** Accurate stock market forecasting contributes to economic stability by attracting domestic and foreign investments, thereby boosting growth in Nepal's financial ecosystem.

## 1.5 Scope and Limitations of the Study

### 1.6.1 Scope

1. **Model Development:**

The study focuses on analyzing advanced machine learning techniques such as LSTM, GBM, and CNN-LSTM for stock price prediction.

1. **Geographic Focus:**

The study focuses exclusively on the Nepal Stock Exchange (NEPSE), providing insights relevant to Nepal’s unique economic and financial ecosystem.

1. **Data Sources:**

The research utilizes a combination of historical stock data, macroeconomic indicators, and sentiment analysis of news headlines.

### Limitations

1. **Data Availability:**

Limited access to high-quality and granular data for NEPSE may constrain the robustness of the models.

1. **Market Volatility:**

The unpredictable nature of emerging markets like Nepal may introduce noise and reduce prediction accuracy.

1. **Generalizability:**

The framework developed may require significant adaptation for application to other emerging markets due to differences in market dynamics.

# Chapter 2: Literature Review

The use of machine learning (ML) for financial forecasting has been extensively studied in the literature, highlighting its advantages over traditional statistical methods. This review examines studies on machine learning-based stock price prediction with a particular focus on emerging markets and Nepal’s NEPSE, the models used, alternative data sources, and valuation metrics.

## ****2.1 ML Models in Stock Price Prediction****

The effectiveness of ML models in capturing non-linear patterns in financial data has been well-documented. LSTM, a type of recurrent neural network, has emerged as a leading method for time-series prediction due to its ability to retain long-term dependencies in sequential data [3]. Studies have demonstrated the superior accuracy of LSTM in stock price prediction compared to traditional methods like ARIMA [2].

Hybrid models, such as CNN-LSTM, further enhance predictive capabilities by integrating convolutional layers for feature extraction with LSTM layers for temporal modeling. Research comparing standalone and hybrid models indicates that hybrid approaches consistently outperform simpler architectures, making them a promising avenue for financial forecasting [7].

Support Vector Machines (SVM) and Gradient Boosting Machines (GBM) have also been utilized in stock prediction tasks, with their robustness in handling non-linear data proving advantageous in volatile markets [8]. Random forest and ensemble methods have shown comparable results, particularly when paired with effective feature selection techniques [8].

The integration of deep learning techniques has significantly enhanced stock market forecasting. Sahu et al. [9] highlight the application of machine learning (ML), deep learning (DL), reinforcement learning (RL), and deep reinforcement learning (DRL) in predicting stock market behavior. These techniques leverage artificial intelligence to extract high-level patterns and provide robust solutions for algorithmic trading. Similarly, Hu et al. [10] emphasize the efficacy of models like CNN, LSTM, and RNN, among others, in capturing the nonlinear dynamics of financial data. Their study notes the rising trend in hybrid models combining multiple deep learning architectures for improved accuracy.

Peng et al. [11] proposed hybrid models such as ARIMA-MLP and ARIMA-RNN for forecasting stock indices in SAARC countries. These models outperform standalone ARIMA or neural network models by effectively addressing the noisy and nonlinear nature of financial data. Joshi [12] demonstrated the application of technical trading tools like candlesticks, Bollinger Bands, and MACD in predicting the Nepal Stock Exchange's market trends. This combination yielded more reliable buy-sell-hold decisions compared to individual indicators.

Garlapati et al. [13] employed ARIMA and Facebook Prophet models to predict stock prices, focusing on preprocessing data for enhanced performance. Their study validated the potential of time-series models in analyzing patterns and trends for effective stock price forecasting.

Ho et al. employed ARIMA, neural networks (NN), and long short-term memory networks (LSTM) to predict stock prices for Bursa Malaysia. The study found that LSTM models outperformed other methods in accuracy during the pandemic, achieving over 90% accuracy. The research emphasizes the dynamic nature of stock prices and the necessity of robust prediction methods under volatile market conditions [15].

Agarwal et al. explored RNNs and LSTMs for predicting time series data, emphasizing their capability to model temporal dependencies. The research demonstrated that LSTM addresses the vanishing gradient problem inherent in traditional RNNs, achieving a prediction accuracy of 91.97% for testing datasets. The paper underlined the superior performance of LSTM over traditional machine learning methods, such as ARIMA and regression models, in capturing the temporal nuances of stock data [16].

