Comparative Analysis of Neural Network Techniques for Forecasting Stock Market Trends in New Zealand and Australia

Project Report
School of Computing

May Thu Tun
Student ID - 1570765

Supervisors
Dr. Soheil Varastehpour & Dr. Masoud Shakiba

Abstract

This project conducts a comprehensive comparative analysis of various deep learning models, including LSTM, GRU, CNN, CNN-LSTM, and GRU-LSTM, for predicting stock market trends in the New Zealand and Australian markets. Stock market prediction is crucial due to the inherent volatility and non-stationary nature of financial markets, which makes it a challenging task for investors and traders who need reliable tools to inform their investment strategies and trading decisions. Accurate predictions can significantly enhance decision-making, potentially leading to higher returns and reduced risks. Utilizing daily historical data from the S&P/NZX 50 and S&P/ASX 200 indices, the models are evaluated using metrics such as RMSE, R-squared, and MAPE. The results highlight the superior performance of hybrid models, particularly the GRU-LSTM, in capturing the complex temporal and spatial dependencies of stock price data. For the S&P NZX Daily dataset, the GRU-LSTM model achieves an RMSE of 39.92, an R^2 score of 0.984, and a MAPE of 0.226. Similarly, the S&P ASX Daily dataset achieves an RMSE of 11.75, an R^2 score of 0.995, and a MAPE of 0.121. These findings provide valuable insights for investors and traders, enhancing their ability to make informed investment decisions by leveraging the predictive power of advanced deep learning models. The study underscores the importance of hybrid approaches in financial forecasting, demonstrating their enhanced predictive accuracy and robustness in real-world applications.

Contents

CHAPTER 1

Introduction

In the constantly evolving global economy, financial activities have grown increasingly complex. This complexity makes understanding and predicting market trends essential for investors and traders, who need reliable tools to inform their investment strategies and trading decisions. The stock market, valued at 98.5 trillion USD in 2022 [1], incorporates this challenge due to its inherent volatility and non-stationary nature. Predicting stock index prices has consistently been one of the most difficult tasks for professionals in finance and related fields due to the market's inherent volatility and noise [2].

This project addresses these challenges by performing a comprehensive comparative analysis of various deep learning models for predicting stock market trends in the New Zealand and Australian markets. Time series prediction involves creating models to forecast future values based on past data. Since the connection between past and future values is often not straightforward, this process essentially models the conditional probability distribution using past observations [3]. Traditional statistical methods and econometric models such as the Autoregressive (AR) Model, Autoregressive Moving Average (ARMA) Model, and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model have been widely used to forecast time series data, showing considerable success in stock price prediction [4]. Still, their limitations in capturing nonlinear relationships in data have paved the way for more sophisticated machine learning and neural network techniques. Recent advancements in machine learning have led to the development of various models designed to improve prediction accuracy by leveraging nonlinear patterns [4]. Neural networks, especially deep learning models, have shown promise in dissecting these complex patterns, especially for time series forecasting [5].

Recently, the use of machine learning and deep learning approaches in stock market forecasting has significantly increased. A comprehensive review revealed that deep learning and hybrid models constitute 45% of the techniques used in this area [6]. Although these models show great promise in capturing the nonlinear dynamics of financial data, there are still notable gaps in the research. This project will explore several deep learning models, including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Networks (CNN), and hybrid models, including CNN-LSTM and GRU-LSTM. By leveraging daily historical data from the S&P/NZX 50 and S&P/ASX 200 indices, the project aims to evaluate the performance of these models using metrics such as Root Mean Squared Error (RMSE), R-squared (R^2) and Mean Absolute Percentage Error (MAPE). Ultimately, the research aims to identify the most effective model for stock price prediction, providing valuable insights for investors and traders in making informed decisions.

The structure of the report starts with chapter 1 providing an overview of the project's motivation and objectives. In chapter 2, a review of existing literature on traditional machine learning approaches, deep learning models, and their applications in financial forecasting is presented. Chapter 3 describes the datasets used, including the S&P/NZX 50 and S&P/ASX 200 indices, and details the preprocessing steps to prepare the data for modelling. The methodology is discussed in chapter 4, which covers the deep learning models employed, including LSTM, GRU, CNN, CNN-LSTM, and GRU-LSTM, and explains the evaluation metrics used to assess model performance. Chapter 5 presents the findings from the comparative analysis of the models, including a detailed discussion of the performance metrics and insights gained from the evaluation. Chapter 6 summarizes the key findings, contributions of the study, and implications for financial forecasting and suggests directions for future research. At the end of chapter 6, the personal learning and reflection section reflects on the personal learning experience gained from the project and the development of technical and research skills. This structured approach outlines a comprehensive examination of deep learning models in stock market prediction, highlighting their potential and guiding future research in this critical area.

CHAPTER 2

Literature review

Over the years, various solutions have been proposed to tackle the challenges of stock market prediction, with notable methods including machine learning, deep learning, time series forecasting, and ensemble algorithms [7]. These methods have proven effective in improving prediction accuracy and reducing root mean square error (RMSE). Sonkavde et al. [8] conducted a systematic review and analysis of these approaches based on their performance, and the most effective solutions for stock market forecasting combine fundamental analysis, technical analysis, sentiment analysis, and deep learning models. Ensemble techniques, in particular, offer highly promising forecasting results. Fig 2.1 represents the algorithms employed for forecasting the stock market [8].

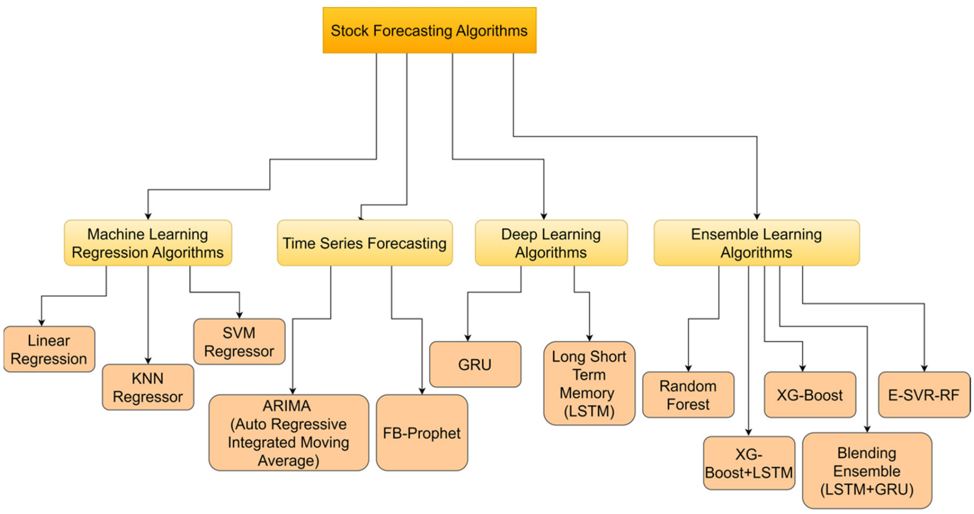


Figure 2.1: Algorithms of Stock Market Prediction [8]

2.1 Traditional machine learning approaches

Traditional machine learning methods, specifically focusing on tree-based models like Boosting, Random Forest, and Bagging, have proven effective in handling classification and regression tasks [9]. These methods have been utilised because they can handle large datasets and provide reliable predictions. For example, traditional models such as ARIMA and GARCH have been extensively used for time series analysis due to their statistical properties and well-established Box-Jenkins methodology. However, they are limited by their linear nature and often struggle to capture the complex, nonlinear relationships inherent in financial markets [10]. The literature by [4] highlights the comparative analysis of traditional models such as ARIMA, GARCH, and hybrid deep learning models on datasets of SP500, Nikkei225 and CSI300. The attention-based model (UA) outperformed the other models (MLP, LSTM, CNN) in all metrics, indicating its superior ability to allocate importance to relevant variables in financial time series. The UA model had the smallest MAPE (0.0067) and the highest correlation coefficient (R = 0.9891). Zhang proposes a hybrid model of ARIMA and ANN where an ARIMA model captures the linear components, and the nonlinear residuals are modelled using an ANN [10]. The empirical results from the three datasets (Sunspot, Canadian Lynx, Exchange Rate) demonstrate that the hybrid ARIMA-ANN model substantially improves forecasting accuracy over traditional ARIMA and standalone ANN models.

Huang et al. [11] introduce the Local Support Vector Regression (LSVR) model to incorporate local volatility into the SVR model, and the study highlights the significant advantages of the Local Support Vector Regression (LSVR) model in financial time series prediction. As another example of SVR, Shen et al. [12]propose a novel prediction algorithm that leverages the temporal correlation among global stock markets and various financial products to predict next-day stock trends using SVM. This approach addresses the limitations of using isolated market data by incorporating external information from other financial markets, commodities, and currencies, thereby enhancing prediction accuracy. Khan et al. (2022) conducted stock market prediction based on the integration of social media and financial news data, combined with machine learning and feature selection techniques, achieved 80-83.22% accuracy using Pyspark, MLlib, linear regression, and random forest [13], with the highest accuracy of RF classifier increased to 83.22%.

