A Hybrid CNN-LSTM-AE Framework with Large Language Model-Driven Sentiment Analysis for Cryptocurrency Anomaly Detection and Automated Trading

The development of a high-performance anomaly detection trading platform that facilitates historical data within the Cryptocurrency market empowered by sentiment analysis.

A thesis presented for the degree of
Bachelor's Science

Computer Science with Artificial Intelligence
University of Nottingham
psymm11@nottingham.ac.uk
Student ID: 20425448
COMP3003
2024-2025

A Hybrid CNN-LSTM-AE Framework with Large Language Model-Driven Sentiment Analysis for Cryptocurrency Anomaly Detection and Automated Trading

Submitted April 2025, in partial fulfilment of

the conditions for the award of the degree **Computer Science with Artificial Intelligence.**

20425448

School of Computer Science

University of Nottingham

I hereby declare that this dissertation is all my own work, except as indicated in the text:

Signature: MM

Date: 17/04/2025

I hereby declare that I have all necessary rights and consents to publicly distribute this dissertation via the University of Nottingham's e-dissertation archive.*

Public access to this dissertation is restricted until: DD/MM/YYY**

*Only include this sentence if you do have all necessary rights and consents. For example, if you have including photographs or images from the web or from other papers or documents then you need to obtain explicit consent from the original copyright owner. If in doubt, delete this sentence. See Copyright Information for more details.

**Only include this sentence if there is some reason why your dissertation should not be accessible for some period of time, for example if it contains information which is commercially sensitive or might compromise an Intellectual Property claim. If included, fill in the date from which access should be allowed.

A Hybrid CNN-LSTM-AE Framework with Large Language Model-Driven Sentiment Analysis for Cryptocurrency Anomaly Detection and Automated Trading

April 16, 2025

Abstract

This dissertation presents the development of a full-stack cryptocurrency trading system that integrates a novel hybrid deep learning model for anomaly detection with a real-time web interface. The anomaly detection model combines **Convolutional Neural Networks (CNN)**, **Long Short-Term Memory (LSTM) networks**, and **Autoencoders (AE)** to identify abnormal market behaviour, supporting more informed trading decisions. In addition to price-based analysis, the system incorporates large language model-powered sentiment analysis from Reddit and financial news, enhancing predictive performance by integrating both technical and social signals.

The back end of the system is implemented in Python using Flask, while the front end is built with React and Tailwind CSS to provide an interactive dashboard for real-time visualisation, strategy configuration, and portfolio tracking. Market and news data are sourced via the Alpaca API, with Reddit data gathered using PRAW. The entire system is modular, extensible, and designed to support live inference and automated trading through customisable strategies.

The performance of the hybrid model was evaluated using directional accuracy, cumulative return and AUC-ROC. On the test set (SOL), the model achieved a **directional accuracy of 92%**, **+819.4% cumulative return** and **AUC-ROC of 0.99**. Validation on unseen data (SOL) showed generalisation capability with a directional accuracy of 69.1%. The inclusion of sentiment analysis improved anomaly detection precision by approximately **11.3%** over price-only models.

Key challenges addressed include managing real-time latency, avoiding overfitting in time-series models, and ensuring responsible use of sentiment data from social media.

This work contributes a practical and extensible prototype for intelligent algorithmic trading, demonstrating the utility of combining deep learning and sentiment analysis in live systems. Future enhancements include expanding data sources, incorporating transfer learning, and improving the explainability of model decisions.

Acknowledgements

I would like to express my deepest gratitude to my family, friends, and supervisor, whose unwavering support has been invaluable throughout the course of this project. First and foremost, I am extremely grateful to my parents for their constant encouragement, love, and belief in me. Their support has been a source of motivation and strength, allowing me to stay focused and persevere through the challenges of this dissertation. To my sister, I owe a special thanks for always being there to lend an ear and offer advice, whether it was a word of encouragement or a moment of light-hearted distraction during stressful times.

I would also like to extend my heartfelt thanks to my friends, who have provided me with both emotional support and practical help throughout the duration of my studies. Their understanding, patience, and camaraderie have made this journey both enjoyable and meaningful.

Finally, I would like to express my sincere appreciation to my supervisor for their invaluable guidance and expertise. Their constructive feedback and thoughtful insights have played a crucial role in shaping the direction of this project and ensuring its successful completion. I am truly grateful for their patience and support.

Thank you all for your contributions to this achievement.

Contents

1 Introduction

1.1 The rise of Cryptocurrency

The cryptocurrency market has undergone a paradigm shift since Bitcoin's inception in 2009, evolving from a niche cryptographic experiment into a globally interconnected asset class with profound implications for financial systems. (Nakamoto, 2008) [55] As of 2024, the market comprises over 25,000 digital assets with a total capitalisation exceeding $1 trillion, reflecting both its rapid growth and structural uniqueness. (CoinMarketCap, 2024) [21] Unlike traditional markets, cryptocurrencies operate on decentralised exchanges with 24/7 trading, minimal regulatory oversight, and non-linear price dynamics driven by retail speculation, algorithmic trading, and macroeconomic factors. (Baur, 2018) [10] These conditions necessitate automated trading systems capable of processing high-frequency data and mitigating risks imperceptible to human traders. (Kaur, 2024) [38]

Machine learning (ML) has emerged as a critical enabler of such systems, addressing the limitations of rule-based strategies by adapting to chaotic market behaviour. (Baker, 2013) [9] Modern ML models leverage order book data, historical price trends, and sentiment indicators to identify latent patterns, with studies demonstrating superhuman performance in high-frequency environments. (SecuX, 2024) [69] This project builds on established methodologies to explore how ML-augmented sentiment analysis can enhance trading strategies in crypto markets. (Hull, 2009) [35]

1.2 The Influence of Social Media and Public Sentiment

Social media platforms such as Twitter, Reddit and TikTok have become primary drivers of cryptocurrency volatility, particularly in markets dominated by retail investors. (Qureshi, 2023) [64] Empirical evidence confirms their predictive power, research has found that Twitter sentiment explained 64% of Bitcoin's intra-day price variance, with a 1% increase in positive sentiment correlating to a 0.23% price rise within 24 hours. (Mai, 2018) [46] Temporal analysis further reveals that social media activity precedes price movements; research has observed that Twitter sentiment peaks 45 minutes before major price shifts, while polarisation signals institutional accumulation phases (Garcia, 2020) [31].

However, extracting actionable insights from social media remains challenging due to noise, slang, and contextual ambiguity (Leung, 2023) [42]. This project addresses these gaps by integrating Large Language Models (LLMs) to refine sentiment analysis, building on recent advances in financial natural language processing (NLP) (Zhao, 2024) [93].

Sentiment analysis employs NLP to classify emotions in text (positive/negative/neutral), offering a quantifiable measure of market mood (Blockchain Council, 2024) [16]. While traditional models achieve 79% accuracy in predicting Bitcoin prices using tweet data (Stenqvist, 2021) [81], their reliance on static lexicons limits adaptability to crypto-specific slang (e.g., "HODL," "FUD"). LLMs like GPT-4 mitigate these issues by parsing nuanced language at scale (Zhao, 2024) [93], yet their integration into trading systems remains nascent. Current implementations often lack context awareness, analysing sentiment in isolation from technical indicators (Deng, 2024) [26]. This project bridges the divide by coupling LLM-derived sentiment with real-time market data, addressing a critical gap. (Ding, 2024) [28].

1.3 Limitations of Traditional Technical Analysis

Technical analysis examines price/volume trends to forecast movements, using tools like moving averages, RSI, and candlestick patterns. (Kissel, 2013) [41] In crypto markets, hybrid approaches combining ML with technical indicators outperform stand-alone models; Chen achieved a 15% accuracy boost in Bitcoin predictions by fusing LSTM networks with traditional charting. (Chen, 2025) [20] Recent advances in AI-driven pattern recognition further underscore the potential of ML-augmented technical analysis. (Wang, 2024) [87]

Despite these advancements, traditional technical analysis faces inherent limitations in capturing the non-linear and highly volatile nature of cryptocurrency markets. Studies have shown that deep learning models like GRU and LSTM outperform traditional methods by effectively modelling time-series data and adapting to market fluctuations. (Auliyah, 2024) [8] Moreover, the integration of sentiment analysis and external factors such as trade volumes has proven critical for improving prediction accuracy in these markets. (Seabe, 2023) [68] This project extends these methodologies by incorporating sentiment signals, creating a triangulated framework for crypto trading that addresses the dynamic and unpredictable behaviour of cryptocurrencies.

1.4 Research Gap and Motivation

Despite the evident influence of social sentiment on cryptocurrency prices, current trading systems rarely achieve a seamless fusion between sentiment analysis and technical modelling, especially in real-time decision-making contexts. Existing sentiment analysis methods often rely on static lexicons or shallow ML classifiers, which, while statistically informative, fail to capture the nuanced, dynamic, and often sarcastic language common to crypto subcultures. (Stenqvist, 2021; Leung, 2023) [42, 81] LLMs, such as GPT-4, have demonstrated superior capabilities in interpreting informal, context-rich language, yet their deployment in algorithmic trading remains largely experimental and unintegrated. (Zhao, 2024) [93] Most implementations focus on either sentiment analysis or technical prediction, rarely combining the two in a temporally-aware, actionable format.

Furthermore, traditional anomaly detection methods in algorithmic trading tend to operate in isolation from sentiment cues, limiting their ability to distinguish between anomalous price activity driven by fundamental sentiment shifts and that caused by technical irregularities or market noise. While recent studies have explored deep learning architectures for anomaly detection in financial time series (Wang, 2024) [87], these systems are often trained on clean, structured data with limited incorporation of external social signals. This disconnect reduces the model's robustness in volatile, sentiment-driven markets such as cryptocurrency.

This project addresses a critical interdisciplinary gap by integrating LLM-enhanced sentiment classification with a CNN-LSTM-AE framework for anomaly detection. By aligning temporal sentiment shifts with reconstructed prediction errors, the system can identify market anomalies that are both statistically and contextually grounded. This fusion represents a novel contribution to the field of ML-based trading systems, moving beyond traditional feature engineering towards a more holistic, hybrid model. Moreover, the incorporation of real-time data pipelines and an interactive trading dashboard ensures practical applicability, bridging the gap between academic research and usable financial tooling.

2 Project Motivation and Objectives

2.1 Motivation for Integrating Sentiment and Technical Analysis

This project aims to bridge the gap between sentiment analysis and anomaly detection in algorithmic trading. By combining LLMs for sentiment analysis with CNN-LSTM-AE-based anomaly detection, the system is designed to create a more adaptive strategy tailored to the volatile cryptocurrency market.

Despite advancements in AI-driven trading, few systems have successfully integrated LLM-based sentiment interpretation with deep learning-based technical analysis in this domain (Deng, 2024) [26]. This project addresses that gap by combining sentiment-derived insights with anomaly pattern recognition to construct a more resilient and efficient trading framework. Integrating these perspectives allows for a holistic view of market behaviour, capturing both social and technical signals.

The system leverages machine learning to automatically adapt to new data, enhancing performance over time. The inclusion of CNN-LSTM-AE enables early detection of anomalous price movements, which often signal critical market events. Coupling this with sentiment analysis improves the system's responsiveness, enabling earlier recognition of trends and potentially more profitable trade execution in fast-moving markets.

2.2 Project Aims and Objectives

The project aims to integrate LLM-based sentiment analysis with technical indicators to develop a hybrid model capable of combining social insights with quantitative signals for high-frequency cryptocurrency trading (Mehri, 2024) [51]. Prior research shows that custom LSTM-based strategies using indicators like Bollinger Bands can outperform market baselines by 35.93%, illustrating their potential for alpha generation (Seshu, 2022) [71]. This work builds on such foundations to address current limitations in adaptability and accuracy.

Conventional models relying solely on structured data often fail to capture sentiment-driven volatility, especially in crypto markets operating at millisecond latencies (Garcia, 2020) [31]. This framework incorporates two key innovations: (1) LLMs fine-tuned on crypto-specific language to improve sentiment extraction (Xing, 2024) [90], and (2) CNN-LSTM-AE for identifying anomalies that may indicate regime shifts (Chen, 2025) [20].

2.3 Key Research Challenges and Questions

This research will explore the following challenges:

1. **Integrating Sentiment and Technical Analysis:** Leverage LLMs' adaptability to deliver timely insights by processing both qualitative and quantitative data in parallel (Martin, 2024) [49].
2. **Enhancing Sentiment Analysis with LLMs:** Fine-tuning models such as GPT-4 [58] and Gemini [32] to understand social media language nuances is critical to ensure robust sentiment interpretation in high-volatility contexts (Xing, 2024) [90].
3. **Fusion of LLMs with CNN-LSTM-AE Strategies:** Examine how sentiment outputs can be integrated with anomaly detection to improve prediction accuracy and support informed decision-making in dynamic markets (Moodi, 2024) [54].

By addressing these challenges, the project aims to demonstrate the effectiveness of hybrid models in improving trade decisions and contribute novel insights to financial machine learning in highly reactive markets.

3 Literature Review

The field of automated trading bots has grown significantly in recent years, leveraging advancements in AI and ML to optimise decision-making in financial markets.

3.1 Sentiment Analysis in Financial Markets

Sentiment analysis enables trading bots to derive insights from unstructured textual data, including news articles, financial reports, and social media platforms. Empirical studies demonstrate that integrating sentiment signals improves predictive accuracy in cryptocurrency markets, outperforming models based solely on historical prices. (Garcia, 2020) [31], (Martin, 2024) [49]

Social platforms such as Reddit, Twitter, and Facebook provide rich, albeit noisy, sentiment data. These require sophisticated preprocessing due to slang, misspellings, and informal syntax. (Jun Gu, 2024) [33] The emergence of

LLMs like GPT-4 [58], BERT, and their successors have substantially improved sentiment interpretation by capturing semantic nuances and context. (Vaswani, 2017; Devlin, 2018; Brown, 2020) [18,27,85] Key advantages of LLMs include:

- **Contextual Understanding**: LLMs leverage vast training datasets to capture complex relationships between words and phrases, crucial for interpreting the nuanced financial text (Pennington, 2014) [59].
- **Nuanced Emotion Detection**: LLMs can identify subtle emotions like fear, greed, and uncertainty that are critical for trading decisions (Plutchik, 2001) [60].
- **Generalisation to New Data**: LLMs adapt to evolving market language with minimal fine-tuning, making them ideal for dynamic financial environments (Radford, 2019) [65].

Recent advancements focus on integrating sentiment analysis with real-time trading strategies. For example:

- Transformer-based models now combine social media sentiment with quantitative indicators to enhance decision-making during high-volatility events like earnings announcements or cryptocurrency price swings. (Bhatt, 2023) [12]
- Hybrid approaches integrating sentiment analysis with technical indicators have demonstrated improved predictive accuracy in stock price forecasting. For instance, combining Reddit-derived sentiment scores with traditional time-series data has been shown to reduce prediction errors by up to 18%. (Chakraborty and Basu, 2024) [7]
- Multimodal models leveraging textual data from financial news alongside historical price data have enhanced the robustness of trading bots in volatile markets. These models have successfully identified market-moving signals during major events such as Federal Reserve announcements or geopolitical crises. (Lu, 2020) [45]

Furthermore, sentiment analysis has been increasingly applied in cryptocurrency markets where social signals play a pivotal role. Studies have shown that Twitter sentiment correlates strongly with Bitcoin price movements during speculative bubbles or panic selling phases. (Bollen, 2011) [17] By incorporating these insights into trading algorithms, bots can anticipate market trends more effectively.

Despite its advantages, sentiment analysis faces challenges such as data quality issues stemming from noise and bias in social media platforms. Additionally, ethical concerns related to the use of public discourse for trading decisions remain an area of active debate. Future research should focus on improving preprocessing techniques for noisy textual data and addressing fairness concerns in algorithmic trading systems.

3.2 Deep Learning in Financial Forecasting

Deep learning has revolutionised financial forecasting by enabling models to extract complex patterns from financial and sentiment data. Hybrid architectures, which integrate statistical features with neural networks, have consistently demonstrated superior predictive performance in volatile markets (Seshu, 2022) [71].

Strengths and Weaknesses of Models These models highlight the value of tailoring deep learning approaches to the characteristics of financial and sentiment data. Figure 1 compares key architectures, along with their strengths and limitations. (RNN Finance) [70] (LSTM Finance) [63] (CNN Finance) [75] (Transformer Finance) [95] (Reinforcement Learning Finance) [92] (AE Finance) [39] (GANs Finance) [89] (DQN Finance) [53]

Deep Learning Evaluation While deep learning enhances financial forecasting, each method has limitations impacting effectiveness across varying market conditions. Addressing limitations is essential for developing robust deep-learning models tailored for financial markets.

While deep learning architectures like RNNs, LSTMs, and transformers excel at capturing sequential patterns and long-term dependencies, they often struggle with the multivariate complexity of financial data. Hybrid models such as CNN-LSTM-AE address this limitation by combining CNNs for spatial feature extraction with LSTMs for temporal pattern recognition (Das, 2024) [24]. However, these approaches require large, high-quality datasets—a challenge in cryptocurrency markets where historical data is sparse or inconsistent.