Pokhrel et al. introduced a unified framework, Deep-SDM, which integrates LSTM, GRU, and CNN architectures for sequential data modeling. This framework incorporates data exploration, hyperparameter tuning, and statistical validation, enabling a comprehensive approach to predictive analytics. The research demonstrated its application in predicting the Nepal Stock Exchange (NEPSE) index, achieving significant accuracy improvements. This study highlights the importance of modular and scalable frameworks in financial data analysis [17].

## ****2.2 Alternative Data Sources****

Incorporating alternative data sources, such as sentiment analysis of news headlines, has proven to be a valuable enhancement to ML-based stock prediction models. For example, the use of sentiment analysis tools like VADER and TextBlob has been shown to improve prediction accuracy by capturing market sentiment and its impact on stock price movements [5].

The use of macroeconomic indicators such as interest rates, inflation, and GDP growth further enriches prediction models by providing a comprehensive view of market dynamics. However, studies in emerging markets like Nepal often face challenges in accessing high-quality and granular data, limiting the potential of these approaches [4].

## ****2.3 Focus on Emerging Markets****

Emerging markets present unique challenges for ML-based stock prediction due to high volatility, limited historical data, and significant socio-political influences. Research focusing on these markets has identified gaps in the adaptability of ML frameworks developed for developed markets. For instance, Bhandari et al. highlighted the need for tailored models that address the specific characteristics of NEPSE, such as its smaller market size and higher dependency on macroeconomic factors [2].

Recent work by Dahal et al. explored the role of sentiment analysis in predicting NEPSE index movements, demonstrating that sentiment scores significantly influence model accuracy [5]. However, the study also emphasized the need for further research to integrate other alternative data sources and evaluate the robustness of models under different economic scenarios.

## ****2.4 Evaluation Metrics****

Studies have shown that LSTM and hybrid models consistently achieve lower RMSE and higher R-squared values compared to traditional methods. However, emerging markets like Nepal often introduce higher levels of noise, requiring additional metrics to evaluate robustness and scalability [3].

Evaluation metrics play a critical role in assessing the performance of ML models in stock price prediction. Commonly used metrics include Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared for regression models. For classification tasks, metrics such as sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC) are utilized [6].

Shah et al. [9] proposed a hybrid deep learning model combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to predict the closing prices of the Nifty 50 index. This architecture leverages CNN for spatial feature extraction and LSTM for capturing temporal dependencies in stock market data, resulting in high prediction accuracy. The study builds on previous works that employed machine learning and deep learning techniques, such as Random Forests, ensemble methods, and sentiment analysis, for stock market predictions. This framework demonstrates the efficacy of integrating spatial and temporal learning for financial forecasting [9].

## ****2.5 Gaps in Literature****

While significant advancements have been made in ML-based stock prediction, gaps remain, particularly in the context of emerging markets. Limited research has focused on Nepal's NEPSE, with most studies relying on frameworks developed for larger, more developed markets. Additionally, the integration of alternative data sources such as local news sentiment and macroeconomic indicators remains underexplored, presenting an opportunity for novel contributions [4][5].

Most existing studies focus on either traditional technical analysis or machine learning independently, with limited exploration of their combined potential. Additionally, while various machine learning models have been tested, the specific impact of different candlestick patterns on prediction accuracy using ensemble methods has not been thoroughly investigated [14].

By addressing these gaps, this research aims to develop a comprehensive ML framework tailored to NEPSE, leveraging state-of-the-art models and data sources to improve prediction accuracy and reliability.

# Chapter 3: Research Methodology

## 3.1 Proposed Workflow

The proposed workflow for this research integrates state-of-the-art machine learning techniques with data-driven approaches tailored to the Nepal Stock Exchange (NEPSE). The workflow is structured into the following stages:

1. **Data Acquisition:**

* Collect historical stock prices (open, high, low, close, volume) from NEPSE.
* Gather macroeconomic indicators such as GDP, inflation, interest rates, and exchange rates from sources like the Nepal Rastra Bank.
* Extract sentiment data from financial news articles and headlines using web scraping and APIs (e.g., VADER, TextBlob) [5].

1. **Data Preprocessing:**

* Normalize numerical data to ensure consistent scaling.
* Impute missing values using interpolation or advanced methods like KNN imputation.
* Perform text preprocessing for sentiment analysis, including tokenization, stop-word removal, and lemmatization [5][8].

1. **Feature Engineering:**

* Derive technical indicators such as moving averages, RSI, and Bollinger Bands from historical stock data [1][3].
* Use Principal Component Analysis (PCA) to reduce dimensionality and identify the most impactful features.
* Incorporate sentiment scores as additional input features for ML models [5].