2.2 Deep learning and ensemble methods

A significant trend towards utilising advanced deep learning architectures has been noted, with models such as Multilayer Perceptrons (MLP), Recurrent Neural Networks (RNN), LSTM, and CNN becoming increasingly common. Ghosh et al. [14] showed that CNNs substantially outperform traditional linear models like ARIMA and other neural networks, including RNN, LSTM, and MLP, in forecasting stock prices in various markets, such as the National Stock Exchange of India and the New York Stock Exchange. Specifically, the study highlighted that CNNs, due to their strong spatial feature extraction capabilities, achieved much lower Mean Absolute Percentage Error (MAPE) values—7.63 for RNN, 6.36 for LSTM, 5.30 for CNN, and 6.71 for MLP when predicting MARUTI stock prices. This suggests the significant potential of neural networks in capturing the nonlinear dynamics of stock market data. Their research methodology involved training neural networks using data from Tata Motors on the NSE and testing them on different companies from the NSE and NYSE. The analysis covered stocks from sectors like Automobile, Banking, and IT, thoroughly assessing the models' predictive abilities.

Sonkavde et al. (2023) [8] review the strengths and limitations of several ML and DL models (LSTM, GRU) and ensemble models (such as combinations of RF, XG-Boost, etc.) and the study uses TAINIWALCHM and AGROPHOS stock data. LSTM and GRU models are particularly effective at identifying complex patterns in time series data. The importance of selecting appropriate hyperparameters and model evaluation metrics like RMSE and MAE are highlighted. Analyzing the dataset for TAINIWALCHM stock yielded RMSE scores of 4.4249 for KNN, 5.6241 for LSTM, and an improved score of 2.0247 for the ensemble algorithm that combines Random Forest, XG-Boost, and LSTM. Bhandari's study [15] evaluates predictive accuracy using RMSE, MAPE, and the correlation coefficient by comparing single-layer and multilayer LSTM models. The results suggest that a single-layer LSTM model, particularly with around 150 neurons, offers superior fit and higher prediction accuracy than its multilayer counterparts, highlighting the efficiency of simpler model architectures in dealing with the nonlinear and volatile nature of stock market data.

Despite their advantages, deep learning models do face challenges. Although models like CNN and LSTM perform well, they often encounter overfitting issues due to the limited size of financial datasets [16]. Researchers have investigated models with fewer parameters and simpler architectures, such as GRU, to mitigate these issues. GRU, an enhanced version of RNN with fewer parameters than LSTM, enables faster training and reduces the risk of overfitting [17]. Conversely, LSTM networks manage extended dependencies better than GRU and RNN [17]. The literature [16] highlights critical problems like model overfitting, especially with limited financial datasets, which compromises performance on unseen data and raises concerns about the models' applicability across different market conditions and sectors. The authors propose a GRU-based model, StockNet, incorporating an innovative data augmentation approach within the network to enhance data variety and robustness. The model comprises two main modules: the Injection module, which prevents overfitting by augmenting data within the network, and the Investigation module, focusing on forecasting stock indexes. Stocknet is evaluated using historical data from the CNX-Nifty index of the Indian stock market, showing a reduction in RMSE by 76.76% and 69.74% compared to DNN and SimpleRNN, respectively.

In some instances, model selection is based on specific market challenges. In 2021, Touzani and Douzi [18] proposed an LSTM-GRU-based trading strategy to forecast stock closing prices and address the Moroccan market's issues, such as lower liquidity and trading discontinuity. They review the application of LSTM and GRU models across various fields, emphasizing their effectiveness in analyzing sequence-dependent data. The model includes two main components: prediction using LSTM and GRU models and decision-making rules based on these predictions. A simulation over multiple 3-year periods was conducted to validate its effectiveness, using returns to estimate expected outcomes. Islam et al. [19] explored the integration and efficacy of combining GRU and LSTM for forecasting currency exchange rates, a notable challenge due to the FOREX market's complexity and volatility. Their proposed model outperformed LSTM, GRU, and SMA methods over a 10-minute timeframe for currency pairs, achieving lower scores in MSE, RMSE, and MAE metrics.

Recent research has introduced hybrid models combining different neural network strengths to enhance prediction accuracy and generalization. For example, Jimmy Ming-Tai Wu et al. [20] propose a novel framework combining CNN and LSTM methods. This hybrid model, SACLSTM, utilizes CNNs' strengths in extracting spatial features from historical stock data and leading indicators like futures and options and LSTMs to capture temporal dependencies. Stock market data was uniquely represented as two-dimensional arrays during preprocessing, simulating image input for CNNs. Leveraging spatial and temporal data, the model improved generalization across stock markets.

Another example of the CNN and LSTM hybrid model is discussed by Song and Choi [21]. They proposed CNN-LSTM, GRU-CNN, and ensemble models, leveraging the spatial feature extraction abilities of CNNs with the sequential pattern recognition strengths of LSTM and GRU methods. These models were tested to forecast the closing prices of major stock indices such as DAX, DOW, and S&P 500. Generally, the hybrid models outperformed RNN, LSTM, and GRU, resulting in a 77.8% improvement in MSE and MAE. Including average high and low prices as a feature positively impacted the models' performance.

Furthermore, attention-based models have gained traction for their ability to prioritize relevant information, thereby enhancing prediction efficiency and accuracy. For example, the CNN-BiLSTM-Attention model by Zhang et al. [22] demonstrated superior performance in predicting stock prices compared to traditional and deep learning models. The integration of CNN and BiLSTM allowed the model to effectively capture spatial and temporal features, surpassing conventional models with the lowest RMSE (64.848) and highest R^2 (0.985) metrics, demonstrating the effectiveness of incorporating attention mechanisms in neural network models for stock price prediction.

The following table overviews the machine learning and deep learning literature review.

Author, Year	**Model**	**Dataset**	**Findings**
Zhang 2003 [10]	Hybrid ARIMA and ANN	Sunspot data, Canadian lynx data, British Pound/US Dollar exchange rate data	The hybrid model outperformed standalone ARIMA and ANN models in forecasting accuracy. The hybrid model significantly improved MSE and MAD across different datasets, effectively capturing linear and nonlinear patterns in time series data.
Huang et al., 2006 [11]	LSVR	DJIA, NASDAQ, S&P 500 (2004)	The LSVR model outperformed standard SVR by adapting the margin locally, resulting in lower MSE values and better prediction accuracy for volatile financial time series.

Shen et al. 2023 [12]	SVM, MART	Global stock indices, currencies, commodities (2000-2012)	The proposed SVM-based algorithm achieved high prediction accuracy (74.4% for NASDAQ, 76% for S&P 500, and 77.6% for DJIA). The trading model demonstrated higher profitability and lower risk compared to benchmarks, with an average profit of $814.60 over 50-day periods, corresponding to an 8% return rate.
Khan et al. 2020 [13]	Random Forest, Deep Learning, Ensemble Methods	Social media (Twitter), financial news (Business Insider), historical stock prices (Yahoo Finance)	Combining social media and news data improved prediction accuracy. The Random Forest classifier achieved the highest accuracy of 83.22%. New York and IBM stocks were more influenced by social media, while London and Microsoft stocks were more influenced by financial news.
Nabipour et al. 2020 [9]	Decision Trees, Bagging, Random Forest, RNN, LSTM	Tehran Stock Exchange data	Deep learning models (RNN and LSTM) outperformed tree-based models in predicting stock prices. LSTMs were particularly effective due to their ability to capture long-term dependencies, resulting in higher accuracy and lower error rates.
Gao et al. 2020 [4]	MLP, LSTM, CNN, UA	SP500, CSI300, Nikkei225 indices (2008-2016)	The UA model outperformed other models in predicting stock index prices, demonstrating the best performance with the lowest MAPE and highest correlation coefficient across all datasets. UA model's attention mechanism effectively allocated importance to relevant variables in financial time series.
Zhang et al. 2023 [22]	CNN-BiLSTM-Attention	S&P 500 stock data	CNN-BiLSTM-Attention model achieved the lowest RMSE (64.848) and highest R^2 (0.985) compared to traditional and other deep learning models. The Attention mechanism improved feature relevance, and the model demonstrated robustness against market volatility.

Ghosh et al. 2024 [14]	CNN, RNN, LSTM, MLP	NSE, NYSE	CNNs outperformed ARIMA and other neural networks, achieving lower MAPE values, demonstrating the lowest MAPE and superior performance in capturing nonlinear trends in stock prices across different sectors. CNN: 5.30, LSTM: 6.36, RNN: 7.63, MLP: 6.71 for MARUTI stock prices.
Sonkavde et al. 2023 [8]	LSTM, GRU, RF, XG-Boost	TAINIWALCHM, AGROPHOS(2018-2023)	LSTM and GRU effective in time series patterns. Ensemble model RMSE: KNN 4.4249, LSTM 5.6241, ensemble 2.0247.
Bhandari et al. 2022 [15]	Single-layer and Multi-layer LSTM	S&P 500 index, macroeconomic variables, technical indicators (2006-2021)	The single-layer LSTM with 150 neurons achieved the highest prediction accuracy, outperforming multi-layer LSTM models. The study highlights the importance of hyperparameter tuning in improving model performance.
Gupta et al. 2022 [16]	GRU (StockNet)	CNX-Nifty index	StockNet reduced RMSE by 76.76% and 69.74% compared to DNN and SimpleRNN, respectively.
Touzani and Douzi 2021 [18]	LSTM-GRU	Moroccan market	LSTM-GRU based strategy effective in forecasting stock closing prices, validated over multiple 3-year periods.
Islam et al. 2021 [19]	GRU, LSTM, SMA	FOREX, market	GRU-LSTM model, outperformed LSTM, GRU, and SMA, achieving lower MSE, RMSE, and MAE.
Wu et al. 2023 [20]	CNN-LSTM (SACLSTM)	Various stock markets	SACLSTM improved generalization across stock markets by leveraging spatial and temporal data.
Song and Choi 2023 [21]	CNN-LSTM, GRU-CNN, Ensemble	DAX, DOW, S&P 500	Hybrid models outperformed RNN, LSTM, GRU, showing a 77.8% improvement in MSE and MAE.