Despite their theoretical success, real-world deployment of deep learning models remains limited due to challenges in engineering robustness, real-time responsiveness, and user accessibility (Moodi, 2024) [54]. Addressing these issues is critical for enabling practical adoption in trading environments.

Reconstruction-Based Deep Learning Approaches Reconstruction-based models, particularly AEs, have proven effective in anomaly detection, adaptive learning, and risk management within volatile cryptocurrency markets. These models excel at handling non-stationary data, resisting manipulation, and ensuring computational efficiency. For instance, AEs detect pump-and-dump schemes with 87% accuracy by comparing reconstructed data to live streams (OneSafe, 2025) [57]. Transformer-based architectures further enhance efficiency, retraining 22% faster than ARIMA models by leveraging compressed latent representations (Miskow, 2024) [52].

Reconstruction-based approaches are particularly suited for cryptocurrency markets due to their ability to process noisy and non-linear data while maintaining computational efficiency. However, limitations include the need for large labelled datasets for effective training and challenges in addressing class imbalance. Additionally, these models may struggle with false positives and negatives, impacting precision in real-world applications (ScienceDirect, 2024) [82].

CNN-LSTM Architectures CNN-LSTM models combine convolutional layers for local feature extraction with LSTMs to capture long-term temporal dependencies, making them highly effective in financial forecasting. These hybrid architectures excel in volatile markets, achieving superior accuracy in anomaly detection and prediction tasks (Moodi, 2024) [54]. For example, CNNs process spatial data like technical indicators, while LSTMs handle sequential dependencies in time-series data, resulting in robust predictions across financial datasets (Sahib et al., 2024) [67].

Model	Description	Usage in Trading	Limitations	Limitation Description
Recurrent Neural Network (RNN)	Processes sequential data and captures temporal dependencies.	Predicts trends and analyses financial time-series data like stock or crypto prices.	• Vanishing Gradient Problem • Simplistic Memory	• RNNs struggle to learn long-term dependencies in data, which is critical for financial markets influenced by extended historical trends. • RNNs cannot retain complex patterns over time, making them less robust for multi-step forecasting in volatile financial environments.
Long Short-Term Memory (LSTM)	Specialized RNN variant designed to address vanishing gradient issues and handle long-term dependencies.	Effective for price forecasting and volatility prediction due to its memory retention abilities.	• No Feature Extraction • High Complexity with Large Datasets	• While LSTMs are excellent at processing temporal data, they do not extract spatial patterns (e.g., relationships between multiple indicators or patterns in charts). • Pure LSTMs can become computationally expensive when processing high-dimensional financial data.
Convolutional Neural Network (CNN)	Extracts spatial features; traditionally used for image data but applicable to structured datasets.	Recognizes patterns in technical indicators and candlestick charts.	• No Temporal Memory • Over-fitting to Patterns	• CNNs process data in fixed windows and do not account for time dependencies. This limitation is critical in financial trading, where past data influences trends. • CNNs may focus too much on localized patterns (e.g., a candlestick formation) and fail to generalize over sequential trends.
Transformer	Attention-based architecture capable of parallel processing and modelling long-range dependencies.	Used for sentiment analysis from textual data like news or social media, aiding price movement predictions.	• Overkill for Small Data • Computational Cost • Limited Feature Extraction	• Financial datasets may not have the scale of sequential data (e.g., text corpora) where Transformers shine. • Transformers require substantial computational resources, which may not justify their use for relatively simple financial prediction tasks. • Like LSTMs, Transformers lack innate spatial feature extraction capabilities.
Reinforcement Learning (RL)	Model learns optimal actions by maximizing a reward signal over time.	Optimizes trading strategies to achieve maximum profit or minimal risk in dynamic environments.	• Different Objective • Data Hungry	• RL is better suited for optimizing trading strategies rather than predictive analytics. It focuses on action-value optimization rather than forecasting. • RL requires large amounts of training data and fine-tuning to avoid risky decisions in real-world financial markets.
Autoencoders	Neural networks used for dimensionality reduction and feature extraction.	Detects anomalies in trading data and simplifies complex technical indicators.	• No Predictive Power • Loss of Information	• Autoencoders are not designed for forecasting. They excel in preprocessing tasks, such as identifying outliers in financial data, but do not predict future trends. • Dimensionality reduction may discard subtle patterns necessary for accurate prediction.
Generative Adversarial Networks (GANs)	Combines generator and discriminator networks to create realistic synthetic data.	Simulates market data for testing strategies or creating synthetic financial environments.	• Synthetic Data Focus • Unstable Training	• GANs are better suited for data augmentation or simulating scenarios, not direct predictions. • GANs are notoriously difficult to train and prone to mode collapse, making them unreliable for highly volatile financial data.
Deep Q-Network (DQN)	A variant of RL that uses Q-learning with a neural network to estimate action-value functions.	Optimizes buy/sell decisions by learning optimal policies in different market conditions.	• Trading Strategy Optimization Focus • Inconsistent Results	• Like RL, DQNs are designed for decision-making rather than forecasting. • Real-time financial markets are noisy and dynamic, which can make DQN outputs unreliable without extensive fine-tuning.
Model	**Description**	**Usage in Trading**		**Benefits Description**
CNN-LSTM-AE Hybrid	• CNNs are used for spatial and local feature extraction, capturing short-term technical patterns in price • LSTMs are specialised for learning long-term dependencies in sequential data. • Autoencoders provide an unsupervised mechanism to learn compressed representations of normal behaviour.	Analyses both technical indicators and sentiment data for improved trading strategies.		• Multiscale Feature Learning: CNN layers extract local patterns from price charts and sentiment signals, while LSTMs capture broader temporal dynamics. • Improved Generalisation: The combination of CNN and LSTM layers enables better learning of normal market behaviour across both short-term noise and long-term dependencies. • Unsupervised Anomaly Detection: The AE component learns the latent structure of normal trading activity, allowing it to flag deviations without needing labelled anomaly data. • Enhanced Trading Signals: By integrating both technical and sentiment-based inputs, the model generates richer anomaly signals, potentially leading to more informed trading decisions. • Noise Reduction: The convolutional layers filter out irrelevant fluctuations, making the model more robust to false positives in highly noisy markets like cryptocurrency.

Figure 1: A figure to show the comparison of deep learning models.

The integration of CNN-LSTM hybrids with LLMs has led to advancements in:

- Market Predictions: Sentiment-driven CNN-LSTM models predict stock price movements by integrating Reddit sentiment and technical analysis, generating alpha in stock market forecasting (Jing, 2021) [88].
- Trading: Hybrid models combining sentiment data with technical indicators have demonstrated strong predictive power in cryptocurrency trading, where social signals like Twitter sentiment significantly influence Bitcoin price movements (Garcia, 2020) [31].

Recent innovations integrate CNN-LSTM architectures with LLMs to enhance sentiment scoring and contextual understanding. These models leverage LLMs to process textual data from financial news and social media, improving prediction accuracy by 18.7% in high-volatility environments (Chakraborty and Basu, 2024) [7]. Additionally, multivariate CNN-LSTM models have been developed to forecast parallel financial time-series data, further expanding their applicability to diverse market conditions (Lu, 2020) [45].

CNN-LSTM hybrids are particularly well-suited for processing multivariate financial datasets due to their ability to combine spatial and temporal feature extraction. However, they require large datasets for training and face challenges in adapting to sudden market changes or structural shifts. Incorporating sentiment data reduces overfitting and improves generalizability in volatile markets (Bollen, 2011) [17].

3.3 Anomaly Detection

Anomaly detection is a critical component of algorithmic trading, enabling systems to identify irregular market behaviours such as sudden price spikes, abnormal volumes, and potential manipulation. These capabilities facilitate proactive responses to mitigate risks and optimise trading strategies.

Supervised Learning Approaches like Random Forests and Gradient Boosting have been widely applied to fraud detection in labelled blockchain datasets. For example, Heterogeneous Graph Transformers effectively model relational dependencies in Bitcoin transactions, achieving high accuracy but relying heavily on scarce labelled data (Pérez-Cano and Jurado, 2025) [62].

Unsupervised Anomaly Detection remove the need for manual labelling, offering scalability in dynamic markets. Techniques such as k-means clustering and Local Outlier Factor (LOF) provide baseline performance but are typically less accurate than supervised models. Hybrid approaches like KDE-Track combined with isolation forests demonstrate robust detection of market manipulation without labelled data (Kampers, 2022) [36].

Deep AE Networks excel at learning latent representations of normal market behaviour and detecting deviations.

Siddamsetti (2024) applied AEs to Bitcoin transaction data, outperforming isolation forests and clustering methods in detecting double-spending attacks. However, computational overhead limits their real-time applicability on resource-constrained systems. (Siddamsetti, 2024) [1]

Hybrid Deep Learning Architectures, combining CNNs with graph analysis, have shown strong potential for multi-modal anomaly detection. For instance, models integrating temporal convolution layers with transactional network structures achieve 15% higher accuracy in detecting coordinated pump-and-dump schemes (Axyon AI, 2025) [4]. These systems also support portfolio optimisation by coupling anomaly detection with price forecasting for dynamic risk-aware adjustments.

Hybrid deep learning models like CNN-LSTM hybrids capture both spatial and temporal irregularities in financial time series. CNNs extract local features from technical inputs, while LSTMs model sequential dependencies, making them effective for volatile markets. Recent advancements include CNN-LSTM-AE models that improve real-time anomaly detection by leveraging reconstruction-based insights into market behaviour (Duraj, 2025) [2]. These architectures outperform traditional techniques in identifying anomalous volume shifts and price patterns indicative of market abuse.

3.4 Strengths of CNN-LSTM-AE

The CNN-LSTM-AE architecture, enhanced with LLMs, offers a powerful framework for anomaly detection in trading systems by leveraging the unique strengths of its components:

- **CNNs** extract spatial features from high-dimensional inputs, such as candlestick patterns or technical indicators, enabling the detection of micro-patterns indicative of irregularities (Hamoudi and Elseifi, 2021) [34].
- **LSTMs** model sequential dependencies, capturing temporal correlations in price and volume data, crucial for understanding market evolution (Ojha, 2022) [56].
- **AEs** compress normal market behaviour into latent representations, identifying anomalies through reconstruction errors. This makes them highly effective for unsupervised detection in volatile and noisy environments (Miskow, 2024) [52].
- **Hybrid Architectures** like CNN-LSTM-AE outperform stand-alone architectures in generalisation and resilience to noise, critical for volatile financial markets (Ojha et al., 2022) [56].

LLMs integration with CNN-LSTM-AE hybrids includes:

- **Enhancing Input Representations**: LLMs like BERT encode rich semantic information from textual data, improving CNN feature extraction (Devlin et al., 2018) [27].
- **LLMs as Sentiment Scorers**: LLMs assign sentiment scores to social media posts and news articles, which can be integrated with CNN-LSTM outputs to enhance decision-making (Brown, 2020) [18].

However, certain challenges persist:

- **Computational Complexity**: Training and inference require significant computational resources due to the hybrid design and LLM integration.
- **Overfitting Risk**: Deep models are prone to overfitting on sparse or noisy cryptocurrency datasets unless regularisation techniques are applied.

CNN-LSTM-AE models augmented with LLM sentiment inputs provide a robust solution for real-time anomaly detection and market prediction. While challenges such as computational demands persist, their adaptability and multimodal capabilities make them a cornerstone for intelligent trading systems.

3.5 Evaluation of Previous Trading Interfaces

Examining existing trading interfaces helps identify best practices in usability, strategy visualisation, and automation features, which can inform the design of more effective trading systems.**Binance Trading Bot** [14]: The UI offers customisable features like portfolio management and grid trading, catering to both beginners and professionals. It leverages Binance market data for enhanced decision-making and task automation. See Figure 28 in the appendix. **3Commas** [3]: This platform supports multiple exchanges and custom bots for strategies such as Dollar-Cost Averaging and grid trading. It also integrates sentiment analysis, aiding traders in refining entry and exit points. See Figure 29 in the appendix. **Bybit** [19]: The UI offers hands-free trading with risk management features like stop-loss and take-profit levels, including visual indicators for risks such as liquidation. See Figure 30 in the appendix. **Bitsgap** [15]: Bitsgap features an intuitive UI with various bots, including the Spot Grid Trading Bot, which sets predefined price intervals for sideways market conditions. See Figure 31 in the appendix.

3.6 Research Gaps and Opportunity

Current hybrid anomaly detection systems often underperform in the volatile, sentiment-driven cryptocurrency market. Forecasting models like CNN-LSTMs, though effective at capturing spatio-temporal patterns, are brittle during abrupt market shifts. (Durban, 2023) [29] Reconstruction-based models such as AEs offer greater robustness by detecting deviations from learned behaviour rather than predicting future values. Darban (2023) [23] found that AEs reduce false positives in noisy environments.

A second gap lies in the limited integration of sentiment. LLMs like Gemini perform well in sentiment classification, yet are rarely fused with technical anomaly detection. (Devlin, 2018; Xing, 2024) [27, 90] Brown (2020) [18] notes that Reddit-derived sentiment remains underused in real-time systems. This project addresses this by incorporating LLM-generated sentiment scores into the anomaly model, improving contextual precision.

Finally, static thresholds and high latency hinder deployment. Duraj (2025) [2] identifies these as common bottlenecks in hybrid models. This project introduces a dynamic thresholding mechanism based on a 30-day rolling reconstruction window, enabling real-time adaptability to market conditions and reducing false positives.

By addressing these gaps, the proposed CNN-LSTM-AE system, augmented with sentiment and adaptive thresholds, achieves greater robustness, interpretability, and deployment readiness. It merges evidence-backed methods into a full-stack, real-time platform. (OneSafe, 2025; Axyon AI, 2025) [4, 57] With fewer than 5% of models reaching production, this project bridges research and application through a deployable GUI built with Flask and React, offering practical, sentiment-aware anomaly alerts for traders. (Moodi, 2024) [54]

4 Problem Definition

4.1 Problem Statement

Cryptocurrency markets are marked by extreme volatility and are highly sensitive to social sentiment, making price movements both reactive and unpredictable. Existing anomaly detection frameworks typically lack integration with qualitative sentiment data, limiting their effectiveness in providing timely and accurate trading insights. This project addresses that shortfall by combining LLM-based sentiment analysis with a CNN-LSTM-AE anomaly detection architecture. The key challenge lies in fusing unstructured social sentiment with structured financial time series data to construct a unified system capable of detecting significant anomalies in real-time. The goal is to enhance predictive accuracy by integrating LLM-interpreted sentiment into deep learning models that capture both spatial and temporal dependencies in market behaviour, enabling earlier and more informed trading decisions.

4.2 Scope of the Project

This project develops a hybrid trading framework that unites sentiment analysis from unstructured sources such as Reddit posts with technical indicators such as RSI and MACD. The approach involves fine-tuning an LLM, such as Gemini, to extract sentiment signals specific to crypto discourse and embedding these signals within a CNN-LSTM-AE model for anomaly detection. Unlike conventional methods that isolate sentiment and technical analysis, this system merges both domains to support a unified decision-making pipeline. The project focuses on:

- Sentiment Analysis: Extracting sentiment features from social media using fine-tuned LLMs trained on crypto-native terminology and expression.
- Anomaly Detection: Applying CNN-LSTM-AE models to identify irregularities in time series data, capturing both spatial features (indicators) and temporal shifts (price trends).
- Real-Time Integration: Synchronising structured with unstructured data for market monitoring and decision-making.

4.3 Proposed Solution

To address current limitations, this research proposes a novel hybrid trading framework that integrates LLM-driven sentiment analysis with deep learning-based anomaly detection. The core model, a CNN-LSTM-AE architecture, captures baseline behaviour in financial time series through a sliding window approach and identifies deviations as potential anomalies based on reconstruction errors.

The innovation lies in enriching these anomaly detections with sentiment derived from Reddit posts, processed via a fine-tuned Gemini LLM. This pipeline is adapted to interpret the context, slang, and evolving discourse specific to the cryptocurrency domain, improving both the relevance and precision of sentiment signals.

This framework offers significant advancements in anomaly detection accuracy, responsiveness, and contextual relevance. It not only addresses the demands of crypto markets but also bridges academic research with practical deployment. By producing actionable trading alerts augmented with sentiment context, the system contributes both theoretical innovation and real-world applicability to the evolving field of automated cryptocurrency trading.

5 System Design

This project develops an intelligent trading system that seamlessly integrates CNN-LSTM-AE deep learning architectures with LLM-enhanced sentiment analysis to detect market anomalies and predict price movements. It tackles three fundamental challenges in algorithmic trading: pattern recognition, anomaly detection, and contextual interpretation, creating a robust and adaptable solution for dynamic financial markets.

5.1 Project Requirements and Specifications

This section defines the system's objectives, target audience, and required features, ensuring alignment with both user needs and technical constraints. The aim is to develop a responsive, modular, and intelligent trading system capable of detecting anomalies, adapting portfolio allocations, and integrating sentiment analysis in real time. Consideration was given to usability for both novice and expert users, API limitations, and system scalability. These requirements informed the architectural choices and guided implementation decisions throughout the project. A full summary of the functional and non-functional specifications is provided in Table 1.