1. **Model Development:**

* Train baseline models (ARIMA, Linear Regression) to establish a benchmark.
* Develop advanced ML models:
  + 1. LSTM for capturing temporal dependencies in stock price data [2][4].
    2. GBM for robust, feature-intensive modeling [3].
    3. Hybrid CNN-LSTM for enhanced feature extraction and sequential modeling [7].

1. **Model Optimization:**

* Use grid search or Bayesian optimization for hyperparameter tuning.
* Implement techniques like dropout and early stopping to prevent overfitting [1].

1. **Model Evaluation:**

* Assess models using metrics such as RMSE, MAPE, and R-squared for regression, and AUC-ROC for classification tasks [6].
* Perform cross-validation to ensure model robustness.

1. **Insights and Deployment:**

* Analyze feature importance and model outputs to provide actionable insights for investors and policymakers.
* Deploy the model in a decision-support system, if applicable, to facilitate real-time predictions [5].

1. **Validation and Feedback Loop:**

* Validate predictions with real-world stock data to refine the models.
* Use feedback to iteratively improve feature selection, data preprocessing, and model architecture.

## 3.2 Research approach

The research adopts a data-driven approach, integrating advanced machine learning techniques with empirical analysis. This approach is structured into the following phases:

1. **Exploratory Phase:**

* Conduct a comprehensive review of existing literature to identify research gaps, particularly in the context of NEPSE and emerging markets.
* Define research objectives, scope, and limitations based on insights from previous studies and the unique challenges of Nepal's financial market.

1. **Data Acquisition and Preparation:**

* Collect multi-dimensional datasets, including historical stock prices, macroeconomic indicators, and sentiment data from news headlines.
* Apply preprocessing techniques such as normalization, imputation, and feature engineering to prepare the data for analysis.

1. **Model Development and Training:**

* Develop baseline models like ARIMA and linear regression for benchmarking.
* Build and fine-tune advanced ML models (e.g., LSTM, GBM, CNN-LSTM) tailored to NEPSE’s market characteristics.
* Incorporate macroeconomic and sentiment analysis features into the models to capture holistic market dynamics.

1. **Evaluation and Validation:**

* Use a robust set of evaluation metrics (e.g., RMSE, MAPE, R-squared) to assess model performance.
* Perform cross-validation and back-testing to ensure model reliability and generalizability.

1. **Insights and Deployment:**

* Generate actionable insights for investors and policymakers by analyzing model outputs and feature importance.
* Develop a prototype decision-support system, if feasible, to deliver real-time stock predictions.

1. **Feedback and Iterative Refinement:**
   * Validate findings against real-world stock performance and refine models based on observed discrepancies.
   * Incorporate stakeholder feedback to enhance the practical relevance and usability of the research outcomes.

This approach ensures a systematic exploration of NEPSE's stock market dynamics while leveraging machine learning's predictive capabilities to address existing research gaps.

## 3.3 Tools and Methods

### ****3.3.1 Tools****

1. **Programming Languages and Libraries:**

**Python:** Primary programming language for data processing, model development, and analysis. Libraries such as TensorFlow, Keras, and PyTorch for building machine learning models. Scikit-learn for implementing traditional ML models and preprocessing techniques. Pandas and NumPy for data manipulation and numerical computations.

1. **Sentiment Analysis Tools:**

**VADER (Valence Aware Dictionary and sEntiment Reasoner):** A rule-based tool for sentiment analysis of textual data.

**TextBlob:** A library for text processing and sentiment score generation.

1. **Visualization Tools:**

**Matplotlib and Seaborn:** For creating comprehensive visualizations of datasets and model outputs.

Dash or Plotly for interactive data dashboards.

1. **Data Sources and Collection Tools:**

APIs and web scraping tools (e.g., Beautiful Soup, Scrapy) for extracting stock data and financial news.

Public datasets from Nepal Rastra Bank and NEPSE.

1. **Computational Infrastructure:**

Local GPU-enabled machines or cloud platforms such as Google Colab and AWS for training computationally intensive models.

### ****3.3.2 Methods****

1. **Data Preprocessing:**

Data cleaning and normalization ensure consistency and reduce noise by standardizing and correcting errors. Missing values are handled through interpolation and imputation techniques. For textual sentiment analysis, preprocessing steps like tokenization, stop-word removal, and lemmatization are used to break text into meaningful components, eliminate irrelevant words, and reduce words to their base form, improving analysis accuracy.