Table 2.1: literature on machine learning and deep learning methods for financial forecasting

Given these advancements, it is essential to conduct comparative studies, particularly in less researched markets such as New Zealand and Australia. These analyses will address current research gaps and provide deeper insights into how different deep learning models function under various market conditions. Consequently, this project seeks to enhance financial forecasting by assessing the effectiveness of several deep learning configurations in predicting stock market trends within these particular conditions.

CHAPTER 3

Data description and preparation

3.1 Data description

The dataset utilised in this study comprises daily stock data for the NZX 50 and ASX 200 indices. The NZX 50 is a stock market index representing the 50 biggest companies on the New Zealand Stock Exchange (NZX). The ASX 200 index lists the 200 top companies on the Australian Securities Exchange (ASX).

3.1.1 NZX 50 index

The NZX 50 dataset includes the daily stock prices of the top 50 companies by market capitalisation listed on the New Zealand Stock Exchange. The data spans from January 1, 2017, to December 31, 2023. The key attributes in this dataset are:

- **Date**: The trading date.
- **Open**: The stock's opening price on the given date.
- **High**: The stock's highest price on the given date.
- **Low**: The stock's lowest price on the given date.
- **Close**: The stock's closing price on the given date.
- **Volume**: The number of shares traded on the given date.
- **Adjusted Close**: The closing price adjusted for stock splits and dividends.

3.1.2 ASX 200 index

The ASX 200 dataset includes the daily stock prices of the top 200 companies by market capitalisation listed on the Australian Securities Exchange. The data spans from January 1, 2017, to December 31, 2023. The key attribute in this dataset are:

S&P/NZX 50 INDEX GROSS (GROSS (^NZ50)

☆ Follow

11,856.56 -116.45 (-0.97%)

At close: June 7 at 5:47 PM GMT+12

Jan 01, 2017 - Dec 31, 2023 | Historical Prices | Daily

Currency in NZD

Date	Open	High	Low	Close	Adj Close	Volume
Dec 29, 2023	11,768.68	11,777.37	11,745.10	11,770.49	11,770.49	5,856,100
Dec 28, 2023	11,768.68	11,777.37	11,745.10	11,761.26	11,761.26	1,841,100
Dec 27, 2023	11,678.43	11,701.87	11,646.18	11,687.97	11,687.97	537,000
Dec 22, 2023	11,627.99	11,634.43	11,598.18	11,634.43	11,634.43	8,811,600
Dec 21, 2023	11,627.99	11,627.99	11,600.25	11,601.26	11,601.26	2,360,300
Dec 20, 2023	11,579.80	11,579.80	11,544.85	11,571.03	11,571.03	2,712,100
Dec 19, 2023	11,617.37	11,617.37	11,574.01	11,604.29	11,604.29	3,697,900
Dec 18, 2023	11,564.98	11,564.98	11,492.62	11,508.56	11,508.56	2,401,800
Dec 15, 2023	11,552.88	11,579.17	11,481.21	11,550.20	11,550.20	89,953,600
Dec 14, 2023	11,552.88	11,579.17	11,518.32	11,548.40	11,548.40	3,703,200
Dec 13, 2023	11,475.77	11,510.46	11,475.77	11,496.82	11,496.82	56,265,600
Dec 12, 2023	11,382.58	11,435.81	11,382.58	11,435.81	11,435.81	3,140,900
Dec 11, 2023	11,449.47	11,449.47	11,402.16	11,409.97	11,409.97	4,052,300
Dec 8, 2023	11,496.61	11,496.61	11,408.25	11,495.64	11,495.64	19,096,000
Dec 7, 2023	11,496.61	11,496.61	11,444.75	11,445.47	11,445.47	2,408,200
Dec 6, 2023	11,463.49	11,474.29	11,435.90	11,437.26	11,437.26	1,256,000
Dec 5, 2023	11,356.99	11,357.87	11,331.07	11,343.85	11,343.85	8,946,000

Figure 3.1: NZX 50 Historical Data [23]

- **Date**: The trading date.
- **Open**: The stock's opening price on the given date.
- **High**: The stock's highest price on the given date.
- **Low**: The stock's lowest price on the given date.
- **Close**: The stock's closing price on the given date.
- **Volume**: The number of shares traded on the given date.
- **Adjusted Close**: The closing price adjusted for stock splits and dividends.

3.1.3 Data sources

The stock data for the NZX 50 and ASX 200 indices were sourced from Yahoo Finance [23], a reputable financial data provider, ensuring accuracy and reliability. The datasets include comprehensive historical records, allowing for a detailed analysis of stock market trends and the performance of individual stocks.

3.1.4 API data loading

The Yahoo Finance API was used to retrieve the stock data. The following code snippet demonstrates how the data was loaded:

```
# Load data
start = '2017-01-01'
end = '2023-12-31'
stock = '^NZ50'
data = yf.download(stock, start, end)
```

- **Start and end dates**: The data retrieval period was set from January 1, 2017, to December 31, 2023.
- **Stock ticker**: The stock ticker symbol ‘N̂Z50‘ was used to specify the NZX 50 index.
- **Data retrieval**: The ‘yf.download‘ function from the Yahoo Finance API was used to download the daily stock data for the specified period.

This method ensured that the dataset included up-to-date and accurate historical records necessary for the analysis.

3.2 Data preprocessing

Before comparative analysis employing LSTM, GRU, CNN and hybrid models, the stock data underwent several preprocessing steps to ensure its quality and suitability for modelling. The dataset was divided into training, validation, and testing subsets to facilitate the development and evaluation of the models. The features selected for the model include Open, High, Low, Close, Adjusted Close, and Volume, with the Open price designated as the target variable for prediction.

- **Missing values**: Any missing values in the dataset were identified and appropriately handled through imputation and removal.
- **Date formatting**: Dates were standardised to a common format to facilitate time series analysis.
- **Normalisation**: Stock prices were normalised to account for differences in scale and improve machine learning models' performance. MinMax scaling [24] was employed to adjust the feature values so they fall within the range of 0 to 1. This normalization technique aids in speeding up the convergence of the deep learning models.

$$x = \frac{x - x_{\max}}{x_{\max} - x_{\min}}, \tag{3.1}$$

- **Sequence Creation**: Following this normalisation, sequences of historical data were created to use as input for time series prediction of the target variable. The following code creates sequences of historical data, with a specified number of time steps (60 time steps for this project), to be used as input for the prediction models.

  ```
  def create_sequences(data, time_steps, target):
      data_X = []
      data_y = []
      for i in range(time_steps, len(data)):
          data_X.append(data[i-time_steps:i, 0:data.shape[1]])
          data_y.append(data[i, target])
      return np.array(data_X), np.array(data_y)

  # Define time steps
  time_steps = 60
  ```

This comprehensive dataset provides a robust foundation for evaluating the efficacy of various deep-learning models in predicting stock market trends within the NZX 50 and ASX 200 indices.

CHAPTER 4

Methodology

This chapter involves the theories and methods of deep learning approaches and evaluation metrics employed for predicting open prices of New Zealand and Australian stock indices. The comparative analysis of the project is conducted based on three sections, including CNN, RNN variants (LSTM and GRU) and their integrated methods. Figure 4.1 illustrates the framework of the project, including applied deep learning models.

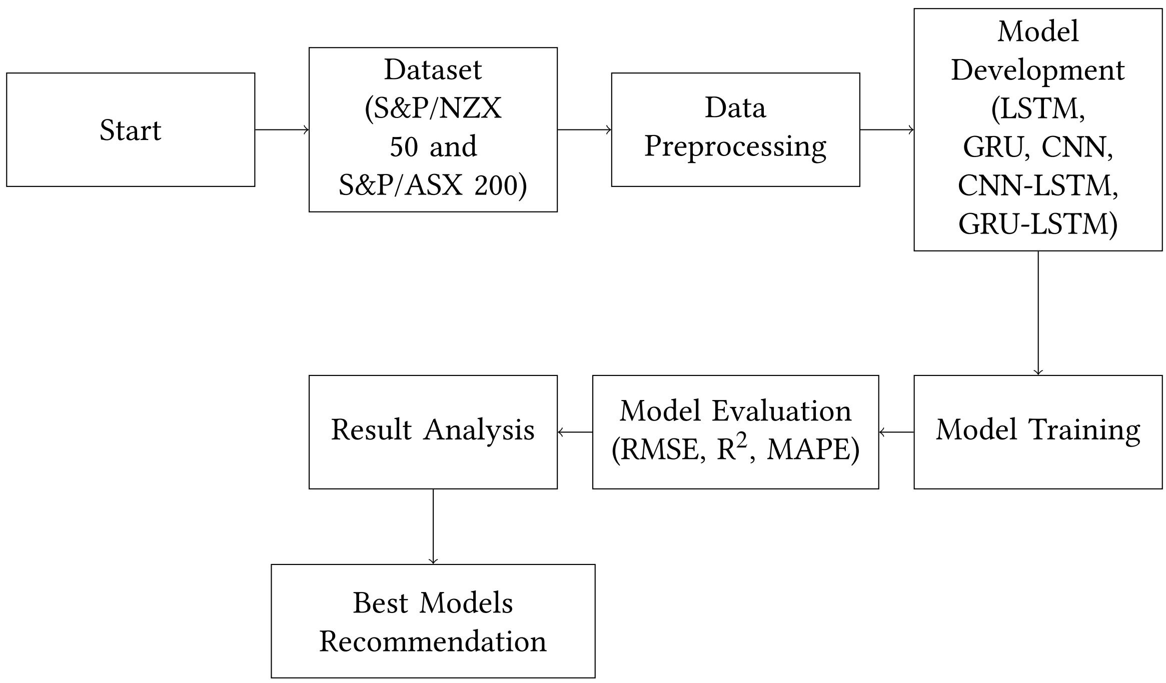


Figure 4.1: Project Framework

4.1 CNN

CNNs are a class of deep neural networks primarily used for processing grid-like data such as images. However, they also effectively extract features from sequential data like time-series data. The architecture of CNNs was initially introduced by LeCun et al. in 1995 [25]. As illustrated in Fig 4.2, key components of CNNs include:

- **Convolutional layers**: These layers apply convolution operations to the input, using filters to extract local patterns. The CNN model proposed by the authors [26] comprises the convolutional layers designed to extract high-level features from daily financial data. It uses filters that span across all the initial variables for each day, effectively combining them into a more informative representation. These layers further process the extracted features to capture patterns over multiple days, allowing the detection of short-term trends and patterns.
- **Pooling layers**: These layers perform down-sampling operations, reducing the dimensionality of the data while preserving important features [27]. These layers help mitigate the risk of overfitting, and max pooling is commonly used to retain the most significant features while reducing computational complexity [26].
- **Fully Connected layers**: These layers connect every neuron in one layer to every neuron in another layer, often used for the final prediction tasks. The high-level features are flattened into a one-dimensional vector, which is then passed through one or more fully connected layers to integrate the features and produce the final prediction [26].

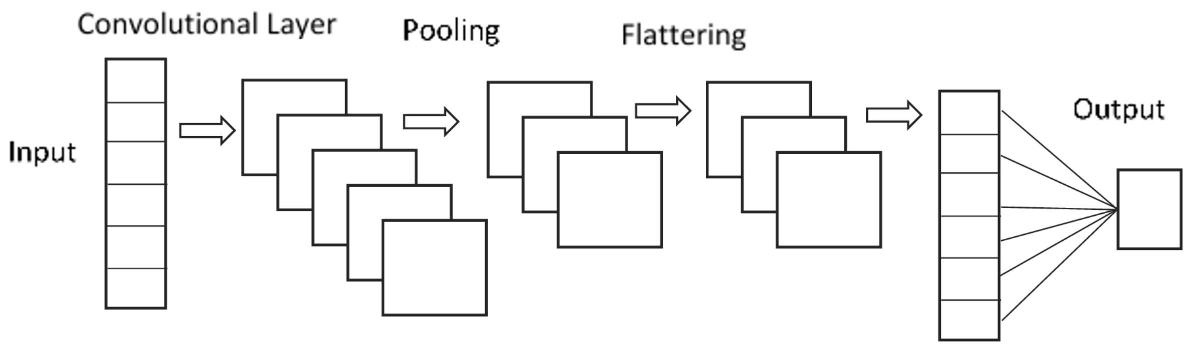


Figure 4.2: CNN Architecture
[27]

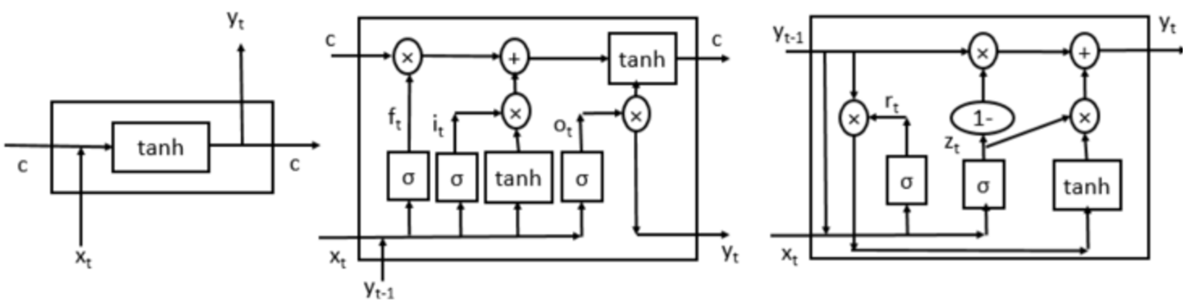
In the context of stock price prediction, 1D-CNNs are used to handle sequential data by capturing temporal patterns within the stock price time series. The dropout technique, initially developed for training deep neural networks, is also utilised. Dropout aims to prevent the model from overfitting to the training data. During each training cycle, each neuron has a probability of being excluded from training, known as the dropout rate. This approach reduces the model's flexibility, encouraging the learning algorithm to develop a model that generalises well to predict new, unseen data [26]. This study's CNN model architecture consists of two convolutional layers followed by max-pooling layers, a flatten layer, a dropout layer and two dense (fully connected) layers as shown in table 4.1.

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 59, 64)	832
max_pooling1d (MaxPooling1D)	(None, 29, 64)	0
conv1d_1 (Conv1D)	(None, 28, 128)	16,512
max_pooling1d_1 (MaxPooling1D)	(None, 14, 128)	0
flatten (Flatten)	(None, 1792)	0
dropout_7 (Dropout)	(None, 1792)	0
dense_3 (Dense)	(None, 50)	89,650
dense_4 (Dense)	(None, 1)	51

Table 4.1: CNN Model Summary

4.2 RNN and variants

RNNs are a type of artificial neural network designed to handle sequential data by maintaining a memory of previous inputs, making them suitable for tasks involving temporal dependencies. RNNs have a feedback loop that allows the output from a time step to become an input to the next time step. This creates a kind of "memory" in the network, enabling it to retain and use information from previous time steps when processing new data [28]. The RNN continually updates and uses the hidden state to process each new example in the sequence. However, standard RNNs suffer from issues like the vanishing gradient problem, which makes training difficult for long sequences [27]. LSTM and GRU are RNN variants introduced to resolve RNN problems. Fig 4.5 demonstrates the model architecture of RNN and variants.



(a) RNN (b) LSTM (c) GRU

Figure 4.3: Architecture of RNN and Variants
[27]

4.2.1 LSTM

LSTM networks are a variant of RNNs designed to overcome the limitations of traditional RNNs' vanishing and exploding gradient problems. They use a series of gates (input, forget, and output gates) to regulate the flow of information, allowing them to maintain long-term dependencies more effectively [27]. The cell state is a key component of LSTM networks that allows them to store information over long periods. It passes information along unchanged unless modified by the gates. The input gate decides which information should be added to the cell state. The new cell state is computed by combining the previous cell state, scaled by the forget gate, and the candidate cell state, scaled by the input gate [21]. This allows the network to retain relevant information over long sequences. The forget gate determines which information from the previous cell state should be discarded. The output gate determines the next hidden state based on the cell state. However, LSTMs can be computationally expensive due to their complex architecture and the possession of a large number of parameters.

Layer (type)	**Output Shape**	**Param #**
lstm (LSTM)	(None, 60, 100)	42,800
dropout (Dropout)	(None, 60, 100)	0
lstm_1 (LSTM)	(None, 60, 100)	80,400
dropout_1 (Dropout)	(None, 60, 100)	0
lstm_2 (LSTM)	(None, 60, 100)	80,400
dropout_2 (Dropout)	(None, 60, 100)	0
lstm_3 (LSTM)	(None, 100)	80,400
dropout_3 (Dropout)	(None, 100)	0
dense (Dense)	(None, 1)	101

Table 4.2: LSTM Model Summary

Table 4.2 demonstrates the structure of LSTM applied for this study. The LSTM model consists of four lstm layers, followed by dropout layers and a dense layer.

4.2.2 GRU

GRUs are a simpler variant of LSTMs that combine the input and forget gates into a single update gate and use a reset gate to control the memory. This makes GRUs computationally more efficient while addressing the vanishing gradient problem [27]. They are designed to handle sequence learning tasks effectively by addressing the vanishing and exploding gradient problems found in traditional RNNs. Unlike LSTMs, which maintain separate cell and hidden states, GRUs merge these into a single state. This simplification helps streamline the architecture while maintaining its ability to capture complex temporal dependencies [21]. However, LSTMs are generally better at remembering longer sequences, making them more suitable for tasks requiring long-term memory. The reset gate is a crucial component of GRUs. It determines the influence of the previous hidden state on the current state. Due to their capacities to efficiently handle long-term dependencies and being less prone to gradient-related issues, they become a powerful technique in the domain of sequence learning, especially when data is limited.

Layer (type)	Output Shape	Param #
gru (GRU)	(None, 60, 100)	32,400
dropout_4 (Dropout)	(None, 60, 100)	0
gru_1 (GRU)	(None, 60, 100)	60,600
dropout_5 (Dropout)	(None, 60, 100)	0
gru_2 (GRU)	(None, 100)	60,600
dropout_6 (Dropout)	(None, 100)	0
dense_1 (Dense)	(None, 50)	5,050
dense_2 (Dense)	(None, 1)	51

Table 4.3: GRU Model Summary

4.3 Hybrid models

This project will benchmark the hybrid models discussed in 2 to extract the strengths of these individual deep learning models and compare the performance with the single deep learning models.