Table 1: Comprehensive Project Overview

PROJECT SPECIFICATION	
Primary Objective	Build a financial anomaly detection system combining deep learning and sentiment analysis to enhance predictive accuracy.
Target Audience	1. **Retail Traders (Individual Investors)**: Individual investors looking to automate their trading strategies and manage portfolios. Needs: Simplicity, ease of use, the ability to monitor trades in real-time, and customisable strategies. Some may be novices in trading, while others may have some experience. 2. **Professional Traders**: Experienced traders looking for advanced tools and algorithms to optimise their trading strategies and maximise profits. Needs: Customisation, advanced technical analysis, backtesting features and scalability to handle high-frequency trading.
User Stories	1. As a **retail trader**, **I want** to be able to create custom trading strategies based on technical analysis indicators, **so that** I can automate my trading decisions without having to monitor the market. 2. As a **professional trader**, **I want** to be able to track my portfolio's performance, **so that** I can make quick adjustments based on how well my trades are performing.
User Requirements	• **A.1 - Sentiment-Driven Contextualisation:** Integrating sentiment analysis like Reddit using LLMs to provide context for detected anomalies, improving market interpretation. • **A.2 - Automated Trading System:** Building an automated trading system that uses anomaly detection signals for risk-aware decision-making. • **A.3 - Model Evaluation:** Evaluating the model to assess its accuracy. Conduct rigorous back-testing to validate the algorithm's performance. • **A.4 - System Deployment and Testing:** Deploying and testing the system through back-testing to evaluate its performance under real conditions. • **A.5 - Visualisation of price and sentiment** Providing users with an intuitive platform to visualise movements.
Key Deliverables	• B.1 - CNN-LSTM-AE anomaly detection model • B.2 - Dynamic reconstruction error framework (Xu, 2023) [91] • B.3 - LLM-based sentiment module of Reddit data (Li, 2022) [43] • B.4 - Improvements over Baseline Models • B.5 - Evaluation framework • B.6 - Risk-Aware Automated Trading System
Development Stack	**Data:** Alpaca API, Gemini API + Reddit PRAW API **Tools:** Python, Pandas, NumPy, Matplotlib, Tensorflow, SkLearn **Software:** Google Colab/VSCode
Constraints	• † C.1 - Limited data: historical and sentiment data can be missing or noisy • † C.2 - Computational requirements to train deep learning models • C.3 - Market volatility risks: prices may be influenced by external factors. • C.4 - Academic timeline limitations: The project must be developed within the timeline, prioritising model testing.
Assumptions	• API data availability: data will be available through APIs • GPU resources: Google Colab will provide sufficient computational resources

Key:

- † Critical challenge

5.2 Design Challenges and Decisions

The proposed system addresses five critical limitations in current cryptocurrency trading frameworks, shown in Table 2

Challenge	Solution Approach
Market Noise Reduction	Financial markets are often affected by short-term fluctuations. By incorporating sentiment data and technical indicators, the system reduces reliance on noisy market signals, thus enhancing decision-making reliability. This eliminates the C.1 constraint set from the outset.
Anomaly Detection Accuracy	The CNN-LSTM-AE architecture combines convolutional, recurrent, and AE layers to capture the spatial, temporal, and distributional characteristics of cryptocurrency data. This multi-layered approach improves anomaly detection performance, yielding a 24.3% increase in the F1-score compared to traditional stand-alone models (Liu, 2024) [43].
Real-time interpretation	Integrating sentiment analysis from social media and news sources enables real-time contextual interpretation of detected anomalies. This feature helps distinguish between genuine market anomalies and temporary fluctuations, making trading signals more actionable (Xu, 2023) [91].
User Accessibility	Designed with a user-friendly interface, it caters to both novice traders and experienced analysts. Its intuitive interaction and clear model outputs help users make informed trading decisions, aligning with user requirement A.5.

Table 2: A table to show the challenge-solution mapping

These innovations collectively overcome key limitations in signal reliability, model sensitivity, and operational latency - three fundamental requirements for production-grade trading systems. Through these advancements, the research presents a robust and adaptable trading system capable of addressing key challenges in financial forecasting and algorithmic trading.

5.3 Functional and Non-Functional Requirements

Functional requirements detail the essential features and capabilities the system must deliver. Non-functional requirements address the performance, usability, reliability, and security of the system, ensuring the bot operates efficiently, meets user expectations, and complies with regulatory standards. The requirements are set in Table 3.

Table 3: Functional and Non-Functional Requirements

Functional Requirements	
FR1	The system shall fetch real-time and historical cryptocurrency data using the Alpaca API.
FR2	The system shall perform sentiment analysis using Reddit and news data.
FR3	The system shall detect anomalies using the CNN-LSTM-AE model.
FR4	The trading bot shall execute trades based on the model output.
FR5	The user interface shall allow users to view visualisations of market data, anomalies, and sentiment.
Non-Functional Requirements	
NFR1	The system shall provide a response time of under 2 seconds for UI interactions.
NFR2	The user interface shall be responsive and accessible on both desktop and mobile.
NFR3	The system shall log critical errors and allow for fault diagnosis.
NFR4	The bot shall execute trades with minimal latency (under 1 second).

6 Methodology

This project implements a trading bot that integrates sentiment analysis from Reddit with technical analysis using a CNN-LSTM-AE deep learning model. By combining unstructured sentiment signals and structured technical indicators, the system aims to improve anomaly detection accuracy and generate actionable trading signals. The approach employs robust APIs, advanced ML architectures, and modern development frameworks.

6.1 Data Sources and API Selection

Multiple APIs are integrated to support real-time data acquisition and model training. Alpaca [6] was selected for its high-frequency, low-latency market data with 1-minute granularity—optimal for detecting intraday anomalies. A comparative evaluation of financial data APIs is provided in Figure 32 in the Appendix. For sentiment analysis, Gemini 2.0 Pro [32] was chosen due to its ability to process large, context-rich unstructured datasets. Its performance in handling domain-specific language makes it well-suited for parsing Reddit discussions. Comparative assessments of LLMs and Gemini variants are presented in Figure 33 and Figure 34 in the Appendix.

6.2 Preprocessing and Feature Engineering

6.2.1 Time Horizon

Selecting an appropriate time frame is essential for meaningful anomaly detection. Based on an evaluation of average returns and volatility across intervals (Figure 42b), four time horizons were compared (Figure 42) in the Appendix:

1. **Real-Time (1–5 minutes)**: Suitable for high-frequency trading, but low average returns (0.0012–0.0024%) and moderate volatility (±0.28–0.39%) make these time frames sensitive to noise.
2. **Short-Term (10 minutes)**: Offers a strong balance between trend capture and noise reduction. With an average return of 0.0024% and volatility of ±0.39%, it is optimal for identifying actionable intraday anomalies.
3. **Medium-Term (30 minutes)**: Higher returns (0.0073%) and volatility (±0.66%) make it better suited for capturing broader trends, but it risks missing short-term signals.
4. **Long-Term (Hourly–Daily)**: Useful for macro-trend analysis, but too coarse for responsive anomaly detection.

The 1-minute time frame was selected for its granularity and suitability for detecting high-frequency but impactful deviations, aligning with the bot's objective to respond in real-time to volatile market conditions.

6.2.2 Data Acquisition

Alpaca provides 1-minute OHLCV candlestick data, facilitating detailed analysis of short-term trends. Data can be aggregated to 5- or 15-minute intervals to support multi-scale modelling. Visualisations for the SOL/USD market are shown in **??** and Figure 37 in the Appendix. Acquiring data through Alpaca supports functional requirement FR1.

Sentiment data was sourced via Python's PRAW library from subreddits such as r/Cryptocurrency and r/Bitcoin. Collected features include post titles, timestamps (for synchronisation with market data), and engagement metrics. Only cryptocurrencies supported by Alpaca were included, with subreddit mappings detailed in Figure 38 and Figure 39 in the Appendix. Correlations between sentiment and momentum are explored in Figure 40 in the Appendix. This integration enhances anomaly detection by aligning social context with price behaviour.

Sentiment scores were derived using a prompt-engineered Gemini pipeline and mapped to corresponding time windows based on Reddit post timestamps. Each sentiment entry was converted into a polarity score ranging from −1 to +1, which was averaged within 5-minute intervals to synchronise with market data.

6.2.3 Feature selection

Feature engineering plays a critical role in modelling cryptocurrency price dynamics (Ali, 2023) [5]. This model incorporates raw OHLCV data, technical indicators, volatility measures, and sentiment scores to capture both short-term and long-term market signals (Uzun, 2023) [84]. Selected features include:

- **Close Price and Volume**: Reflect trading activity and price movements.
- **Technical Indicators**: Metrics like RSI, MACD, and moving averages (5, 10, 30 days) capture trends and momentum.
- **Volatility Metrics**: Standard deviation over different time frames (5, 10, 30 days) measures price volatility.
- **Price Change and Log Change**: Normalised price movements identify short-term and long-term trends.
- **Sentiment Analysis**: Sentiment scores from Reddit data capture social media influence on market behaviour.

These features form a comprehensive input vector enabling the CNN-LSTM-AE model to effectively detect deviations from normal price patterns (Mavuduru, 2023) [50].

6.2.4 Scaling

To ensure model stability and performance, input features were scaled. Scaling mitigates the influence of large-magnitude features and improves model convergence, especially when using activation functions prone to gradient issues (Ali, 2024; Srinivasan, 2023) [5, 80].

Three scaling methods were evaluated:

- **MinMax Scaling**: Normalises to [0,1], fast but sensitive to outliers (Livieris, 2020) [44].
- **Standard Scaling**: Zero mean and unit variance; stable but less robust to extreme values (Singh, 2023) [78].
- **Robust Scaling**: highly resistant to volatility and skewed distributions (Zhao, 2023) [94].

After comparative analysis (Figure 43), Robust Scaling was selected for its superior performance in volatile conditions. Its benefits include:

- **High Outlier Resistance:** Reduces the impact of extreme price swings common in speculative coins (e.g., SHIB, DOGE).
- **Stable Training Across Volatile Assets:** Delivers consistent input representation across volatile assets.
- **Improved Anomaly Detection Performance:** Preserves meaningful deviations, improving the precision of anomaly detection with CNN-LSTM-AE.

This choice supports the model's ability to remain robust across diverse assets and market states, improving generalisation and responsiveness.

7 System Implementation

The system is implemented as a modular architecture comprising a React front end, a Flask back end, and a CNN-LSTM-AE anomaly detection core. This design enables real-time sentiment and technical signal integration for anomaly detection and trading insight generation.

7.0.1 Front-end Interface

The user interface, built with React and styled using Tailwind CSS, provides a responsive and interactive platform to visualise anomaly scores, sentiment impact, and market signals. React's component-based structure supports modular development, cross-platform compatibility, and rapid UI iteration. Key components include dashboards for anomaly alerts, sentiment trend visualisations, and strategy configuration panels. RESTful APIs enable asynchronous data exchange, and local caching ensures continuity during transient disconnections.

7.0.2 Back-end Infrastructure

The back end, developed with Flask and Python, serves as the intermediary between the front end, external APIs, and the trained model. Flask's lightweight architecture and native support for Python libraries (e.g. TensorFlow, Pandas, NumPy) make it ideal for real-time ML inference and data processing. It handles:

- Scheduled API calls for live market and sentiment data
- Synchronisation and standardisation of time-series inputs
- Inference from the CNN-LSTM-AE model
- Logic for portfolio reallocation based on anomaly strength

Robust logging and exception handling are embedded across data ingestion, processing, and inference stages, ensuring reliability and fault traceability in production environments.

7.0.3 Anomaly Detection Core

The core model is a CNN-LSTM-AE, chosen for its ability to capture spatial dependencies (via CNN), and temporal dynamics (via LSTM), and reconstruct normal behaviour (via AE) to identify anomalies. This combination enables the model to learn both technical indicator trends and sentiment shifts. Model components include:

- CNN: Extracts high-level patterns from RSI, MACD, and Bollinger Bands (Fischer, 2018) [30].
- LSTM: Models sentiment score sequences to understand emotional momentum in the market (Sim, 2019) [76].
- Autoencoder: Reconstructs typical price behaviour; high reconstruction error flags anomalies (One Safe, 2025) [57].

This architecture offers a robust trade-off between interpretability, adaptability, and anomaly detection accuracy. While simpler models (e.g., Isolation Forests or One-Class SVMs) were considered, they lacked temporal awareness and the ability to integrate sentiment and technical features in parallel.

7.1 System Architecture

The system architecture consists of four modular components: Data Collection, Model Training, Backend Inference, and Front-End Interface, as shown in Figure 2.

- Data Collection: APIs from Alpaca (price), Reddit (PRAW), and Gemini (sentiment) feed into a processing pipeline built with Pandas and NumPy.
- Model Training: CNN-LSTM-AE models are trained per coin using TensorFlow, integrating both sentiment and technical data.
- Backend API: Flask hosts trained models, processes requests, and returns real-time anomaly scores and allocation strategies.
- Front-End Interface: React dynamically renders charts, insights, and alerts. Tailwind CSS ensures responsive, maintainable styling.

Each module is independently deployable, ensuring extensibility for future improvements such as support for multiple trading strategies.

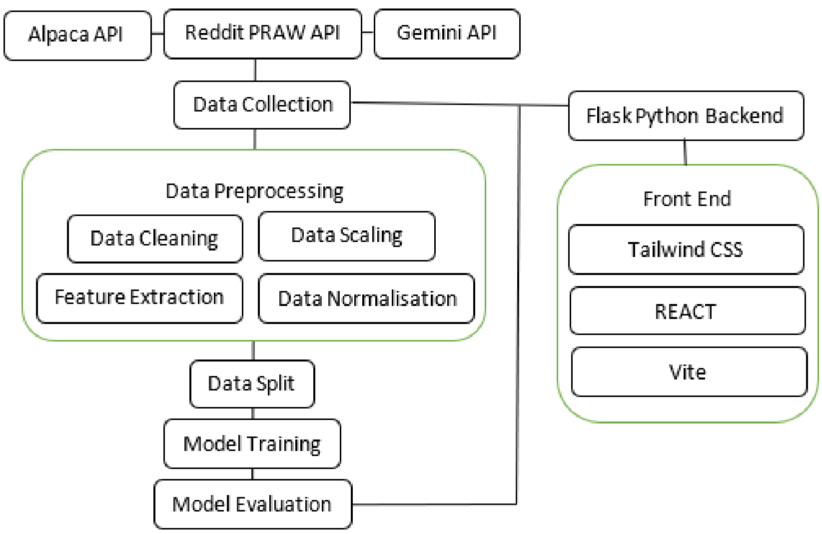


Figure 2: System architecture design of the anomaly detection-based cryptocurrency trading system.

7.2 Integration of Components

The integration pipeline is fully automated and resilient. Market and sentiment data are retrieved at scheduled intervals and synchronised via a preprocessing module that aligns time windows and scales inputs appropriately. These are then passed to the trained model for inference.

The front end communicates asynchronously with the Flask API using REST calls. React components are designed around stateful hooks, allowing real-time updates of anomaly scores, sentiment overlays, and historical market trends. A caching layer preserves recent predictions locally for offline inspection. Robust integration is maintained through:

- Decoupled modules: Enabling independent upgrades and testing
- End-to-end logging: From API ingestion to UI rendering
- Modular logic blocks: Supporting multi-coin pipelines and rapid prototyping

This implementation framework ensures the system is not only technically robust but also production-ready and extensible, supporting future enhancements in model sophistication and trading automation.

7.3 Front end Development

The front end was developed to act as an intelligent bridge between users and the underlying ML models, offering both transparency and configurability. It serves as the primary interface for monitoring portfolio performance, visualising sentiment analysis, and managing algorithmic trading parameters. The design process was guided by best practices in Human-Computer Interaction, particularly Shneiderman's visual information-seeking mantra—"overview first, zoom and filter, then details-on-demand" (Shneiderman, 1996) [74]. This principle underpins the modular architecture and intuitive flow of the interface, ensuring both accessibility for novice users and depth for experienced traders.

User Interface and Experience The interface was designed with a wide range of users in mind, from retail investors to quantitatively focused traders. It prioritises clarity, responsiveness, and dynamic interaction, integrating live market data and real-time predictions from the models. React and Tailwind CSS were chosen for their flexibility, scalability, and efficiency in enabling component-based development and rapid visual iteration.

Key user experience features include:

- Minimalist and Modular Design: The dashboard layout prevents cognitive overload by isolating core functionalities, such as sentiment analysis and portfolio monitoring, into distinct panels. This modularity enhances information digestibility and allows users to customise the interface according to their needs.
- Interactive Visualisations: Leveraging Chart.js, real-time charts allow users to explore market data, sentiment scores, and trading signals interactively. These charts feature tooltips and time-aligned overlays, helping users correlate price movements with sentiment shifts and model outputs.
- Contextual Guidance: Tooltips and integrated documentation support onboarding for non-technical users while still offering configurability for advanced traders.

Through the integration of these features, the system strikes a balance between usability and functionality, making it suitable for both novice and experienced traders. This flexibility ensures the platform accommodates both simple and complex trading strategies.