1. **Feature Engineering:**

Derived features, such as technical indicators like moving averages, Bollinger Bands, and RSI, are created to capture key market trends. Principal Component Analysis (PCA) is then applied to reduce dimensionality, retaining the most informative features while simplifying the data. Additionally, sentiment scores derived from financial news are incorporated as extra predictors to enhance the model's ability to capture market sentiment and improve forecasting accuracy.

1. **Model Development:**

Baseline models like ARIMA and linear regression are used for benchmarking. Advanced models include LSTM for capturing temporal dependencies in time-series data, GBM for handling feature-intensive regression tasks, and a hybrid CNN-LSTM model that combines convolutional feature extraction with sequential data modeling for improved performance.

1. **Model Optimization:**

Grid search and Bayesian optimization are used to fine-tune hyperparameters like learning rates, number of layers, and batch sizes for optimal model performance. Regularization techniques, such as L1/L2 regularization and dropout, are applied to prevent overfitting, ensuring the model generalizes well to new data.

1. **Model Evaluation:**

Metrics like RMSE, MAPE, and R-squared are used to evaluate regression model performance, providing insights into accuracy and fit. Cross-validation is employed to assess the model's robustness and generalizability, ensuring it performs well on unseen data.

1. **Workflow Automation:**

Python scripts and pipelines are used to automate data collection, preprocessing, and model training, streamlining the workflow. Reusable modules are created to ensure scalability and reproducibility, making future research more efficient and consistent.

## 3.4 Data Collection

The data collection process for this research involves a multi-dimensional approach to ensure comprehensive coverage of variables influencing stock prices:

1. **Historical Stock Data:** Daily price data (open, high, low, close) and trading volumes from the Nepal Stock Exchange (NEPSE) are gathered as primary time-series inputs for stock price forecasting.
2. **Macroeconomic Indicators:** Variables such as inflation rates, interest rates, GDP growth, and exchange rates are sourced from the Nepal Rastra Bank and other government publications. These indicators provide contextual information on the broader economic environment influencing stock market dynamics.
3. **Sentiment Data:** Textual data, including news articles, headlines, and social media mentions relevant to the Nepalese financial market, are collected using APIs and web scraping tools. Examples include VADER for sentiment scoring and TextBlob for polarity analysis.
4. **Data Validation:** Data from different sources are cross-verified to ensure reliability. Inconsistencies and missing data are identified and addressed through preprocessing steps.
5. **External Datasets:** Global financial indices and data from related regional markets are optionally included to study their correlation and influence on NEPSE.

This structured data collection ensures the inclusion of both quantitative and qualitative dimensions, providing a robust foundation for predictive modeling.

# Chapter 4: Epilogue

## 4.1 Research Timeline

The timeline of the research is detailed in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **Task** | **Start Date** | **End Date** | **Duration** |
| Research Topic Selection | Nov 14 | Nov 25 | 11 days |
| Literature Review | Nov 25 | Dec 25 | 30 days |
| Proposal Writing | Dec 25 | Jan 28 | 33 days |
| Data Collection and Preprocessing | Jan 28 | Feb 20 | 22 days |
| Algorithm Implementation | Feb 20 | Feb 28 | 8 days |
| Evaluation and Analysis | Feb 28 | Mar 11 | 43 days |
| Final Review and Submission | Mar 11 | Mar 12 | 1 day |

*Table 1: Research Timeline*

## 4.2 Expected Output

The anticipated outcomes of this research include:

1. **Development of a Predictive Framework:**

A robust machine learning framework tailored to the Nepal Stock Exchange (NEPSE) that effectively combines historical stock data, macroeconomic indicators, and sentiment analysis from financial news. This framework will leverage advanced models like LSTM and hybrid CNN-LSTM to address NEPSE-specific challenges such as limited data availability and high market volatility.

1. **Comparative Performance Analysis:**

A comprehensive evaluation of various ML models (e.g., LSTM, GBM, and SVM) in predicting NEPSE stock prices, including baseline comparisons with traditional methods like ARIMA and linear regression. Insights into model effectiveness based on metrics like RMSE, MAPE, and R-squared, aiding stakeholders in selecting appropriate models for different applications.

1. **Integration of Sentiment Analysis:**

Demonstration of how incorporating sentiment scores from financial news enhances prediction accuracy. Quantitative analysis of the impact of investor sentiment on stock price movements in Nepal's emerging market context.