4.3.1 CNN-LSTM

CNNs are highly effective in learning internal representations of time-series data. A one-dimensional convolutional layer helps filter out noise, extract spatial features, and reduce the number of parameters in the model [21]. On the other hand, RNNs are considered the best deep-learning models for forecasting time-series data due to their ability to handle sequential dependencies. The CNN-LSTM model combines the strengths of CNNs and LSTMs to forecast time-series data effectively. The benchmarked paper [29] proposes a CNN-LSTM hybrid model to forecast stock prices with the model architecture of the Input layer, one-dimensional convolutional layer, pooling layer, LSTM layer and fully connected dense layer as shown in Fig 4.4.

The CNN layer extracts features from the input data by applying convolution operations. The pooling layer reduces the size of the feature maps, retaining only the most significant features. The LSTM layer processes these features to learn temporal dependencies. Finally, The dense layer outputs the predicted stock price based on the learned features and dependencies. The CNN-LSTM model was tested using the daily trading data of the Shanghai Composite Index and compared with MLP, CNN, RNN, LSTM, and CNN-RNN. The proposed model achieved the highest prediction accuracy with the lowest MAE and RMSE. The project applies the same architecture and tests on daily stock prices of NZX and ASX indices. Table 4.4 explains the CNN-LSTM model structure implemented for this study. There are two convolutional layers with max-pooling layers, an LSTM layer, a dropout layer and two dense layers.

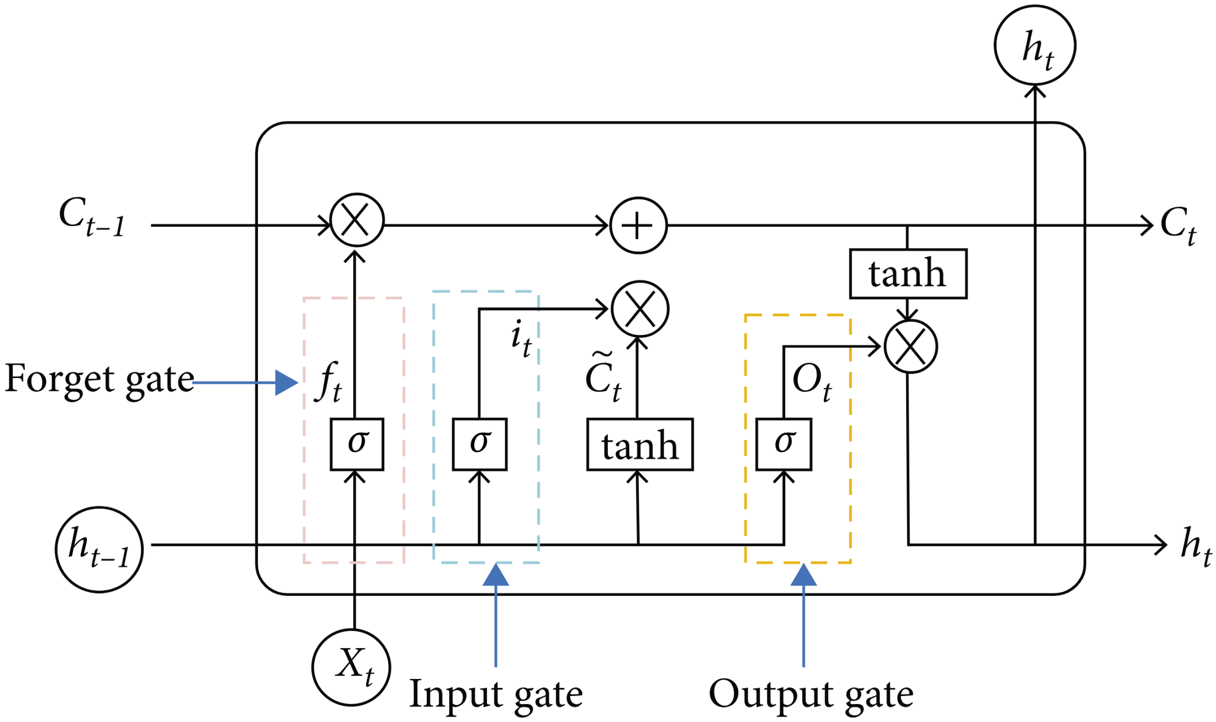


Figure 4.4: Architecture of CNN-LSTM
[29]

Layer (type)	Output Shape	Param #
conv1d_2 (Conv1D)	(None, 59, 128)	1,664
max_pooling1d_2 (MaxPooling1D)	(None, 29, 128)	0
conv1d_3 (Conv1D)	(None, 28, 128)	32,896
max_pooling1d_3 (MaxPooling1D)	(None, 14, 128)	0
lstm_4 (LSTM)	(None, 128)	131,584
dropout_8 (Dropout)	(None, 128)	0
dense_5 (Dense)	(None, 50)	6,450
dense_6 (Dense)	(None, 1)	51

Table 4.4: CNN-LSTM Model Summary

4.3.2 GRU-LSTM

Another hybrid model this project aims to employ is the GRU-LSTM model. The paper [19] proposes an integrated model of GRU-LSTM to predict foreign exchange currency prices. The model leverages the strengths of both GRU and LSTM to improve prediction accuracy for time series data. By combining GRU and LSTM, the model takes advantage of GRU's efficiency and LSTM's ability to learn long-term dependencies. Compared to standalone models, the hybrid model shows improved accuracy in predicting currency prices. GRU and LSTM make the model well-suited for handling the complexities of time series data, particularly in the volatile

FOREX market.

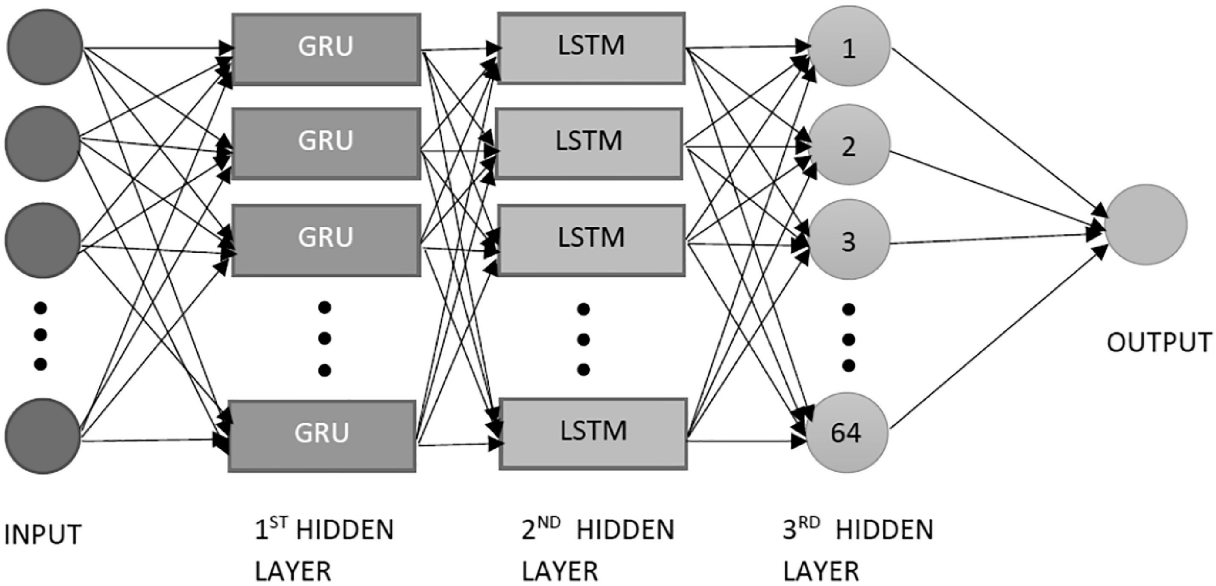


Figure 4.5: Architecture of GRU-LSTM
[19]

The hybrid model consists of input, GRU, LSTM, dense, and output layers in the model architecture. The model begins with an input layer that takes in the time series data of currency prices. As shown in table 4.5, the first hidden layer is a GRU layer with 20 hidden neurons. GRUs are used because they efficiently handle the vanishing gradient problem and are less computationally intensive than LSTMs. This layer processes the input data and generates weighted values passed to the next layer. The second hidden layer is an LSTM layer with 256 neurons, which were chosen for their ability to capture long-term dependencies in the data, which is crucial for accurate time series forecasting. Finally, dense layers are responsible for transforming the data into the output format, and the output is generated by the output neuron, which provides the predicted price.

Layer (type)	Output Shape	Param #
gru_3 (GRU)	(None, 60, 50)	8,700
lstm_5 (LSTM)	(None, 256)	314,368
dense_7 (Dense)	(None, 50)	12,850
dense_8 (Dense)	(None, 1)	51

Table 4.5: GRU-LSTM Model Summary

4.4 Evaluation metrics

In this project, three evaluation metrics are employed to assess the performance of the stock market prediction models: the R-squared (R^2) score, RMSE, and MAPE. These metrics comprehensively evaluate the models' accuracy and effectiveness in predicting stock prices.

R-squared (R^2) Score

The R-squared score, also known as the coefficient of determination, is a statistical measure representing the proportion of the variance for a dependent variable explained by an independent variable or variables in a regression model. The R^2 score ranges from 0 to 1, where 1 indicates perfect prediction, and 0 indicates that the model does not explain any variability in the response data [30]. A higher R^2 value indicates a better fit for the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where y_i are the actual values, $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the mean of the actual values.