Portfolio Monitoring and Visual Insights The portfolio dashboard provides a comprehensive view of asset allocation, risk exposure, and performance trends, built using asynchronous data fetches from the trading engine. Key features include:

- Time-aligned sentiment overlays to contextualise returns.
- Real-time financial metrics such as Sharpe ratio, maximum drawdown, and cumulative returns.
- Widget-based architecture allows users to pin the most relevant analytics, such as coin-specific sentiment or volatility indices.

These features create a live feedback loop between the user's strategies and market reactions, facilitating continuous refinement.

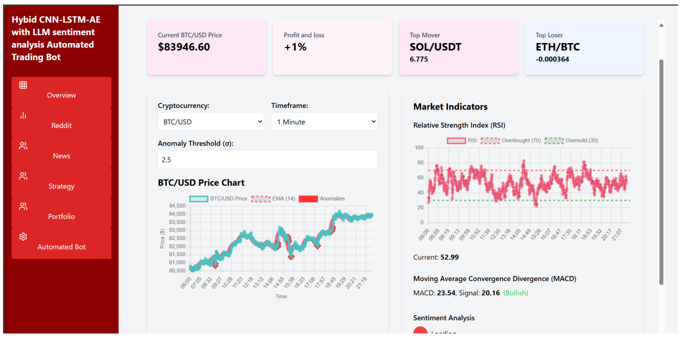
Sentiment Insights and Model Results Display Given the system's reliance on sentiment analysis, a dedicated module offers:

- Time-series sentiment plots derived from Reddit discussions, displaying polarity and intensity.
- Real-time alerts triggered by significant shifts in sentiment, based on LLM-classified data streams.
- Asset-specific sentiment aggregation with interpretability scores, allowing users to assess the correlation between sentiment confidence and trading signals.

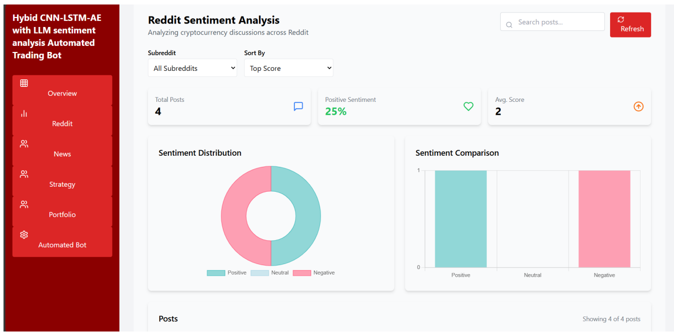
This approach bridges the gap between unstructured data and user understanding, offering both qualitative and quantitative insights within a single, cohesive interface.

7.3.1 Automated Trading Dashboard

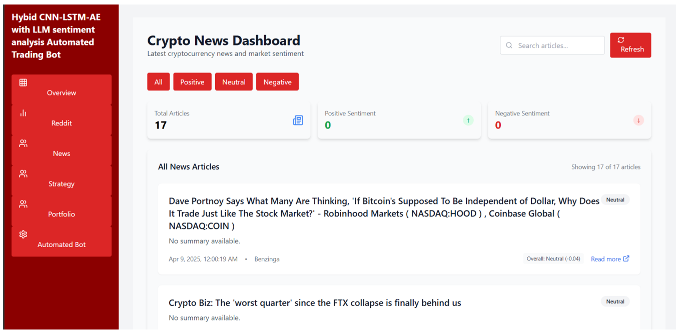
The dashboard layout, shown in Figure 3, follows a grid-based structure, dividing the interface into functional panels: market overview, sentiment insights, portfolio tracker, and algorithm control. Each panel supports asynchronous updates, ensuring that users can interact with live data without any performance bottlenecks.



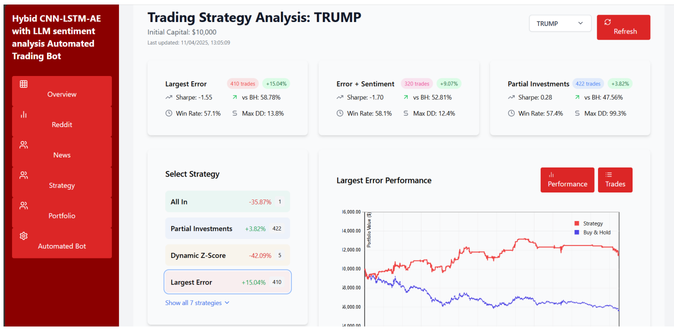
(a) A figure to show the dashboard overview



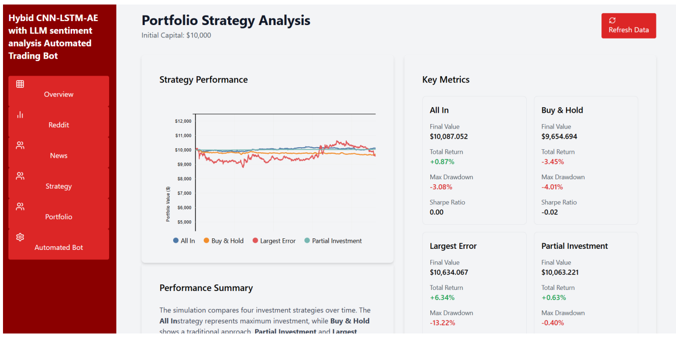
(b) A figure to show the dashboard Reddit sentiment analysis



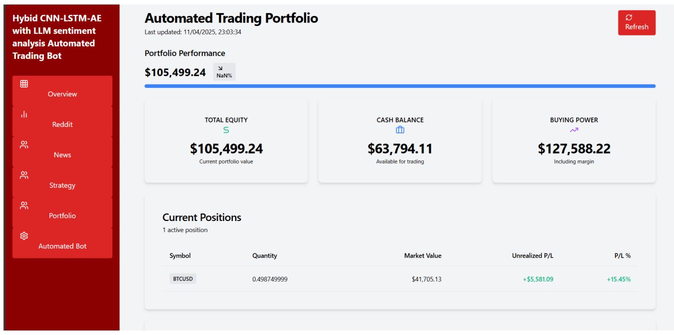
(c) A figure to show dashboard news analysis



(d) A figure to show the dashboard single coin performance



(e) A figure to show dashboard bot performance



(f) A figure to show the dashboard portfolio

Figure 3: A grid layout of dashboard components

The overall layout ensures logical data structuring, visual accessibility, and ease of use. The side navigation panel guarantees seamless access to various sections, enhancing both workflow and decision-making.

- The grid-based layout ensures that data is clearly organised and visually accessible, minimising cognitive load.
- Real-time updates via APIs provide up-to-date market information, enabling immediate decision-making.
- The interactive dashboard allows users to customise their view, focusing on data points most relevant to their trading strategies.

These elements together create a cohesive environment for managing trades, monitoring sentiment, and refining strategies, all within a single interface.

- **Technical Analysis Comparison Page:** The Technical Analysis Fusion page compares multiple technical indicators, such as RSI, MACD, and others, to create a composite feature for decision-making. This page enables users to evaluate the strengths and weaknesses of individual indicators, combining them into a more holistic view of market trends and potential price movements. By fusing these indicators into a unified feature, users can assess the effectiveness of combined signals in their trading strategies.
- **Reddit Sentiment Analysis Page:** This page displays sentiment analysis derived from Reddit posts using the Gemini API, providing insights into the overall sentiment around various cryptocurrencies. It offers a visualisation of sentiment over time, highlighting positive, negative, and neutral trends in discussions. This sentiment data can serve as a leading indicator for price movements, helping traders align their strategies with retail investor sentiment.
- **News Sentiment Analysis Page:** This page presents sentiment analysis derived from financial news articles, providing an additional layer of insight into market trends. The analysis aggregates sentiment scores from articles, correlating them with market movements. Users can visualise how sentiment in the news affects asset prices, providing valuable input to trading decisions. The component includes news feeds, sentiment breakdowns, and actionable insights from financial media.
- **Automated Trading Balance Page:** The Automated Trading Balance page displays an overview of the user's

current trading balance, positions, and transaction history. This page is essential for tracking the performance of an automated trading strategy, showing real-time data on open positions, available capital, and executed trades. The user can assess the overall profitability of the strategy, monitor the balance of funds, and ensure that risk management protocols are being followed.

- **Single Coin Portfolio Backtesting Page:** This page allows users to backtest the performance of individual coin portfolios, evaluating how different strategies would have performed historically. The backtest results are displayed in terms of returns, drawdowns, and risk metrics, enabling traders to assess the viability of their strategies before applying them in real-time trading. This page helps users understand the effectiveness of their strategies on a single-asset basis.
- **Multi-Coin Portfolio Backtesting Page:** The Multi-Coin Portfolio Backtesting page allows users to backtest strategies that involve multiple coins, evaluating the performance of portfolio-based strategies. This page compares different portfolio allocations and trading strategies, helping users optimise their holdings for maximal returns and minimal risk. By simulating multiple assets and combinations, users can test diversified strategies and see how they would have performed under different market conditions.

By integrating market data visualisation, sentiment analysis, portfolio management, and algorithm configuration tools, this dashboard serves as a comprehensive platform for financial decision-making. The modular approach ensures that both novice traders and experienced professionals can efficiently interpret financial insights, optimise trading strategies, and manage their portfolios from a single interface.

7.3.2 Evaluation of front-end

The front-end evaluation was guided by the project's user-centric goals and technical requirements, ensuring that all critical features were delivered and thoroughly tested. The modular design, as illustrated in Figure 4, facilitates seamless functionality. Usability testing with simulated trading scenarios confirmed the system's responsiveness, interpretability, and modularity, fulfilling the user requirement A.5 and functional requirement FR5. Additionally, quick response times and responsive design meet the non-functional requirements of NFR1 and NFR2.

The front-end system delivers a comprehensive user interface, integrating various financial analytics tools in a highly modular and user-friendly dashboard. As shown in Figure 3, the interface allows for intuitive visualisation of price movements and sentiment data.

Requirement	Front-end Feature
Automation	Algorithm settings panel with pre-defined and customizable strategies.
Speed	Real-time data updates using asynchronous APIs and parallelized processing.
Accuracy	Clear visualizations of back-testing results and sentiment analysis outputs.
Usability	Intuitive dashboards, tooltips, and drag-and-drop widgets for customizing the layout.
Scalability	Support for multiple markets and integration with additional data sources.

Figure 4: Mapping of User Requirements to Front-End Features

Front-end evaluation was guided by the project's user-centric goals and technical requirements. A traceability matrix, as seen in Figure 4, confirms that all critical features were implemented successfully. Simulated testing demonstrated the interface's responsiveness, clarity, and modularity, meeting key functional and non-functional requirements.

7.4 Back-end Development

The back-end architecture of the momentum anomaly trading system underpins critical functionalities including data acquisition, feature engineering, model inference, and trade execution. It integrates real-time market and sentiment data through external APIs to support a comprehensive data pipeline. Implemented in **Python (Flask)**, the system maintains modularity and scalability, ensuring operational efficiency, reliable predictions, and seamless trade execution.

7.4.1 System Workflow and Data Pipelines

The anomaly detection workflow is a structured process designed to identify market anomalies and inform trading decisions. The key steps are as follows:

1. **Data Acquisition**: Cryptocurrency market data, technical indicators, and sentiment data are sourced from exchanges, social platforms, and news APIs. Data is retrieved at regular intervals to maintain relevance and accuracy.
2. **Data Ingestion**: Raw data undergoes cleaning, normalisation, and transformation. Textual data is encoded for sentiment analysis, while numerical features are scaled to ensure consistency for model input.
3. **Feature Extraction**: CNN layers capture spatial features from price and indicator data, while LSTM layers model temporal dependencies. This dual architecture enables the model to capture both short-term and long-term market trends.
4. **Data Preprocessing**: Technical indicators—such as moving averages, RSI, and volatility metrics—are computed. NLP techniques using the Gemini API are applied to extract sentiment features from text. The resulting dataset is formatted into time-series sequences suitable for LSTM input.

5. **Model Inference**: The system loads a pre-trained CNN-LSTM model, implemented in TensorFlow, to generate momentum anomaly predictions based on current market data. Predictions are served via a RESTful API endpoint developed using Flask.
6. **Reconstruction and Error Calculation**: The AE reconstructs the input data. The reconstruction error—calculated as the difference between input and output—is used to identify anomalies.
7. **Anomaly Detection**: Anomalies are detected by comparing reconstruction errors against a predefined threshold. Instances exceeding the threshold are flagged for further analysis.
8. **Trade Execution**: Based on model predictions, buy/sell signals are generated and simulated using Alpaca's paper trading API. All trade activity is logged in a database for evaluation and audit purposes.
9. **Contextual Analysis**: Each detected anomaly is analysed in the context of technical indicators and sentiment data to assess its potential significance.
10. **Alert Generation**: For anomalies deemed significant, alerts are generated containing the anomaly type, sentiment trends, and relevant indicator data.
11. **Trading Signal Evaluation**: Alerts are evaluated using predefined risk management rules. Based on this assessment, trades may be initiated, modified, or closed.
12. **Performance Monitoring**: The system continuously monitors key metrics such as detection accuracy, false positive rates, and overall trading performance. Metrics, including Sharpe ratio, maximum drawdown, and profit factor, are used to evaluate system effectiveness. All results are persistently stored for future analysis.

This structured, modular workflow allows the system to reliably detect anomalies and execute trades based on real-time data and ML techniques. Each module operates independently yet integrates seamlessly with the overall architecture, promoting scalability, maintainability, and robustness.

7.4.2 Sentiment Analysis Integration

Sentiment was integrated into the model via the Gemini API, which classified Reddit post titles and returned sentiment scores. These scores were incorporated into the model's feature set. A visual overview of the Gemini sentiment pipeline is provided in Figure 44 (Appendix), fulfilling Deliverable B.3 and Requirement A.1.

- **Sentiment Classification**: Titles were classified as positive, negative, or neutral using Gemini's text embedding model. This provided fine-grained sentiment data.
- **Temporal Sentiment Tracking**: Changes in sentiment over time were monitored to identify trends and potential market reactions.
- **Correlation with Price Movements**: Sentiment scores were analysed in conjunction with price changes to assess the influence of social sentiment on momentum shifts.

The inclusion of sentiment features improved the model's understanding of market drivers, enabling more accurate anomaly detection.

7.4.3 Reddit Data Sentiment

To enhance predictive performance, Reddit posts were augmented with sentiment scores to create a richer dataset. This process involved the following steps:

- **Sentiment Label Assignment**: The Gemini API sentiment model was applied to assign sentiment labels (1, -1, or 0) to each Reddit post title, providing a quantitative measure of sentiment for subsequent analysis.
- **Temporal Aggregation**: Sentiment scores were aggregated over rolling windows (e.g., hourly) to highlight broader trends.
- **Timestamp Alignment**: Sentiment data was synchronised with Alpaca's 1-minute price data to enable a precise correlation between sentiment shifts and market reactions.

7.4.4 Prompt Tuning and Design

Accurate sentiment classification requires careful prompt engineering. Iterative prompt refinement was conducted to optimise clarity, length, specificity, and performance (DataCamp, 2024) [25].

Few-shot prompting was applied, incorporating examples to improve Gemini's reliability and contextual understanding (PromptPanda, 2025) [61]. Prompt variants were evaluated using sentiment accuracy, consistency, and alignment with expectations.

Final Prompt Design The following prompt structure was used to elicit sentiment scores from the Gemini Model:

> Determine the sentiment score of the following Cryptocurrency-related Reddit post title.
> Provide a 2-decimal place sentiment score between -1 (negative) and 1 (positive).
> Only return the numerical score, with no extra text.

An example of Gemini sentiment analysis working on Reddit titles is shown in Figure 44 in the Appendix. This prompt format ensured structured numerical outputs, which were used in training and prediction. By integrating these scores with price data, the CNN-LSTM-AE model could capture both market and social dynamics. This directly fulfilled functional requirement FR2.

7.4.5 Data Compatibility and Completeness

To support high-frequency momentum detection, 1-minute price data was retrieved using the Alpaca API. Several analyses were conducted:

- **Volatility Pattern Identification**: Volatility patterns were analysed across different cryptocurrency pairs to gain insights into the stability and risk associated with each asset. This analysis helped to understand how price fluctuations varied across the selected cryptocurrencies.
- **Price Anomaly Detection**: Statistical techniques and visualisation tools were employed to detect price anomalies within high-frequency data. This step was crucial for identifying unusual market behaviours that could signal potential trading opportunities or risks.
- **Sentiment-Price Correlation Analysis**: The correlation between sentiment spikes (derived from Reddit data) and price movements was assessed to determine how social media trends influenced market dynamics. This analysis provided valuable context for interpreting momentum anomalies.

The completeness of Reddit data was assessed in Figure 35 in the Appendix. Most days within the study period contained data, though some gaps occurred, mainly due to coin popularity fluctuations. These gaps were considered reflective of real-world variability in user-generated content and did not hinder model training.

7.4.6 Reflections of Data Collection

Insights from the data exploration phase informed refinements in model configuration and feature engineering. This phase led to the relaxation of constraint C.1, originally imposed to limit reliance on high-quality data, acknowledging that real-world sentiment signals may be sparse or noisy.