1. **Feature Importance Insights:**

Identification of the most influential features (e.g., technical indicators, macroeconomic variables, and sentiment scores) contributing to stock price predictions. Recommendations for future research and practical implementations based on feature importance analysis.

1. **Decision-Support Tools:**

A prototype or proof-of-concept decision-support system for investors and policymakers, allowing real-time predictions and data visualization. Suggestions for improving investment strategies and policy frameworks based on predictive outputs.

1. **Academic Contributions:**

Filling the research gap by focusing on NEPSE, contributing to the global discourse on machine learning applications in emerging financial markets. A methodological guide for adapting predictive frameworks to other emerging markets with similar challenges.

These outputs will significantly benefit various stakeholders, including investors, financial institutions, policymakers, and researchers, by providing actionable insights and advancing the state-of-the-art in stock price prediction for emerging markets.

# References

1. A. Singh and S. Shakya, "Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE," Procedia Computer Science, 2020.
2. U. Bhandari et al., "Predicting stock market index using LSTM," Machine Learning with Applications, 2022.
3. D. Iyanuoluwa et al., "Machine Learning for Stock Market Forecasting: A Review of Models and Accuracy," Finance & Accounting Research Journal, 2024.
4. P. Shrestha, "Application of Machine Learning and Deep Learning Techniques for Nepal Stock Market Price Prediction," Thesis, 2021.
5. K. Dahal et al., "Predicting the Direction of NEPSE Index Movement with News Headlines Using Machine Learning," Econometrics, 2024.
6. T. Aldhyani et al., "Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms," Electronics, 2022.
7. K. Alkhatib et al., "A New Stock Price Forecasting Method Using Active Deep Learning Approach," J. Open Innov. Technol. Mark. Complex., 2022.
8. T. Shen, "Machine Learning Applications in Stock Price Prediction," Proceedings of Finance in the Age of Environmental Risks and Sustainability, 2024.
9. S. K. Sahu, A. Mokhade, and N. D. Bokde, "An overview of machine learning, deep learning, and reinforcement learning-based techniques in quantitative finance: Recent progress and challenges," Applied Sciences, vol. 13, no. 3, pp. 1956, Feb. 2023, doi: 10.3390/app13031956.
10. Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," Applied System Innovations, vol. 4, no. 1, pp. 9, Feb. 2021, doi: 10.3390/asi4010009.
11. Z. Peng et al., "An application of hybrid models for weekly stock market index prediction: Empirical evidence from SAARC countries," Complexity, vol. 2021, Art. no. 5663302, pp. 1–10, Dec. 2021, doi: 10.1155/2021/5663302.
12. D. L. Joshi, "Enhancing investment strategies: A comprehensive technical analysis approach for informed buy-sell-hold decisions in NEPSE," Nepal Journal of Multidisciplinary Research, vol. 6, no. 4, pp. 72–80, Dec. 2023, doi: 10.3126/njmr.v6i4.62009.
13. A. Garlapati et al., "Stock price prediction using Facebook Prophet and ARIMA models," in Proc. 6th Int. Conf. Convergence Technol. (I2CT), Pune, India, Apr. 2021, pp. 1–6, doi: 10.1109/I2CT51068.2021.9418057.
14. *“Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme,” IEEE Journals & Magazine, Available:* [*https://ieeexplore.ieee.org/document/9481924/*](https://ieeexplore.ieee.org/document/9481924/)
15. *M. K. Ho, H. Darman, and S. Musa, “Stock Price Prediction Using ARIMA, Neural Network and LSTM Models,” J. Phys. Conf. Ser., vol. 1988, no. 1, pp. 1–10, 2021. [Online]. Available:* [*https://doi.org/10.1088/1742-6596/1988/1/012041*](https://doi.org/10.1088/1742-6596/1988/1/012041)*.*
16. *H. Agarwal, G. Mahajan, A. Shrotriya, and D. Shekhawat, “Predictive Data Analysis: Leveraging RNN and LSTM Techniques for Time Series Dataset,” Procedia Comput. Sci., vol. 235, pp. 979–989, 2024. [Online]. Available:* [*https://doi.org/10.1016/j.procs.2024.04.093*](https://doi.org/10.1016/j.procs.2024.04.093)*.*
17. *N. R. Pokhrel, K. R. Dahal, R. Rimal, H. N. Bhandari, and B. Rimal, “Deep-SDM: A Unified Computational Framework for Sequential Data Modeling Using Deep Learning Models,” Software, vol. 3, no. 1, pp. 47–61, 2024. [Online]. Available: https://doi.org/10.3390/software3010003.*