RMSE

The RMSE is a frequently used measure of the differences between values predicted by a model and the values observed. It is the square root of the average squared differences between prediction and actual observation [30]. RMSE is a good measure of how accurately the model predicts the response, and it is most useful when large errors are particularly undesirable.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

MAPE

MAPE measures the accuracy of a model as a percentage, indicating the average magnitude of the errors in a set of predictions without considering their direction [30]. It is particularly useful for understanding the relative performance of the model.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

4.4.1 Importance of Evaluation Metrics

The choice of these metrics allows for a detailed evaluation of the model's performance. The R^2 score provides an overall indication of the model's explanatory power. RMSE gives insight into the magnitude of prediction errors. MAPE offers an understanding of the prediction accuracy in percentage terms, which is particularly intuitive for comparing different models and datasets. According to Chicco et al. (2021) [30], the R^2 score is often more informative than other metrics such as SMAPE, MAE, MAPE, and MSE due to its interpretability and ability to reflect the proportion of variance explained by the model.

These evaluation metrics are essential to ensure the models' accuracy and efficiency in predicting stock prices for evaluating various deep-learning configurations in stock market prediction.

CHAPTER 5

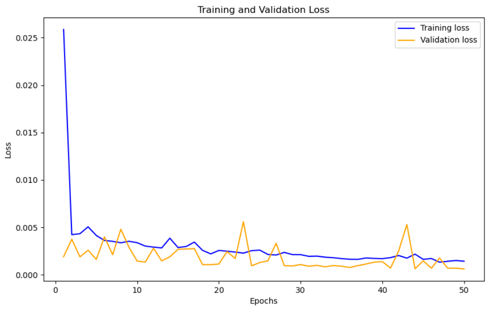
Result and discussion

This chapter explores the comprehensive analysis and performance of hybrid models CNN-LSTM, GRU-LSTM and their single models LSTM, GRU and CNN.

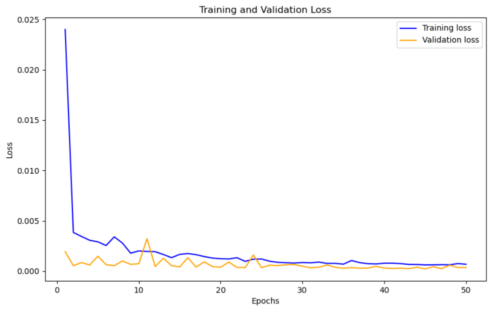
Loss curves experiment on NZX dataset

The training and validation loss curves for the five models (LSTM, GRU, CNN, CNN-LSTM, and GRU-LSTM) provide critical insights into the models' performance during training.

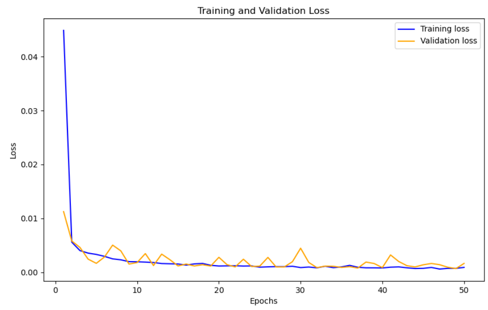
- **LSTM**: The LSTM model significantly reduces training and validation loss within the first few epochs, which then stabilizes. The validation loss closely follows the training loss, indicating that the model generalises well to unseen data. However, some fluctuations in the validation loss suggest that the model might still be experiencing minor overfitting.
- **GRU**: The GRU model initially exhibits a rapid decline in training and validation loss, followed by a consistent pattern where both losses remain low and close to each other. The small gap between the training and validation loss curves implies that the GRU model performs well without significant overfitting.
- **CNN**: The training loss decreases quickly and remains low for the CNN model, but the validation loss exhibits more noticeable fluctuations. The gap between the training and validation loss curves is more pronounced than LSTM and GRU models, indicating that the CNN model might be overfitting the training data. This suggests that CNN alone might struggle with the temporal dependencies of stock data.
- **CNN-LSTM**: The CNN-LSTM model shows a stable decline in both training and validation losses, with the validation loss closely tracking the training loss. This indicates good generalization and the ability to capture the data's spatial and temporal patterns. The combined architecture leverages the strengths of CNNs for spatial feature extraction and LSTMs for temporal sequence learning, resulting in a balanced performance.



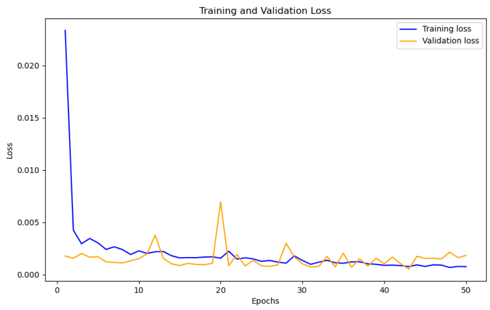
(a) LSTM Model



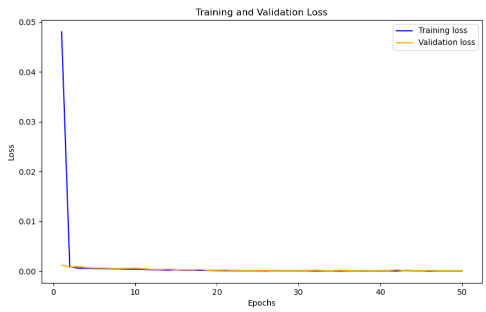
(b) GRU Model



(c) CNN Model



(d) CNN-LSTM Model



(e) GRU-LSTM Model

Figure 5.1: Training and Validation Loss Curves for Different Models on NZX Dataset

- **GRU-LSTM**: The GRU-LSTM model demonstrates the most stable and consistent loss curves among all models. Both training and validation losses decrease sharply and stabilize at very low values with minimal fluctuations. This indicates that the GRU-LSTM hybrid model effectively captures the complex patterns in the stock price data, providing excellent generalisation capabilities. The low and stable validation loss signifies that this model is less prone to overfitting than the others.

5.1 Evaluation metrics comparison

The evaluation metrics for the models applied to both the S&P NZX Daily and S&P ASX Daily datasets reveal significant insights into the performance of each model. Across both datasets, the GRU-LSTM model consistently outperforms other models regarding RMSE, R^2 score, and MAPE. Specifically, for the S&P NZX Daily dataset, the GRU-LSTM model achieves an RMSE of 39.92, an R^2 score of 0.984, and a MAPE of 0.226. Similarly, for the S&P ASX Daily dataset, the GRU-LSTM model shows remarkable performance with an RMSE of 11.75, an R^2 score of 0.995, and a MAPE of 0.121.

Model	Dataset	RMSE	R^2 Score	MAPE
LSTM	S&P NZX	108.10	0.885	0.737
	S&P ASX	65.78	0.848	0.709
GRU	S&P NZX	59.73	0.965	0.383
	S&P ASX	26.67	0.975	0.291
CNN	S&P NZX	151.27	0.775	1.076
	S&P ASX	83.42	0.756	0.937
CNN-LSTM	S&P NZX	103.42	0.895	0.677
	S&P ASX	64.91	0.852	0.705
GRU-LSTM	S&P NZX	**39.92**	**0.984**	**0.226**
	S&P ASX	**11.75**	**0.995**	**0.121**

Table 5.1: Evaluation Metrics Comparison for S&P NZX and S&P ASX Daily Datasets

In contrast, the CNN model exhibits poor performance across both datasets, with the highest RMSE and MAPE values and the lowest R^2 scores. This suggests that CNNs are powerful in extracting spatial features, but they might struggle with the temporal dependencies inherent in stock market data when used in isolation.

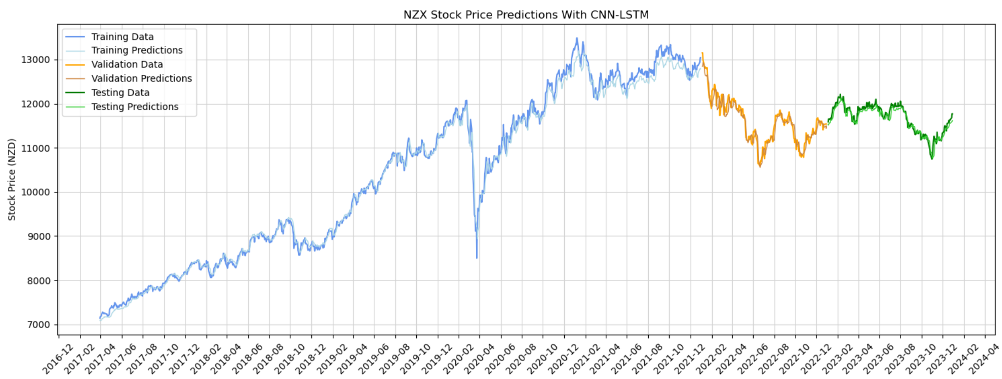
The LSTM and GRU models also perform well, with GRU generally outperforming LSTM. For the S&P NZX Daily dataset, the GRU model achieves an RMSE of 59.73, an R^2 score of 0.965, and a MAPE of 0.383, while for the S&P ASX Daily dataset, it achieves an RMSE of 26.67, an R^2 score of 0.975, and a MAPE of 0.291. These results indicate that GRU models may be more efficient in capturing the temporal dynamics of the stock prices than LSTM.

The hybrid models, CNN-LSTM and GRU-LSTM, demonstrate a balanced approach by leveraging the strengths of both CNNs and RNNs. The CNN-LSTM model shows competitive performance significantly better than the CNN model alone, suggesting that combining CNNs with LSTMs enhances the model's ability to capture both spatial and temporal features.