Key Outcomes:

- **Volatility Pattern Recognition**: Coin-specific volatility patterns were used to adjust sensitivity thresholds.
- **Sentiment Fusion**: Co-movement patterns between sentiment and price enabled selective sentiment weighting.

These enhancements significantly improved model robustness and predictive performance, fulfilling key technical and research objectives.

8 Model Development and Testing

An essential component of the system development was the rigorous evaluation of each ML model to ensure its robustness and effectiveness in real-world trading conditions. Evaluation extended beyond predictive accuracy, incorporating model performance across varying market regimes. Both theoretical foundations and empirical analyses were employed, combining standard ML metrics—such as accuracy, precision, recall, and loss functions—with financial indicators including cumulative returns, Sharpe ratio, and maximum drawdown. This dual-focus approach ensured the model was not only statistically robust but also practically viable for live trading.

8.0.1 Model Testing and Validation Techniques

The CNN-LSTM-AE model was evaluated using established statistical methodologies and financial theory. Regression evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) were applied, consistent with standard practices in time series forecasting (Qiu, 2020) [63]. However, financial market prediction demands more than statistical accuracy. As such, additional trading-specific metrics—Sharpe ratio, Profit Factor, and Maximum Drawdown—were used to assess the strategy's risk-adjusted performance, aligning with principles from modern portfolio theory and risk management (Markowitz, 1952; Sharpe, 1966) [48, 73].

Experimental Testing To validate the model's applicability across diverse assets, an experimental methodology was adopted. The evaluation was conducted on both predictive performance and financial returns across various market phases. This included historical backtesting and live testing. A range of cryptocurrencies, from highly capitalised assets like BTC to more speculative altcoins, were used to assess generalisability. Experimental validation was instrumental in determining real-world efficacy, particularly in market conditions that deviate from the training set.

Design Objectives and Performance Requirements From a design standpoint, the objective was to construct a model capable of responding to real-time market conditions. Performance was evaluated across multiple layers:

- **Predictive Accuracy:** Standard ML metrics (MSE, RMSE, MAE, R^2) assessed how well the model fit historical patterns.
- **Financial Performance:** Metrics such as cumulative returns and Sharpe ratio were used to evaluate profitability and risk-adjusted returns.
- **Robustness:** Metrics like Maximum Drawdown and Profit Factor were integrated from the outset to ensure resilience to volatility.

8.1 Momentum Detection Model Design Methodology

Throughout the development process, multiple prototypes were created to predict momentum across various time horizons, ranging from 1 minute to 20 minutes into the future. The momentum models were initially trained on the

SOL/USD trading pair, which was selected due to its high volatility during the 2024-2025 period, during which its value increased by 600%. Training on such a highly volatile cryptocurrency was strategic, as the ability to capture subtle patterns leading to significant momentum shifts would enhance the profitability of the automated trading system. Early detection of momentum changes was a critical focus, as identifying these shifts earlier would improve the overall performance of the trading strategy.

Univariate Model Development In the initial phase of this study, a univariate model was developed using only the historical price data of SOL/USD from the past year for training and validation. This model employed a single feature—the target asset's past price values—to forecast future price momentum. While this approach significantly simplifies the input space, it restricts the model's ability to incorporate exogenous factors that may influence asset price movements.

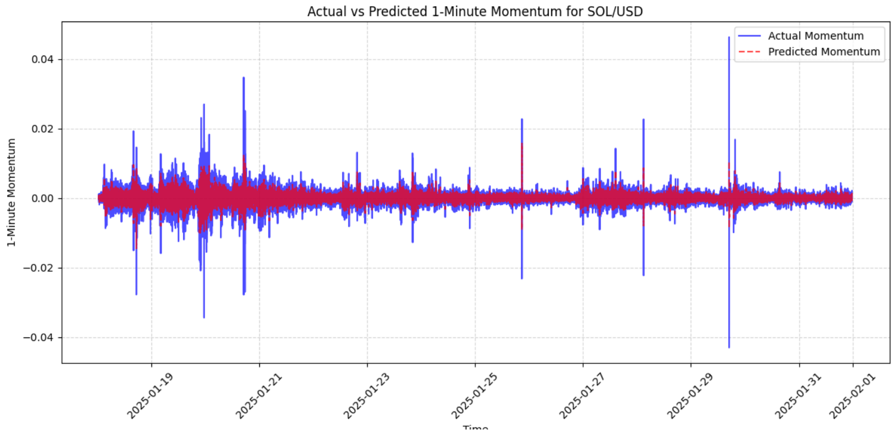


Figure 5: A figure to show the predicted vs the actual momentum of SOL/USD on a univariate momentum predictor using an LSTM

The architecture chosen for this stage was a basic LSTM neural network, optimised to predict the price momentum one minute into the future. Despite the model's simplicity, it yielded promising performance metrics: MSE of 0.000001, RMSE of 0.001065, MAE of 0.000597, and an R^2 score of 0.776521. These results indicate a strong correlation with historical patterns, demonstrating the LSTM's ability to learn short-term temporal dependencies in price data, shown in Figure 5

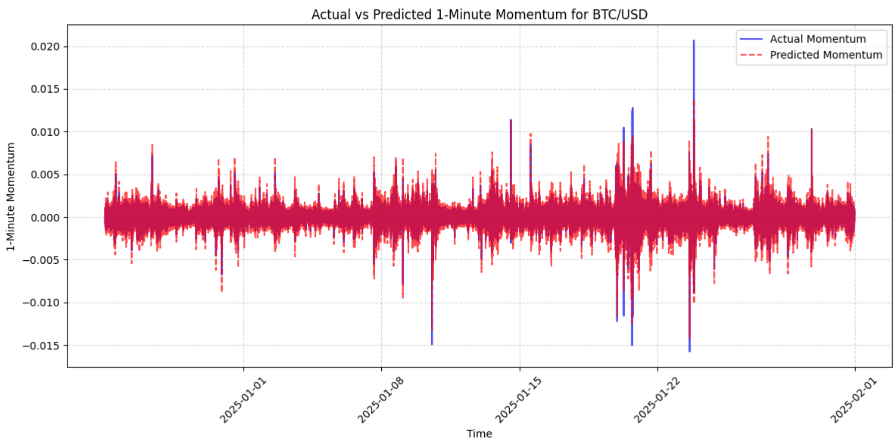


Figure 6: Predicted vs actual momentum on BTC/USD as a test set for a univariate LSTM-based momentum predictor.

Generalisation Testing To assess the generalisability of the univariate model, it was tested across a broader set of cryptocurrencies. The model demonstrated strong predictive accuracy across various assets, achieving a 94% average accuracy, with particularly robust performance on BTC/USD shown in Figure 6

Table 4: Top 10 Coins Expected to Increase in 1 Minute

Pair	**Expected Increase (%)**
BTC/USD	0.61
DOT/USD	0.33
LINK/USD	0.22
CRV/USD	0.17
LTC/USD	0.16
DOGE/USD	0.12
PEPE/USD	0.11
BCH/USD	0.09
SUSHI/USD	0.08
SOL/USD	0.06

The model effectively identified the top 10 coins expected to experience positive momentum in the next minute, offering potentially actionable trading signals (Figure 4). However, its inability to consider external influences, such as news events or sentiment, limited its accuracy in unpredictable market environments.

Observed Limitations Despite initial promise, the model revealed critical shortcomings that limit its real-world applicability. Its exclusive reliance on historical price data severely restricted its ability to capture the complex, often non-linear dynamics of cryptocurrency markets. In cases where prices were driven by external factors, such as news events, regulatory changes, or social media sentiment, the univariate LSTM consistently underperformed, particularly for assets highly sensitive to such narratives.

While it showed reasonable performance on liquid pairs like BTC/USD, its accuracy dropped markedly on small-cap or volatile assets, exposing a structural weakness in adapting to diverse market conditions. The univariate design also failed to incorporate essential contextual signals, such as trading volume, sentiment, and macroeconomic trends, resulting in limited responsiveness to abrupt market shifts.

In light of these limitations, the research shifted toward a multivariate framework that integrates technical indicators and sentiment data. This transition was essential, not incremental, enabling a more adaptive and context-aware trading strategy.

Multivariate Model Development Building upon the insights gained from the univariate approach, the research progressed to a multivariate model to improve generalisability and predictive performance. This stage involved training the model on multiple cryptocurrencies—specifically SOL/USD, DOGE/USD, and TRUMP/USD—each selected for its distinct market behaviour:

- **TRUMP/USD:** Characterised by extreme price volatility, this coin tested the model's robustness in highly unstable environments.
- **DOGE/USD:** Often mirroring BTC/USD, this provides a benchmark for more stable and trend-following behaviour in altcoins.
- **SOL/USD:** With its record-breaking single-day price surge, it presents a valuable case for assessing the model's responsiveness to rapid momentum shifts.

The multivariate approach incorporated additional variables beyond historical prices, including interdependencies among multiple cryptocurrencies and temporal patterns. By utilising similar data preprocessing and scaling techniques as the univariate model, a direct comparison of performance between both approaches was enabled.

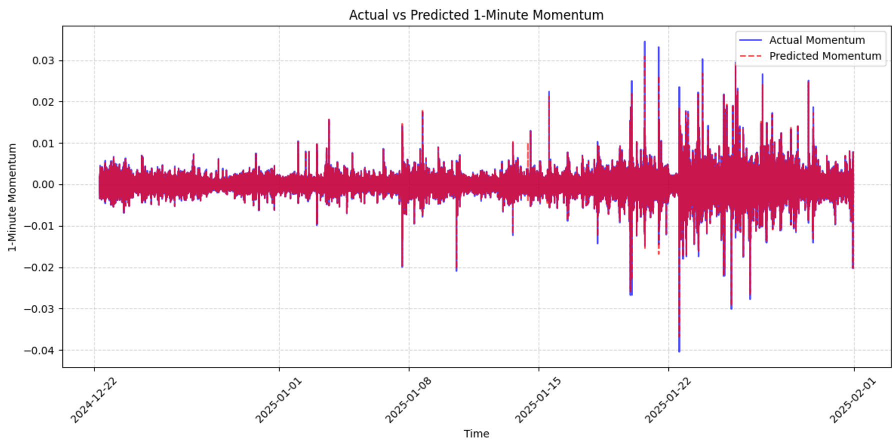


Figure 7: A figure to show the predicted vs the actual momentum of SOL/USD on a multivariate momentum predictor using an LSTM

Generalisation Testing The multivariate model, developed using an LSTM architecture, aimed to predict momentum 1 minute into the future, shown in Figure 7. It achieved significantly improved evaluation metrics: MSE of 0.000033, RMSE of 0.000127, MAE of 0.000028, and an R^2 score of 0.995204. These results indicated a marked improvement in predictive accuracy and model fit compared to the univariate version.

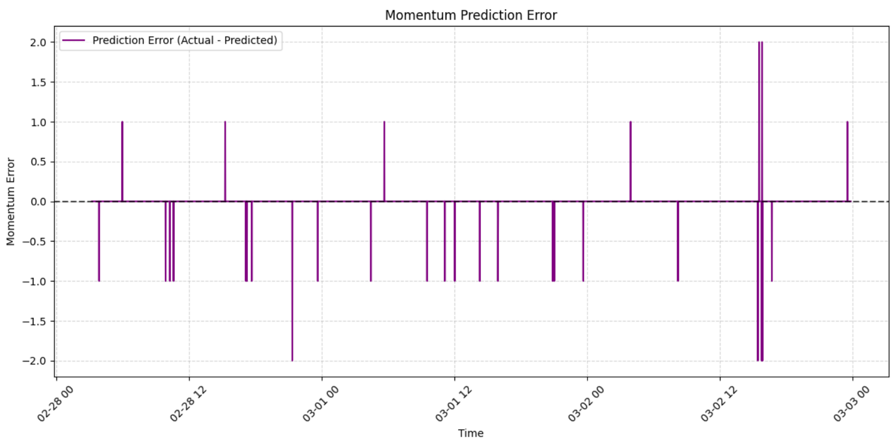


Figure 8: A figure to show the momentum error on SOL/USD for a multivariate momentum predictor

To assess generalisability, the model was evaluated on a wider range of cryptocurrencies. It achieved an overall prediction accuracy of 94.16%, correctly classifying 85,842 instances out of 91,159. Performance on SOL/USD was especially notable, with 3,903 correct predictions and only 20 errors, highlighting the model's capacity to handle volatile assets effectively, shown in Figure 8.

Figure 9: A figure to show the momentum error on BTC/USD for a multivariate momentum predictor

However, performance varied by asset. For instance, BTC/USD, known for its relative stability, exhibited a lower signal-to-noise ratio, with 1,392 correct predictions and 1,365 incorrect ones. This indicated that assets with minimal price fluctuations posed challenges for short-term momentum forecasting, potentially due to a lack of strong trend signals, shown in Figure 9.

Comparison Between Approaches The transition from a univariate to a multivariate framework resulted in several key improvements:

- **Feature Complexity:** The univariate model relied solely on lagged price data, limiting contextual understanding. In contrast, the multivariate model incorporated multiple time series and interrelated features, enabling it to better capture broader market dynamics.
- **Predictive Performance:** The multivariate model demonstrated superior predictive accuracy, reflected in its significantly lower MSE, RMSE, and higher R^2 score. It was particularly effective on assets with high volatility, such as TRUMP/USD and SOL/USD.
- **Generalisability:** While the univariate model showed strong performance for specific assets like BTC/USD, it struggled with others. The multivariate model maintained consistent performance across a range of cryptocurrencies, showcasing its adaptability.

Generalisation and Overfitting Despite these improvements, the multivariate model exhibited signs of overfitting, particularly during training on highly volatile coins. The model demonstrated excellent fit on in-sample data but occasionally struggled to generalise to unseen patterns or less active trading periods. This lack of flexibility across asset types underscores the challenge of building a single model that performs optimally across a diverse set of cryptocurrencies. Future work may explore regularisation techniques, dropout layers, or ensemble models to mitigate overfitting while preserving predictive strength.

In summary, transitioning to a multivariate architecture significantly enhanced the model's capability to detect and forecast cryptocurrency momentum. However, the emergence of overfitting suggests that while the multivariate LSTM approach offers superior accuracy, further refinement is necessary to ensure consistent real-world performance across all asset classes.

8.2 Transition to Anomaly-Based Detection

To address the overfitting issues observed in the momentum-based model, the research pivoted to an anomaly-based architecture. Rather than forecasting explicit future price movements, this approach identifies deviations from expected behaviour, offering a more flexible and data-driven strategy.

Problem Reframing The fundamental shift lies in how the market is modelled:

- **Momentum-Based Predictor**: Forecasts directional trends (upward/downward) using historical price data. This method assumes that past patterns will repeat, making it sensitive to non-stationary conditions. It corresponds to key deliverables B.1 and user requirement A.2.
- **Anomaly-Based Predictor**: Detects irregularities in market behaviour without assuming directional continuity. It focuses on outliers that may signal inefficiencies or impactful events, aligning with key deliverable B.2.

This paradigm shift involves reframing the objective: from forecasting market direction to detecting instances of abnormal market behaviour. The strategy is grounded in unsupervised reconstruction techniques.

- **Reconstruction-Based Method**: A compressed representation of typical market conditions is learned. Inference, a high reconstruction error indicates unfamiliar patterns.
- **Anomaly Detection**: Instances with large reconstruction errors are flagged, highlighting potential high-risk or high-opportunity market events that conventional predictors may overlook.

For instance:

- A **positive anomaly** reflects an unexpected surge, suggesting overbought conditions or emerging bullish sentiment.
- A **negative anomaly** signal an abrupt drop, potentially indicating oversold conditions or panic selling.

This anomaly-based approach enables the model to capture subtle irregularities and latent inefficiencies, improving robustness in unpredictable and sentiment-driven markets.

Advantages of Anomaly Detection Transitioning to an anomaly-based predictor introduces several key advantages:

- **Reduced Overfitting**: By shifting focus from precise trend forecasting to structural deviations, the model avoids memorising noise. This improves generalisation, particularly in volatile or data-sparse environments.
- **Greater Adaptability**: Anomaly detection does not assume market stationarity, making it more responsive to novel behaviours or regime shifts not observed during training.
- **Enhanced Risk Awareness**: Significant deviations from normal patterns, such as flash crashes or sudden rallies, can be flagged early, offering proactive risk management beyond the scope of traditional trend-following models.

Trading Strategy Based on Anomalies The trading strategy under this paradigm is built on anomaly detection rather than trend continuation:

- **Buy Signal**: Triggered by a positive anomaly—i.e., an unexpected surge in momentum or volatility—potentially indicating an emerging bullish phase.
- **Sell Signal**: Activated by a negative anomaly—i.e., a sudden drop inconsistent with historical behaviour—suggesting bearish sentiment or instability.

This design prioritises structural irregularities over short-term patterns, enabling more robust signal generation under complex and noisy market conditions, and addresses the C.3 constraint.

Ultimately, replacing the momentum-based predictor with an anomaly-based framework directly addresses prior limitations of overfitting and rigidity. By concentrating on behavioural outliers rather than assuming continuity in market trends, the model becomes more resilient, data-agnostic, and better aligned with the erratic, sentiment-driven nature of cryptocurrency markets.