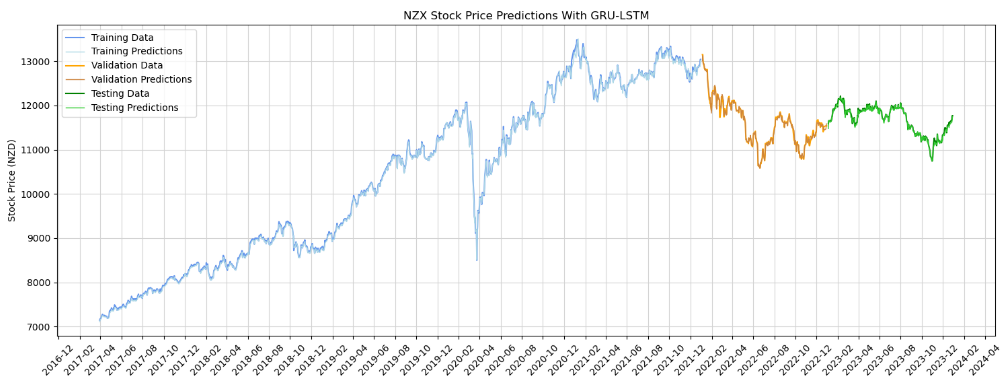
Overall, the GRU-LSTM model's superior performance highlights the importance of hybrid approaches in complex time series prediction tasks such as stock market forecasting. These findings suggest that hybrid models can effectively leverage the strengths of different architectures, leading to improved predictive accuracy and robustness.

5.2 Hybrid models comparative analysis

The two figures 5.2 illustrate the performance of CNN-LSTM and GRU-LSTM models in predicting NZX stock prices, with distinctions between training, validation, and testing phases. The light blue line represents the training data, while the blue line shows the training predictions. The model closely follows the training data, indicating a good fit during the training phase. The orange line represents the validation data, and the corresponding predictions are shown in light orange. A good alignment between the validation data and predictions demonstrates that the model generalises well to unseen data. The green line represents the testing data, while the light green line shows the testing predictions. The model maintains its ability to follow the actual data trends, although some deviations are noticeable, particularly during volatile periods. Overall, the CNN-LSTM model performs strongly across all phases, with accurate predictions during training, validation, and testing. However, slight deviations in the validation and testing phases suggest the model could benefit from further optimization to improve generalisation.



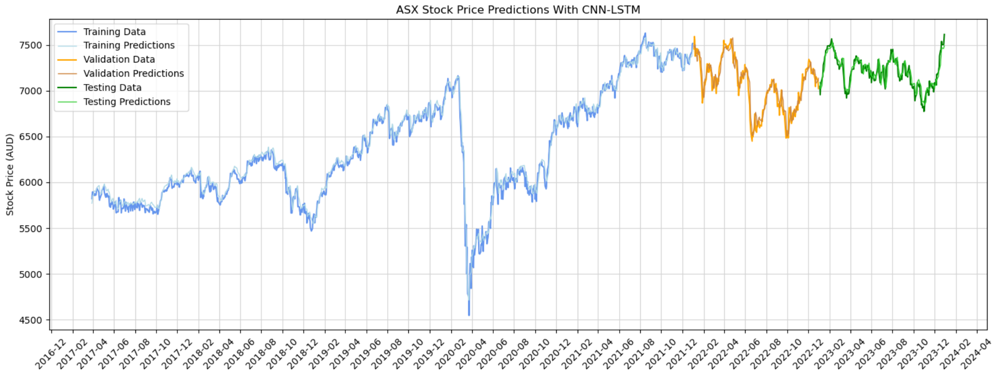
(a) CNN-LSTM



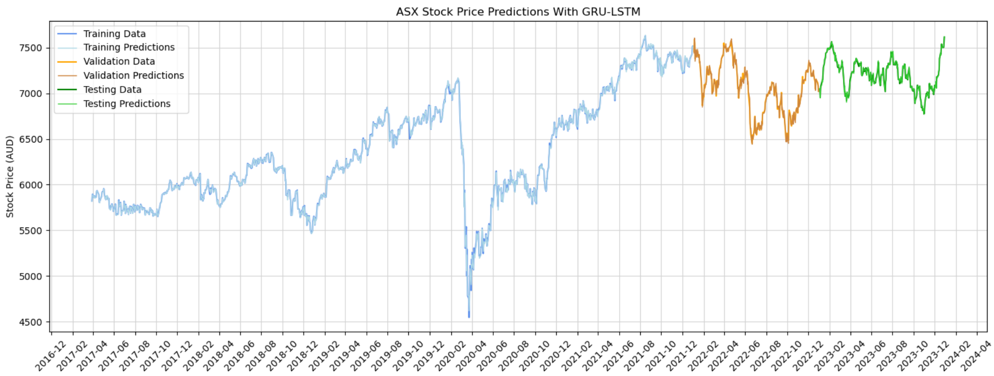
(b) GRU-LSTM

Figure 5.2: NZX Daily Open Price Predictions with Hybrid Models

Like the CNN-LSTM model, the GRU-LSTM model strongly fits the training data, with the light blue line (training data) and the blue line (training predictions) closely aligned.



(a) CNN-LSTM



(b) GRU-LSTM

Figure 5.3: ASX Daily Open Price Predictions with Hybrid Models

The orange line (validation data) and light orange line (validation predictions) indicate a close match. The model maintains accurate predictions, following the validation data with minimal deviation. The testing data and predictions show that the GRU-LSTM model performs exceptionally well on the testing data, with minimal deviations between actual and predicted values. The GRU-LSTM model demonstrates excellent performance across all phases, with predictions closely tracking the actual data throughout the training, validation, and testing phases. This indicates that the GRU-LSTM model effectively captures both short-term and long-term dependencies in the stock price data, making it a robust choice for stock price prediction tasks.

Similar to NZX dataset, the hybrid models show similar trends for ASX dataset 5.3. Both models show strong predictive capabilities for the ASX dataset, with the GRU-LSTM model slightly outperforming the CNN-LSTM model in terms of tracking the actual stock prices more closely across all phases. The GRU-LSTM model's ability to minimize deviations during validation and testing phases suggests better generalization and robustness, making it the preferred model for this task. However, the CNN-LSTM model also performs well and could be further optimised for improved performance.

CHAPTER 6

Conclusion

6.1 Summary of findings

The study demonstrates that hybrid models, specifically CNN-LSTM and GRU-LSTM, outperform traditional and standalone neural network models in predicting stock prices for the S&P NZX and S&P ASX datasets. The GRU-LSTM model consistently achieves the lowest RMSE and MAPE values and the highest R-squared scores, indicating its superior ability to capture complex and nonlinear patterns in the data. In contrast, the CNN model exhibits the poorest performance, suggesting that while CNNs effectively extract spatial features, they struggle with temporal dependencies. The LSTM and GRU models also perform well, with GRU generally outperforming LSTM. Overall, the findings emphasize the importance of hybrid approaches in financial forecasting, demonstrating their enhanced predictive accuracy and robustness.

Additionally, this study addresses a significant gap in the research by focusing on the New Zealand and Australian stock markets, which are relatively underexplored compared to other major markets like the US and Europe. These markets' unique characteristics and dynamics provide valuable insights and highlight the potential for applying advanced deep-learning models in diverse financial environments. Based on these results, the GRU-LSTM model is recommended as the best model for stock price prediction due to its remarkable performance in capturing temporal dependencies effectively in the data, leading to more accurate and reliable forecasts.

6.2 Contributions of the study

This study makes important contributions to the field of stock market prediction and the application of deep learning models in financial forecasting:

Hybrid model implementation and evaluation: The study successfully benchmarks and demonstrates the implementation and efficacy of hybrid models, specifically CNN-LSTM and GRU-LSTM, for predicting stock prices. These models combine the strengths of CNNs in feature extraction and RNNs in capturing temporal dependencies, leading to improved prediction accuracy. A comprehensive comparison of traditional machine learning models, standalone deep learning models, and hybrid models is conducted. The results highlight the superior performance of hybrid models, particularly the GRU-LSTM model, in terms of RMSE, R^2 score, and MAPE.

Generalisation and robustness: The study provides evidence of the generalization capabilities and robustness of the GRU-LSTM model across different datasets (S&P NZX and S&P ASX), making it a reliable tool for real-world stock price prediction.

Practical implications: The findings have practical implications for financial analysts and investors, offering a powerful tool for improving investment decision-making processes. The models developed can be integrated into automated trading systems and risk management strategies. This research contributes to the growing body of literature on applying deep learning in finance, offering insights into model optimisation and the potential of hybrid architectures.

6.3 Implications

The analysis and results of the CNN-LSTM and GRU-LSTM models for predicting stock prices in the NZX and ASX datasets have several significant implications. The superior performance of hybrid models, particularly the GRU-LSTM, demonstrates the effectiveness of combining different neural network architectures to capture spatial and temporal dependencies in financial time series data. The lower RMSE, higher R^2 scores, and reduced MAPE values indicate that these models can provide more accurate predictions than traditional and standalone neural network models. The stability and low deviation of the GRU-LSTM model in both the validation and testing phases suggest a strong generalization capability. This indicates that the model is less prone to overfitting and can adapt well to unseen data, which is crucial for real-world applications where market conditions vary.