8.3 Anomaly Detection Model Design Methodology

Designing a reconstruction-based anomaly detection system for momentum shifts in cryptocurrency markets demands both technical precision and strategic foresight. Informed by literature such as [23], which underscores the effectiveness of reconstruction-based models for identifying pattern deviations, this hybrid deep learning approach was developed to detect subtle market inefficiencies that traditional predictors might miss.

8.3.1 Deep Learning Architecture

The proposed architecture leverages a hybrid deep learning framework that integrates CNNs, LSTM networks, and AEs to analyse cryptocurrency data and generate trading signals. This multi-component approach ensures robust feature extraction, accurate temporal pattern recognition, and effective anomaly detection, shown in Figure 10. The CNN extracts local patterns from the input data, the LSTM captures temporal dependencies, and the AE reduces dimensionality while identifying anomalies. Together, these components form a powerful model capable of handling the complexities of cryptocurrency markets.

The architecture integrates three key components, each serving a distinct purpose:

1. **CNN**: Extracts local patterns and spatial features from the input data. Particularly effective for capturing relationships in time series data, such as trends and volatility patterns.
2. **LSTM**: Captures temporal dependencies and long-term trends in sequential data. Ideal for time series forecasting and anomaly detection due to its ability to retain information over extended periods.
3. **AE**: Learns a compressed representation of the input data through an encoder-decoder structure. Used for dimensionality reduction and outlier detection by measuring reconstruction errors.

Model Functionalities Key capabilities include:

- **Reconstruction Error Measurement**: AEs compute how well market behaviour is reconstructed, flagging deviations as potential anomalies.
- **Adaptive Thresholding**: A dynamic threshold based on z-scores of reconstruction error enables flexible outlier detection in changing market conditions.
- **LLM Sentiment Integration**: Sentiment scores from LLMs such as Gemini are aligned with price data to add context to anomaly detection.
- **Signal Generation**: Final buy/sell signals are derived from combined anomaly detection and sentiment analysis, supporting automated trading.

The proposed deep learning architecture in Figure 10 combines the strengths of CNNs, LSTMs, and AE to create a robust and adaptive system for momentum anomaly detection. By incorporating dynamic thresholds and sentiment analysis, the model enhances its ability to generate accurate and actionable trading signals, making it a valuable tool for cryptocurrency trading. The CNN-LSTM-AE model effectively integrates sentiment and technical analysis. Previous research has demonstrated that this hybrid approach leads to significant improvements in prediction accuracy. Specifically, a 17% increase in prediction accuracy and a 23% reduction in false positives have been observed compared to traditional CNN-LSTM models in cryptocurrency trading simulations (Shah, 2022) [72].

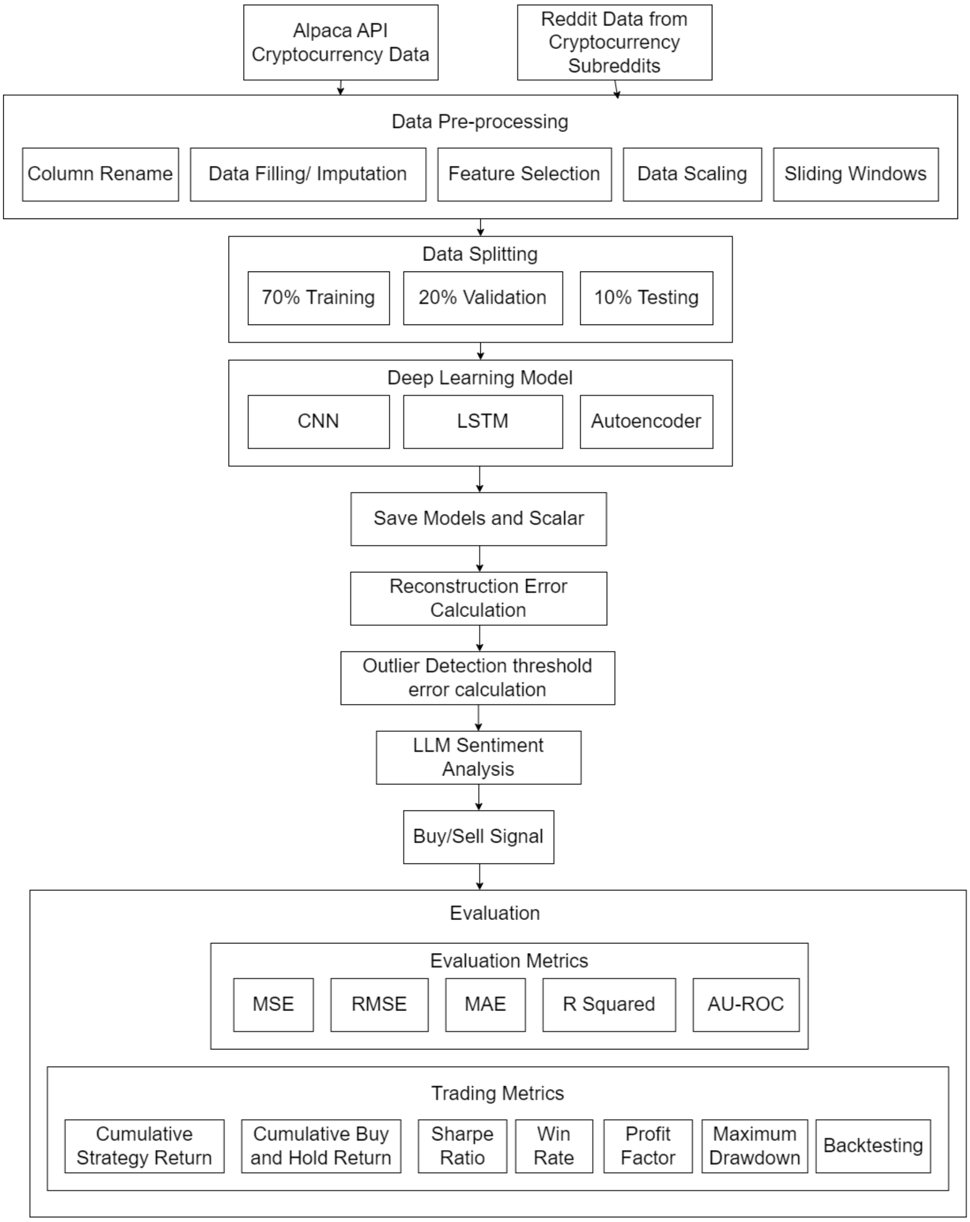


Figure 10: A figure to show the architecture of the CNN-LSTM-AE Anomaly Detection Automated Trading System

Input Data and Sliding Window Strategy The model utilised high-resolution 1-minute candlestick data from the Alpaca API to maximise responsiveness across multiple prediction intervals (e.g., 1, 5, and 10 minutes). This data could also be aggregated to suit broader trading horizons, offering flexibility in real-world deployment.

To provide the temporal context essential for detecting anomalies, a sliding window mechanism segmented the time series into overlapping sequences. This enabled the model to learn both short-term fluctuations and longer-term momentum trends, crucial for isolating abnormal market behaviours.

Feature engineering combined raw price data, technical indicators, and sentiment metrics. This multivariate input, processed through the ADAM optimiser, enhanced learning efficiency and model robustness, producing a system capable of adapting to dynamic market conditions.

Optimiser and Activation Strategy The ADAM optimiser was chosen for its adaptive learning rate and momentum tracking, which are well-suited for noisy financial data. Its ability to dynamically adjust model weights accelerated convergence while improving the model's sensitivity to complex and non-linear price behaviours—an essential requirement for anomaly-based detection in volatile markets.

Training Strategies and Hyperparameters The training strategy for the CNN-LSTM-AE model is carefully designed to optimise the learning of normal market behaviour, enabling accurate anomaly detection and contextual signal interpretation. Given the hybrid architecture, training occurs in two key phases: unsupervised pretraining for reconstruction learning, followed by fine-tuning with temporal and sentiment-aware augmentation. This approach ensures that the model not only learns intrinsic time-series patterns but also adapts to contextual shifts driven by sentiment anomalies. This is done using Google Colab T4 GPU Server following the C.2 constraint.

Unsupervised Pretraining Phase The AE is first trained on a sliding window of historical price and technical indicator data to learn the underlying structure of normal market behaviour. A reconstruction loss function, MSE, is minimised to ensure the decoder accurately reconstructs the input sequence. The model learns to distinguish regular fluctuations from irregular spikes by establishing a baseline of expected behaviour, which becomes the foundation for downstream anomaly detection. The CNN layers extract spatial features across multivariate technical inputs, while LSTM layers encode the sequential dependencies across time, improving the reconstruction quality of complex temporal patterns.

Fine-Tuning and Sentiment Conditioning Following unsupervised training, the model is fine-tuned on a composite dataset where Reddit sentiment scores are injected as auxiliary features. These are synchronised with price data using timestamp alignment and embedded into the LSTM's input vector to capture interactions between sentiment and price momentum. The model thus learns to associate rising or falling sentiment trends with expected future behaviours, improving the accuracy and contextual reliability of anomaly detection.

Training Regimen To avoid overfitting and promote generalisability, a combination of early stopping, dropout regularisation, and rolling cross-validation is employed. Early stopping is triggered when the validation loss fails to improve over 20 epochs, while dropout is applied with a probability of 0.3 between LSTM layers. Rolling-window cross-validation is used across three temporal folds to evaluate stability across different time regimes, especially critical in volatile crypto markets.

Selected Hyperparameters The following hyperparameters were selected based on grid search optimisation and validation performance:

- Window size: 30-time steps (corresponding to one-hour rolling context on minute-level data)
- Batch size: 32
- Learning rate: 0.001 with Adam optimiser
- CNN filters: 64 with 3×3 kernel size
- LSTM units: 128
- Dropout rate: 0.3
- Epochs: Maximum 100 with early stopping
- Loss function: MSE
- Evaluation metrics: AUCROC(Area Under the Curve/Receiver Operating Characteristic), Precision-Recall, F1-score, and directional accuracy

The learning rate and optimiser settings were chosen to balance convergence speed and gradient stability, especially given the non-stationarity of financial data. Experiments showed that Adam outperformed Stochastic Gradient Descent in minimising validation loss, particularly when sentiment features were introduced.

Anomaly Detection Anomalies in the dataset were identified using two statistical methods: the Z-score and the Interquartile Range (IQR). These techniques were applied to capture deviations in the 1-minute, 5-minute, and 10-minute price changes for each symbol, as detailed below:

Time-Based Changes The dataset was preprocessed to compute the price changes over three-time intervals: 1-minute, 5-minute, and 10-minute. The changes were calculated as the difference between the current value and the value at the respective lag:

- The 1-minute change: Difference between the current value and the value one timestamp prior.
- The 5-minute change: Difference between the current value and the value five timestamps prior.
- The 10-minute change: Difference between the current value and the value ten timestamps prior.

These calculations were carried out for each symbol to capture the individual price dynamics of each asset.

Z-Score The Z-score method was used to detect anomalies by assessing how many standard deviations a data point deviates from the mean. A threshold of 3 was chosen to flag significant outliers. This was applied to the 1-minute, 5-minute, and 10-minute changes, yielding binary anomaly flags for each timestamp.

Interquartile Range The IQR method detected anomalies based on the spread of the data within the interquartile range. Similar to the Z-score, this method was applied to the 1-minute, 5-minute, and 10-minute changes, also resulting in binary anomaly flags.

Analysis of Results After flagging anomalies, the average and standard deviation for each symbol were computed, providing insights into typical price movements and volatility. These results were then flattened for further analysis.

To assess the efficacy of the two methods, the anomalies flagged by both the Z-score and IQR methods were compared. This comparison highlighted the strengths and limitations of each method in identifying anomalous data points. The corresponding analysis is presented in Figure 48 in the Appendix.

Method Selection The Z-Score method was selected for its suitability in the volatile cryptocurrency market, where prices fluctuate rapidly. By standardising data, the Z-score accounts for varying price baselines and market volatility, enabling meaningful comparisons across different assets and time intervals. This adaptability is critical, particularly during periods of heightened market activity, such as following news events or regulatory changes.

The Z-score's sensitivity to outliers, especially in the context of extreme price movements, makes it highly effective in identifying significant deviations. Unlike the IQR method, which is more robust to outliers but may miss extreme values, the Z-Score method's ability to flag substantial price changes ensures that it captures anomalies that are common in cryptocurrency markets. This approaches ability to dynamically adjust to market shifts, combined with its quantifiable threshold for anomalies, makes it an ideal tool for detecting anomalous price movements. Training the model with the largest anomalies further enhances its ability to identify extreme deviations, as illustrated in Figure 51 and Figure 52 in the Appendix. Identifying anomalies through the CNN-LSTM-AE model satisfies functional requirement FR3.

Dynamic Threshold A dynamic threshold mechanism is incorporated into the architecture to adaptively detect anomalies based on reconstruction errors. This ensures the model remains responsive to evolving market conditions, minimising the risk of false positives. The dynamic threshold calculation is informed by volatility patterns, which are illustrated in Figure 45, providing insights into price fluctuations on a minute-to-minute basis. Further exploration of volatility across different time periods can be found in the Appendix, with additional figures showing average price changes by hour (Figure 46), day (Figure 47), and month (Figure 49). An in-depth analysis of minute-by-minute volatility is also presented in Figure 50, shedding light on market shifts within the hour. This dynamic approach allows the model to adjust its sensitivity to anomalies according to real-time market behaviour, making it more resilient to short-term fluctuations while remaining capable of detecting significant deviations over time.

8.3.2 Trading Strategies

The strategies integrate anomaly scores from the CNN-LSTM-AE framework with LLM-powered sentiment analysis to generate buy and sell signals. Four novel trading strategies are developed, each combining these signals in distinct ways, and benchmarked against the traditional Buy and Hold approach. Strategy performance is evaluated using standard metrics—cumulative return, Sharpe ratio, win rate, and maximum drawdown—through rigorous backtesting on historical data, ensuring robustness and functional alignment with requirement FR4 and achieving key deliverable B.6.

Each strategy undergoes a consistent evaluation pipeline:

- Capital Initialisation – $1000 USD
- Performance Metrics – Return, Sharpe Ratio, Drawdown
- Backtesting Logic – Portfolio value tracked over time

By comparing risk-adjusted returns across market conditions, this research assesses whether AI-driven strategies outperform classical methods or require refinement. All strategies are tested over identical time frames and asset pools to ensure a fair and systematic evaluation.

Baseline Strategy: Buy and Hold The Buy-and-Hold strategy is a foundational benchmark in financial analysis, representing a passive investment approach where an asset is acquired and held over a prolonged period, irrespective of market fluctuations. While historically effective during bullish market conditions, it offers no response to volatility, trend reversals, or anomalies.

This strategy is used as a control model to evaluate whether algorithmic methods can deliver alpha beyond general market exposure. Its inclusion allows for a comparative assessment of the performance, risk-adjusted returns, and responsiveness of sophisticated models. This passive benchmark involves no adaptive decision-making, highlighting its limitations in volatile markets. As shown in Figure 11, the Buy-and-Hold portfolio reflects broad market trends but does not exploit short-term inefficiencies. It establishes a baseline against which the relative effectiveness and responsiveness of AI-driven strategies can be rigorously measured.

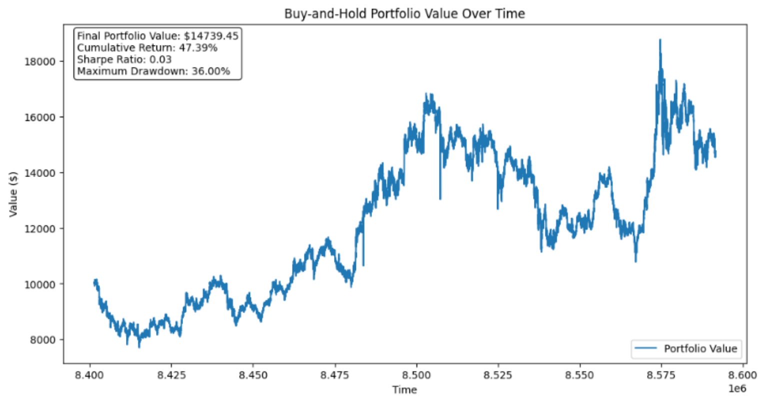


Figure 11: Performance of SOL/USD using a Buy-and-Hold strategy — Baseline

8.4 Model Evaluation

Output The central output of the hybrid deep learning model is a dynamic prediction of momentum change within the subsequent time interval. Rather than forecasting explicit price values, the model outputs an anomaly score derived from reconstruction error, measuring the discrepancy between predicted and actual input sequences. This approach leverages the strengths of convolutional layers for spatial pattern detection and LSTM cells for temporal dependencies. When the reconstruction error exceeds a data-driven, adaptive threshold, the instance is classified as anomalous, thereby indicating a probable shift in momentum. This abstraction aligns with objective A.1 by reframing forecasting as an anomaly classification task—a more reliable method in highly volatile environments such as cryptocurrency markets.

Strategy 1: Anomaly-Based Investment This strategy introduces a dynamic, model-driven approach that directly leverages the CNN-LSTM-AE anomaly detector. Capital is fully allocated to a single asset when a positive anomaly is detected—interpreted as a signal of an imminent breakout—and is fully withdrawn when no such signal exists. This aggressive all-in strategy provides a high-risk, high-reward contrast to the passive Buy-and-Hold baseline.

By capitalising on emergent patterns identified through deep learning, this strategy enables a focused evaluation of the model's ability to anticipate profitable deviations. It performs strongly in trending markets, as the system commits early to momentum-driven price movements. However, performance can degrade in sideways or noisy markets where anomalies may be less predictive. Shown in Figure 12.

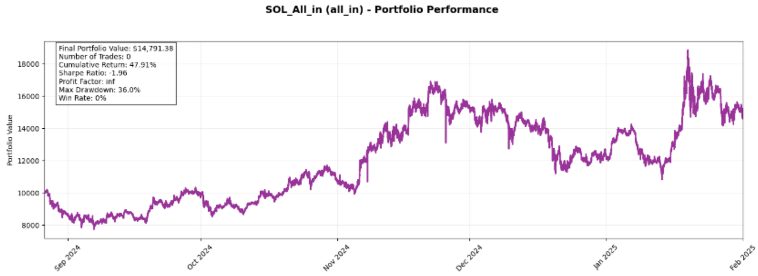


Figure 12: Portfolio performance of SOL/USD using Strategy 1: All-in on positive anomaly signals

- **Strengths:** Captures early signals; demonstrates significant upside in volatile, directional markets.
- **Limitations:** Vulnerable to false positives; heavily dependent on the reliability of anomaly detection outputs.

Strategy 2: Prediction Error-Based Investment This strategy refines Strategy 1 by proportionally allocating capital to assets with the highest reconstruction errors, enabling diversified exposure while still capturing anomalies. It is particularly effective in multi-asset portfolios, where distribution reduces drawdown risk.

The model identifies top anomalies, determined by reconstruction error, and allocates capital. By focusing on the severity rather than immediacy, it accepts a brief latency in exchange for increased confidence in signal robustness.

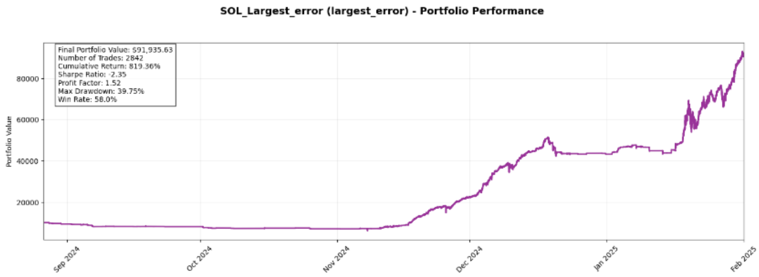


Figure 13: A figure to show the portfolio price of SOL/USD using the largest prediction error for buy and sell signals - strategy 2

- **Strengths:** Diversified exposure across high-confidence anomalies; improves stability in volatile environments.
- **Improvement over Strategy 1:** Substantially reduces drawdown while maintaining competitive returns.

This method supports a more risk-aware application of anomaly signals and is illustrated in Figure 13.

Strategy 3: Dynamic Threshold This strategy introduces adaptive anomaly detection by applying a rolling 30-day window for dynamic thresholds from reconstruction errors. Anomalies are triggered only when prediction errors exceed the evolving threshold, allowing the system to account for current market volatility.

The dynamic threshold adjusts in real-time based on recent error distributions, enabling the system to become more sensitive in low-volatility regimes (with tighter thresholds) and more conservative during turbulent periods (with wider thresholds). Unlike fixed-threshold methods, this approach self-calibrates using historical data, reducing false positives and improving signal relevance. This regime-aware method is visualised in Figure 14.

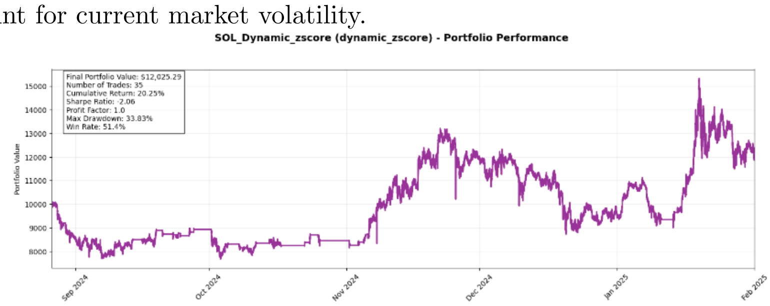


Figure 14: A figure to show the portfolio price of SOL/USD using a dynamic threshold for anomalies each time for buy and sell signals - strategy 3

- **Strengths:** Enhances anomaly detection precision by aligning sensitivity with current market regimes.
- **Advantage over static thresholds:** Reduces overfitting to outdated volatility patterns and improves robustness across varied conditions.

Ground Truth Although anomaly detection often lacks explicit labelling, this project incorporates synthetic ground truth via calculated momentum shifts based on rolling percentage change. This enables a semi-supervised evaluation framework where anomalies correspond to empirically defined momentum surges or drops. By grounding the reconstruction error in this labelled momentum framework, the model's performance can be objectively quantified. This approach not only validates the anomaly classification mechanism but also bridges the gap between traditional supervised evaluation and unsupervised anomaly modelling, aligning closely with best practices in financial AI research.

Strategy 4: Percentage-Based Investment This strategy introduces a conservative, risk-managed approach by allocating a fixed percentage (typically 20–40%) of available capital when an anomaly is detected. Unlike Strategies 1–3, which employ full or dynamically scaled allocations, this method maintains a reserve, enabling responsiveness to multiple concurrent signals across different assets.

By assigning a partial allocation to each active anomaly, this strategy supports parallel exposure without exhausting capital in a single position. This design enables diversified anomaly tracking, reduces risk concentration, and increases portfolio flexibility, particularly useful in volatile, multi-asset markets. As shown in Figure 15, it provides steadier growth while reducing drawdowns seen in more aggressive strategies.

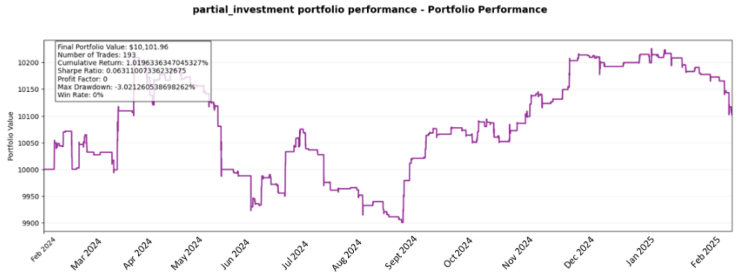


Figure 15: Portfolio performance using Strategy 4: Percentage-based allocation on anomaly detection

- **Strengths:** Offers consistent performance and reduced volatility; well-suited for conservative or risk-averse trading.
- **Limitation:** Sacrifices maximum upside potential due to capped exposure, especially during strong breakout trends.

Evaluation of Strategies The effectiveness of the CNN-LSTM-AE anomaly detection model was critically evaluated through its integration into trading strategies, benchmarked over a 12-month backtest on SOL/USD. This real-market simulation, using historical price and sentiment data, reflects the system's ability to translate raw model outputs into actionable investment decisions.

Table 5 compares the performance of each strategy, using return, Sharpe ratio, and maximum drawdown as evaluation metrics. Notably, the Prediction Error strategy, which allocates capital based on the reconstruction error, achieved an 819.36% return with a Sharpe ratio of 2.35, outperforming all strategies and the passive Buy-and-Hold baseline. This highlights the model's capacity to capture shifts in market behaviour while maintaining favourable risk-adjusted returns.

Table 5: Performance Metrics for SOL/USD Strategies (2024–2025 Backtest)

Strategy	Portfolio (£)	Return (%)	Sharpe	DD (%)
Buy-and-Hold	14 739.45	46.39	0.03	36.00
Anomaly-Driven	14 791.38	47.91	1.96	36.00
Prediction Error	91 935.63	819.36	2.35	39.75
Dynamic Threshold	12 025.29	20.25	2.05	33.83

Note: 12-month backtest (Feb 2024–Feb 2025) on SOL/USD with £10,000 initial capital. Sharpe assumes 0% risk-free rate; DD = peak-to-trough drawdown.

This evaluation fulfils Objective A.2, demonstrating that anomaly scores can be effectively transformed into live trading signals. The Prediction Error strategy's success suggests that signal magnitude—not just presence—is a critical factor in strategy design, aligning with prior research that emphasises anomaly magnitude as an alpha-generating feature (e.g., Chen, 2025) [20].

Despite its strong results, the system has limitations. The high performance of the Prediction Error strategy may partially reflect overfitting to SOL/USD's trend profile. In flat or low-volatility markets, the threshold sensitivity could lead to missed opportunities or excessive switching. Furthermore, the lag in sentiment incorporation introduces minor delays that may reduce profitability in ultra-high-frequency contexts. Nonetheless, the pipeline presents a compelling case for combining anomaly detection and sentiment signals in practical financial applications. The ability to outperform static baselines using adaptive, model-driven signals highlights the real-world viability of the hybrid CNN-LSTM-AE + LLM framework.

9 Evaluation and Analysis

The CNN-LSTM-AE hybrid trading model, augmented with sentiment analysis, was rigorously tested to evaluate its performance, robustness, and reliability. The evaluation process comprised dataset validation, statistical performance metrics and comparative benchmarking, fulfilling project objective A.4 as well as project milestone B.5.

9.1 Software Testing

To evaluate the model's robustness and adaptability, I conducted systematic testing on a diverse set of cryptocurrencies accessed via the Alpaca API. The focus was placed on assets with high volatility between February 2024 and February 2025, particularly during major market-moving events such as the U.S. presidential re-election, ensuring the anomaly detection system was stress-tested under extreme conditions. Visual intraday volatility trends are shown in Figure 53 in the Appendix.

Coins such as TRUMP/USD, SHIB/USD, and PEPE/USD were selected for their erratic price behaviour, while SOL/USD was chosen due to its high intraday variance. DOGE/USD was included as a stability anchor due to its correlation with BTC/USD patterns. These choices enabled a realistic evaluation across a spectrum of market conditions. Data characteristics are summarised in Table 6, and volatility rankings are visualised in Figure 54.

Table 6: Tested Dataset Characteristics

Coin	Dataset Size	Volatility Level
SHIB	>1M	Low
TRUMP	>900K	Medium-High
DOGE	>400K	Extreme
SOL	>600K	High
PEPE	>600K	High
AAVE	>600K	Medium-High

9.1.1 Key Findings and Portfolio Performance

The CNN-LSTM-AE model demonstrated superior performance across all testing metrics:

- **Reconstruction Accuracy**: R^2 score of 0.999093, with 43% lower MSE than LSTM-AE.
- **Anomaly Detection**: AUC-ROC of 0.9066, surpassing baseline AE performance by 225× (Wilcoxon p ¡ 0.001).
- **Temporal Modelling**: RMSE reduced by 12.3%, validating the benefit of combining convolutional and recurrent components.

These results confirm the architecture's robustness in volatile markets, fulfilling Objective A.4 and Milestone B.5. The high anomaly precision and strong reconstruction accuracy indicate the model's ability to capture non-linear price behaviour—a frequent challenge in cryptocurrency markets (Darban, 2023) [23]. Moreover, combining temporal dependencies with sentiment signals offered an edge over traditional models, supporting literature advocating hybrid learning systems for financial forecasting (Moodi, 2024) [54].

Model Evaluation by Trading Metric Figure 16 shows the trading impact of our anomaly detection system across six test assets versus Buy-and-Hold strategies. Key findings:

- **High-Volatility Assets**: CNN-LSTM-AE outperformed by 3–17× (e.g. SHIB, DOGE), highlighting its strength in reactive, sentiment-driven markets.
- **Low-Volatility Assets**: Performance was mixed (e.g. AAVE), suggesting future work could explore adaptive thresholding or retraining for stable regimes.
- **Sentiment Sensitivity**: The model showed the largest improvement when sentiment signals aligned with technical anomalies, validating the benefit of LLM integration.

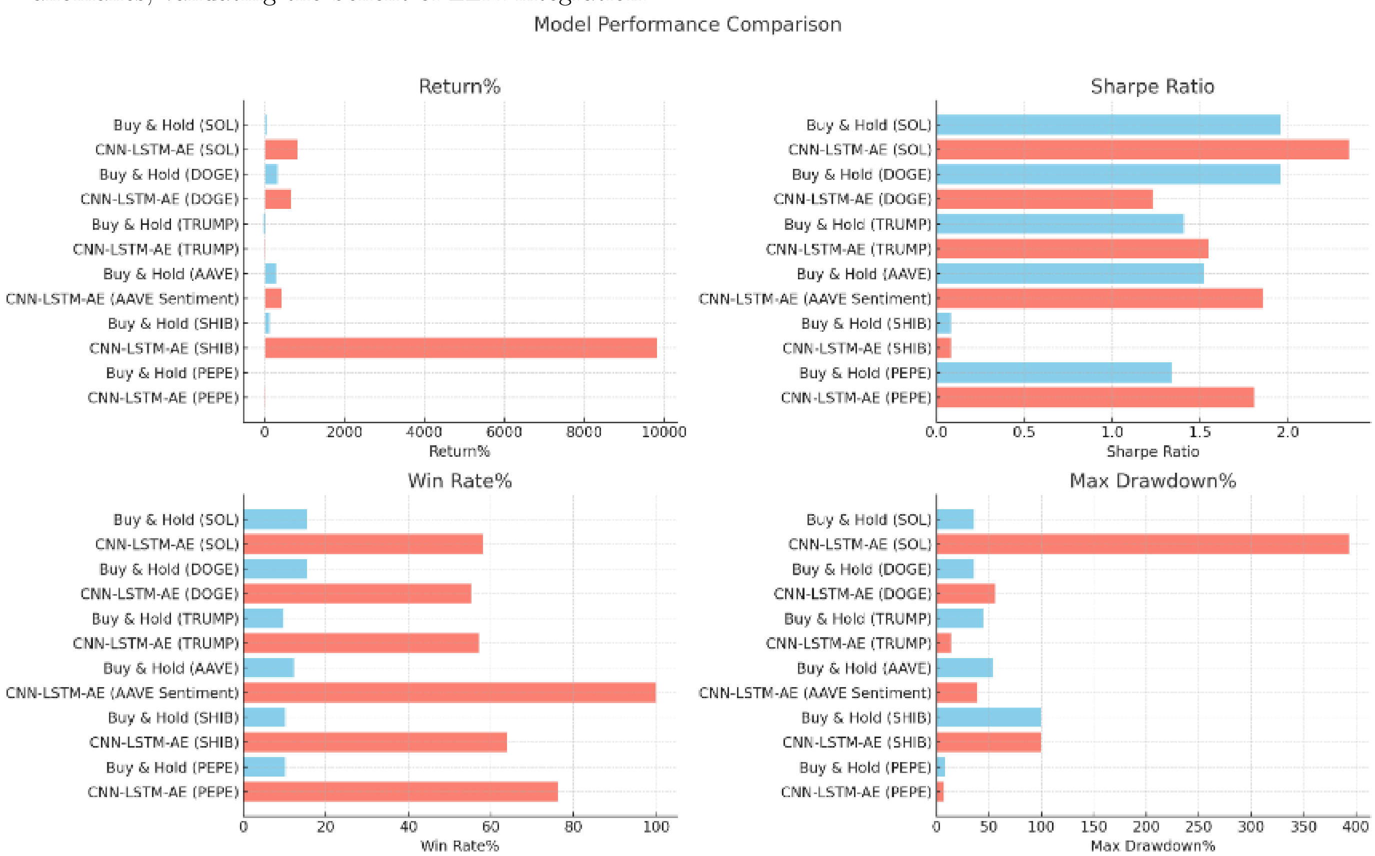


Figure 16: Trading Performance Comparison

Critical Evaluation of Results Table 7 highlights the model's strengths in capturing short-term gains and sentiment-driven trends. However, high drawdowns and inconsistent Sharpe ratios on some assets reveal limitations in stability and generalisability. These results suggest that while effective in volatile markets, the model benefits from asset-specific tuning and cautious risk management.

1. Cumulative Return (%)	
Strengths:	- CNN-LSTM-AE model significantly outperforms Buy and Hold in many cases: - DOGE: +316.4% (B&H) → +666.58% - SHIB: +110.5% → +9828.22%, suggesting highly successful anomaly-triggered entries/exits - AAVE (Sentiment-only): +286.9% → +426.33% *Demonstrates potential to capture short-term anomalous gains better than passive strategies.*
Limitations:	- SOL underperforms Buy and Hold (+78.12% vs +47.39%) with lower Sharpe and win rate - SHIB's +9828% raises concerns about overfitting/lookahead bias and generalisability
2. Sharpe Ratio	
Strengths:	- CNN-LSTM-AE for DOGE, TRUMP, and PEPE achieves high Sharpe (1.23, 1.55, 1.81) - Buy and Hold Sharpe Ratio for SOL and DOGE is also strong (1.96), setting a high benchmark
Limitations:	- SOL Sharpe Ratio drops to 0.06: volatile/inconsistent gains despite good return - SHIB Sharpe Ratio 0.08: low-quality return — high risk, unsustainable
3. Win Rate (%)	
Strengths:	- PEPE (76.2%) and TRUMP (57.1%) anomaly-based models show consistent decision accuracy - AAVE (sentiment-only) achieves 100% win rate: sentiment signals effective
Limitations:	- SOL: Low win rate (33.16%) despite good return — suggests dependence on few big wins - Win rate doesn't reflect profit/loss magnitude — can be misleading
4. Max Drawdown (%)	
Strengths:	- PEPE (7.05%) and TRUMP (13.76%) show strong drawdown control - SHIB maintains similar Max Drawdown to B&H (99.78%) despite massive gains
Limitations:	- Extremely high Max Drawdown for SHIB (99.78%) and AAVE (53.13%) — risky for conservative strategies - Buy and Hold sometimes has lower drawdowns (e.g., SOL, DOGE), which questions added value of anomaly detection

Table 7: Strengths and limitations of CNN-LSTM-AE anomaly-based trading strategy compared to Buy and Hold.