Accurate stock price predictions can significantly enhance decision-making for investors and financial analysts. By leveraging the predictive power of these advanced models, stakeholders can make more informed investment decisions, potentially leading to higher returns and reduced risks. These models' successful implementation and evaluation contribute to the growing literature on machine learning applications in finance. It highlights the potential of deep learning models to outperform traditional forecasting methods, paving the way for further research and development in this field.

6.4 Future work

Future developments could be considered to improve the accuracy and applicability of stock price prediction models. Future work objectives involve creating more comprehensive, accurate, and practical applications for stock price prediction, contributing to more efficient and informed financial markets for underexposed areas such as New Zealand and Australia.

1. **Model optimisation**: Further tuning of hyperparameters and architectural enhancements can be explored to improve the performance of the CNN-LSTM and GRU-LSTM

models. Techniques such as Bayesian optimization or genetic algorithms [31] could be employed.

2. **Enhanced feature engineering**: The project employs technical analysis to generate price predictions. Integrating more complex and diverse features, such as macroeconomic indicators, sentiment analysis from news and social media, and market volatility indices, could provide additional context and improve the predictive accuracy of the models.

3. **Cross-market analysis**: Expanding the analysis to include multiple markets and indices could provide insights into the generalisation of the models across different financial environments. Comparative studies across various geographical regions and market conditions can validate the models' reliability.

4. **Trading and risk management applications**: Developing real-time prediction systems using these models can provide continuous and updated forecasts, which are crucial for high-frequency trading and dynamic investment strategies. Implementing these predictive models within automated trading systems can provide algorithmic trading strategies. Beyond prediction, these models can be adapted for risk management applications, such as predicting market downturns or identifying potential periods of high volatility, aiding in developing robust risk mitigation strategies.

6.5 Personal learning and reflection

Working on this project has been a tremendous learning experience, significantly boosting my technical skills in machine learning and deep learning. I have gained practical experience with hybrid models like CNN-LSTM and GRU-LSTM, which has been incredibly rewarding. My research and analytical abilities have improved throughout this journey, allowing me to critically evaluate existing literature, design experiments, and thoroughly analyse results.

The process involved overcoming challenges such as data preprocessing, model selection, and hyperparameter tuning, which required creative problem-solving and adaptability. These troubleshooting experiences taught me to approach problems systematically and persistently. Managing the project timeline and scope has also helped my project management skills, enhancing my planning, time management, and ability to meet deadlines.

Engaging with peers and supervisors improved my collaboration and communication skills. Presenting my findings and receiving feedback was crucial for refining my research. On a personal level, this project has reinforced the importance of resilience and continuous learning for excellence in both academic and professional pursuits.

Reflecting on this journey, I see my significant progress in understanding and applying advanced machine learning and neural network techniques. The challenges I faced and overcame have deepened my knowledge and strengthened my confidence in tackling complex problems. This project has prepared me to address future academic research and professional practice challenges with a renewed commitment to continuous improvement and lifelong learning. The insights and skills I have developed will become a strong foundation for my future data science and finance career.

Bibliography

[1] "Statista - The Statistics Portal — statista.com." `https://www.statista.com/`. [Accessed 05-06-2024]. → [p1]

[2] M. Binkowski, G. Marti, and P. Donnat, "Autoregressive convolutional neural networks for asynchronous time series," in *International Conference on Machine Learning*, pp. 580–589, PMLR, 2018. → [p1]

[3] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative finance*, vol. 1, no. 2, p. 223, 2001. → [p1]

[4] P. Gao, R. Zhang, and X. Yang, "The application of stock index price prediction with neural network," *Mathematical and Computational Applications*, vol. 25, no. 3, p. 53, 2020. → [p1], [p4], [p7]

[5] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," *Neural Computing and Applications*, vol. 32, no. 6, pp. 1609–1628, 2020. → [p1]

[6] 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, p. 1956, 2023. → [p1]

[7] M. Obthong, N. Tantisantiwong, W. Jeamwatthanachai, and G. Wills, "A survey on machine learning for stock price prediction: Algorithms and techniques," 2020. → [p3]

[8] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy, and S. K. Bhat, "Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications," *International Journal of Financial Studies*, vol. 11, no. 3, p. 94, 2023. → [p3], [p5], [p8], [p34]

[9] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and S. S, "Deep learning for stock market prediction," *Entropy*, vol. 22, no. 8, p. 840, 2020. → [p4], [p7]

[10] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003. → [p4], [p6]

[11] K. Huang, H. Yang, I. King, and M. R. Lyu, "Local support vector regression for financial time series prediction," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 1622–1627, IEEE, 2006. → [p4], [p6]

[12] S. Shen, H. Jiang, and T. Zhang, "Stock market forecasting using machine learning algorithms," *Department of Electrical Engineering, Stanford University, Stanford, CA*, pp. 1–5, 2012. → [p4], [p7]

[13] W. Khan, M. A. Ghazanfar, M. A. Azam, A. Karami, K. H. Alyoubi, and A. S. Alfakeeh, "Stock market prediction using machine learning classifiers and social media, news," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–24, 2022. → [p4], [p7]

[14] B. P. Ghosh, M. S. Bhuiyan, D. Das, T. N. Nguyen, R. M. Jewel, M. T. Mia, D. M. Cao, and R. Shahid, "Deep learning in stock market forecasting: Comparative analysis of neural network architectures across nse and nyse," *Journal of Computer Science and Technology Studies*, vol. 6, no. 1, pp. 68–75, 2024. → [p4], [p8]

[15] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. Khatri, "Predicting stock market index using lstm," *Machine Learning with Applications*, vol. 9, p. 100320, 2022. → [p5], [p8]

[16] U. Gupta, V. Bhattacharjee, and P. S. Bishnu, "Stocknet—gru based stock index prediction," *Expert Systems with Applications*, vol. 207, p. 117986, 2022. → [p5], [p8]

[17] T. B. Shahi, A. Shrestha, A. Neupane, and W. Guo, "Stock price forecasting with deep learning: A comparative study," *Mathematics*, vol. 8, no. 9, p. 1441, 2020. → [p5]

[18] Y. Touzani and K. Douzi, "An lstm and gru based trading strategy adapted to the moroccan market," *Journal of big Data*, vol. 8, no. 1, p. 126, 2021. → [p5], [p8]

[19] M. S. Islam and E. Hossain, "Foreign exchange currency rate prediction using a gru-lstm hybrid network," *Soft Computing Letters*, vol. 3, p. 100009, 2021. → [p5], [p8], [p18], [p19]

[20] J. M.-T. Wu, Z. Li, N. Herencsar, B. Vo, and J. C.-W. Lin, "A graph-based cnn-lstm stock price prediction algorithm with leading indicators," *Multimedia Systems*, vol. 29, no. 3, pp. 1751–1770, 2023. → [p6], [p8]

[21] H. Song and H. Choi, "Forecasting stock market indices using the recurrent neural network based hybrid models: Cnn-lstm, gru-cnn, and ensemble models," *Applied Sciences*, vol. 13, no. 7, p. 4644, 2023. → [p6], [p8], [p16], [p17]

[22] J. Zhang, L. Ye, and Y. Lai, "Stock price prediction using cnn-bilstm-attention model," *Mathematics*, vol. 11, no. 9, p. 1985, 2023. → [p6], [p7]

[23] "Yahoo Finance - Stock Market Live, Quotes, Business & Finance News — finance.yahoo.com." `https://finance.yahoo.com/`. [Accessed 25-05-2024]. → [p10], [p34]

[24] S. Nayak, B. B. Misra, and H. S. Behera, "Impact of data normalization on stock index forecasting," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 6, pp. 13–13, 2014. → [p11]

[25] Y. LeCun, Y. Bengio, *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995. → [p14]

[26] E. Hoseinzade and S. Haratizadeh, "Cnnpred: Cnn-based stock market prediction using a diverse set of variables," *Expert Systems with Applications*, vol. 129, pp. 273–285, 2019. → [p14]

[27] R. Jaiswal and B. Singh, "A hybrid convolutional recurrent (cnn-gru) model for stock price prediction," in *2022 IEEE 11th international conference on communication systems and network technologies (CSNT)*, pp. 299–304, IEEE, 2022. → [p14], [p15], [p16]

[28] M. Hiransha, E. A. Gopalakrishnan, V. K. Menon, and K. Soman, "Nse stock market prediction using deep-learning models," *Procedia computer science*, vol. 132, pp. 1351–1362, 2018. → [p15]

[29] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A cnn-lstm-based model to forecast stock prices," *Complexity*, vol. 2020, pp. 1–10, 2020. → [p17], [p18]

[30] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation," *Peerj computer science*, vol. 7, p. e623, 2021. → [p20]

[31] L. Haldurai, T. Madhubala, and R. Rajalakshmi, "A study on genetic algorithm and its applications," *International Journal of computer sciences and Engineering*, vol. 4, no. 10, p. 139, 2016. → [p29]

List of Figures

List of Tables

Abbreviations

ANN	Artificial Neural Network
AR	Autoregressive Model
ARIMA	Autoregressive Integrated Moving Average
ASX	Australian Securities Exchange
CNN	Convolutional Neural Network
GRU	Gated Recurrent Unit
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
LSTM	Long Short-Term Memory
MAPE	Mean Absolute Percentage Error
MLP	Multilayer Perceptron
NSE	National Stock Exchange
NZX	New Zealand Stock Exchange
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
S&P	Standard & Poor's
SVM	Support Vector Machine
UA	Universal Attention