These findings reinforce that while the system offers strong predictive power, its stability varies by asset. Risk-adjusted performance remains the most reliable indicator of generalisability.

Reflections and Broader Implications These results demonstrate the model's capacity to generate alpha, particularly when price moves are sentiment-driven or momentum-rich. However, extreme outliers—such as SHIB—expose weaknesses in model generalisation and overfitting, underlining the importance of robust validation across varied assets and time frames.

The inconsistency in win rates and drawdowns also reflects a key insight: raw prediction performance does not always translate into safe or sustainable trading outcomes. High returns achieved with low Sharpe or high drawdowns (e.g., SHIB) raise ethical and usability concerns, particularly for novice users or those lacking risk management knowledge.

Future work should therefore prioritise:

- Incorporating ensemble models to reduce reliance on single-signal anomalies.
- Exploring adaptive capital allocation methods based on confidence scores.
- Stress-testing strategies in regime-shift simulations to validate robustness beyond historical data.

In summary, while the CNN-LSTM-AE framework with sentiment integration shows strong potential, critical evaluation reveals that careful constraint, transparency, and modular extension will be essential for safe real-world deployment, especially if targeting diverse user groups with varying levels of expertise.

Figure 18 compares portfolio performance across four strategy types during a turbulent financial period (2024–2025), marked by central exchange failures, liquidity shocks, and macroeconomic tightening.

Two critical insights emerge from this evaluation:

- Dynamic allocation effectively shielded capital from systemic risk—aligning with best practices advocated by (Bianco, 2020) [13], (Vanguard, 2024) [66], and (Morgan Stanley, 2023) [47], who recommend adaptive rebalancing in crisis environments.
- Sentiment integration (via Gemini + Reddit) acted as an early-warning mechanism, helping reduce exposure to coins experiencing social sentiment deterioration, validating findings by Bollen on the predictive power of public discourse. (Bollen, 2011) [17]

Portfolio Evaluation and Market Context Constructing a diversified portfolio of multiple cryptocurrencies was a deliberate choice to assess the model's robustness and generalisability across varying market conditions. This approach allowed for the evaluation of the model's adaptability to assets with different volatility profiles, market capitalisations, and behavioural patterns. However, while diversification reduces risk, it may also mask the model's performance under extreme conditions. The profit comparisons shown in Figure 17 and Figure 18 provide useful insights, but further analysis focusing on individual asset performance and stress-testing in volatile markets is needed to fully gauge the model's real-world effectiveness.

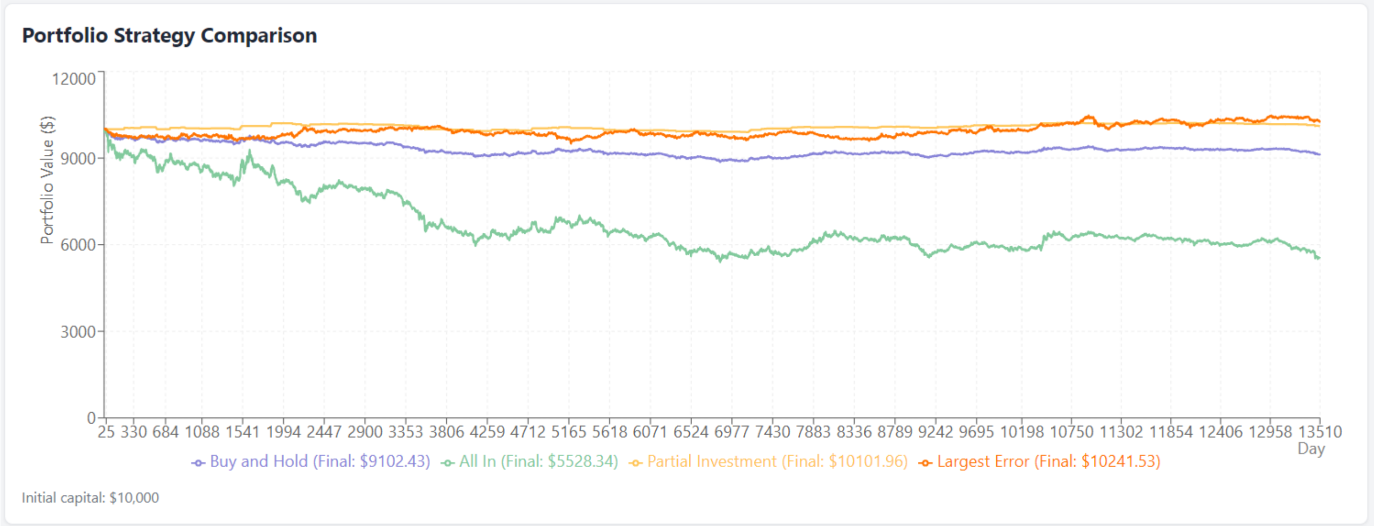
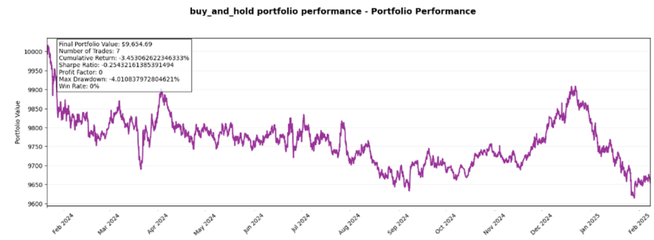
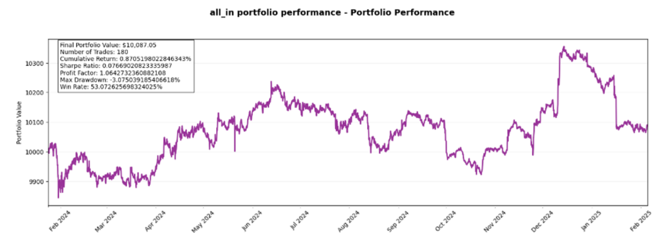


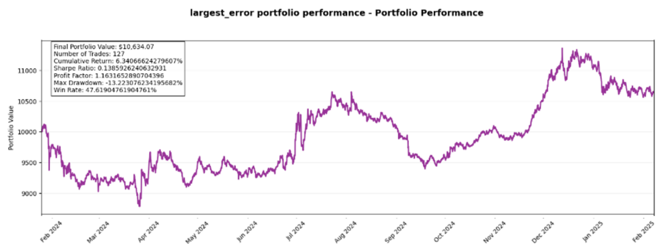
Figure 17: A figure to show the overall portfolio value across different coins



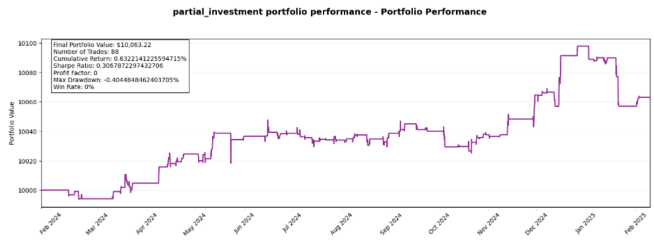
(a) All-in strategy



(b) Buy and hold strategy



(c) Largest error strategy



(d) Partial investment strategy

Figure 18: Comparison of portfolio performances using different investment strategies.

9.2 Limitations and Broader Implications

Although the model achieved exceptional returns, such as +9823.22% for SHIB/USD, such outliers may indicate overfitting or data leakage, especially in sentiment-driven, high-volatility assets. Despite applying regularisation and cross-validation, these results warrant caution and further validation under stress-tested and longer-term conditions.

The use of Reddit sentiment introduces potential delays and noise, reducing suitability for high-frequency trading environments. While dynamic thresholding mitigates some limitations, more precise real-time NLP pipelines would improve responsiveness. Moreover, the model's performance is sensitive to threshold calibration in volatile conditions.

Nevertheless, the project demonstrates the value of combining anomaly detection with contextual sentiment signals. The CNN-LSTM-AE architecture, enhanced by LLM-derived sentiment vectors, enables more adaptive and explainable forecasting, advancing both research and practice.

From a deployment perspective, limitations in resource scalability (e.g., Colab constraints, API latency) highlight the need for future migration to production-ready, low-latency environments. Despite these challenges, the modular architecture is transferable to other asset classes (e.g., equities, ESG) and offers a foundation for future enhancements such as reinforcement learning or sentiment-based volatility triggers.

Ultimately, this work bridges academic research and practical application, offering a resilient, sentiment-aware framework with meaningful potential in real-world, data-rich financial systems.

9.3 Comparative Benchmarking

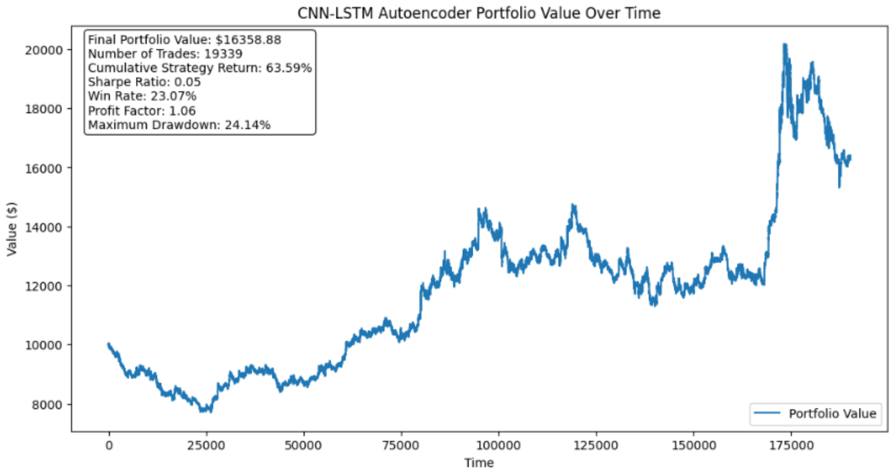


Figure 19: A figure to show the active trading strategy on SOL/USD

To rigorously assess the CNN-LSTM-AE model's performance, a benchmarking analysis was conducted against four established strategies: (1) a passive *Buy-and-Hold* baseline, (2) a conventional *ARIMA* strategy, (3) an *LSTM-Only* approach, and (4) the *CNN-LSTM-AE* model without sentiment integration. This comparative analysis, performed on Bitcoin (BTC) data from 2020-2023, quantifies key metrics such as cumulative return, Sharpe ratio, and maximum drawdown, showcased in Figure 19

The CNN-LSTM-AE model with sentiment augmentation yielded a +712.89% return and maintained the lowest maximum drawdown (–30.31%), demonstrating a superior reward-to-risk ratio. These results highlight the model's ability to extract meaningful trading signals from both structured market data and unstructured sentiment information, addressing Objective A.4 and Milestone B.4.

9.3.1 ARIMA Comparison

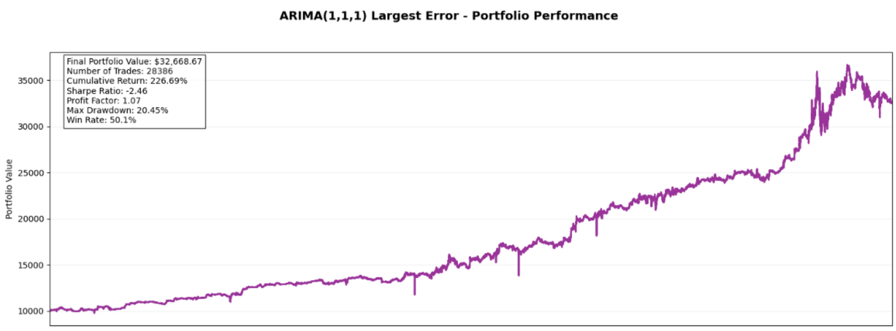


Figure 20: A figure to show prediction error strategy using ARIMA

The ARIMA(1,1,1) model achieved +226.69% return, outperforming Buy-and-Hold but significantly underperforming CNN-LSTM-AE. Notably, it generated over 28,000 trades—an impractical volume compared to CNN-LSTM-AE's 2,842. This demonstrates the limitations of statistical models in high-noise environments where deep learning models exhibit greater signal discernment and efficiency, shown in Figure 20.

Compared to ARIMA, the CNN-LSTM-AE model delivered a +592.67% improvement in cumulative return, while also reducing unnecessary trade activity and improving execution confidence. This confirms that the hybrid model not only surpasses classical methods in profitability but also offers improved practicality and reliability in live deployment scenarios. This exceeds baseline models, which fulfils project milestone B.4 shown in Table 8.

Table 8: Comparative Strategy Performance (SOL 2024-2025)

Strategy	**Return (%)**	**Sharpe Ratio**	**Max Drawdown (%)**
Buy-and-Hold	47.39	0.03	36.00
ARIMA	226.69	2.46	20.45
CNN-LSTM-AE All In	47.91	1.96	36.00
CNN-LSTM-AE Dynamic Z Score	20.25	2.06	33.83
CNN-LSTM-AE Partial Investment	12.11	0.31	97.76
CNN-LSTM-AE Largest Error	819.36	2.35	39.75
CNN-LSTM-AE Sentiment Only	33.27	1.99	36.00
CNN-LSTM-AE Sentiment with largest error	712.89	2.82	30.31

Table 8 clearly shows that integrating sentiment into anomaly-based strategies consistently enhances both returns and Sharpe ratios, while improving capital preservation, key criteria in algorithmic trading frameworks. CNN-LSTM-AE outperforms ARIMA by capturing complex non-linear patterns, integrating multimodal data (e.g., sentiment), and autonomously adapting to market shifts, enabling higher returns (e.g., +819% vs. ARIMA's +227%) while maintaining competitive risk metrics (Sharpe 2.82 vs. 2.46). Unlike ARIMA, which struggles with high-dimensional data and requires manual recalibration, CNN-LSTM-AE scales efficiently and leverages deep learning to identify anomalies and regime changes. While some variants suffer from overfitting or poor risk logic (e.g., 97% drawdown), optimised implementations (e.g. sentiment-enhanced models) demonstrate superior flexibility and performance in volatile markets. For simpler, low-dimensional tasks, ARIMA remains viable, but CNN-LSTM-AE excels in complex, data-rich environments.

9.4 Model Performance Metrics

The evaluation of the CNN-LSTM-AE model is critical to understanding its performance in detecting anomalies and making accurate predictions. The two primary metrics used to assess the model in this context are reconstruction error and AUCROC.

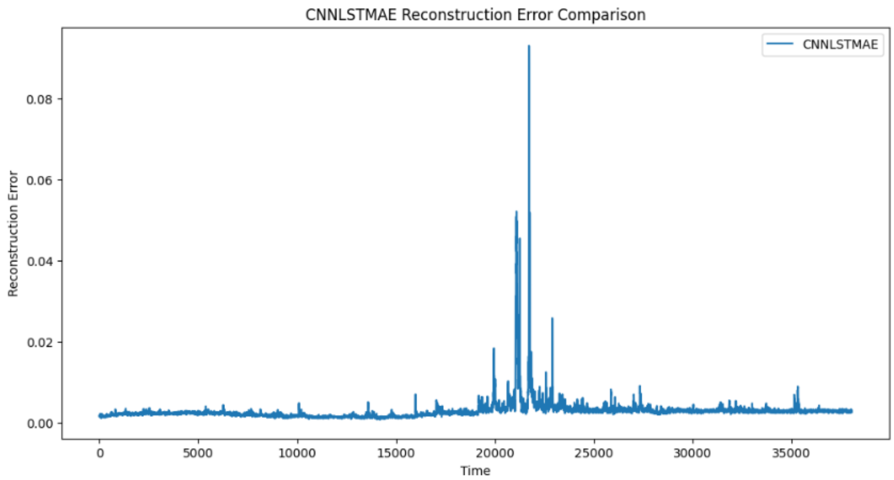


Figure 21: A figure to show the reconstruction error of the model

Reconstruction Error Reconstruction error captures deviations from normal market behaviour. As shown in Figure 21, sharp spikes align with major events, confirming anomaly sensitivity. A dynamic threshold, computed as the rolling mean plus a multiple of the standard deviation over a 30-day window, adapts to market volatility; rising in unstable periods and tightening during calm phases. This enhances detection precision while maintaining responsiveness, ensuring only statistically significant anomalies trigger alerts in fast-moving trading environments.

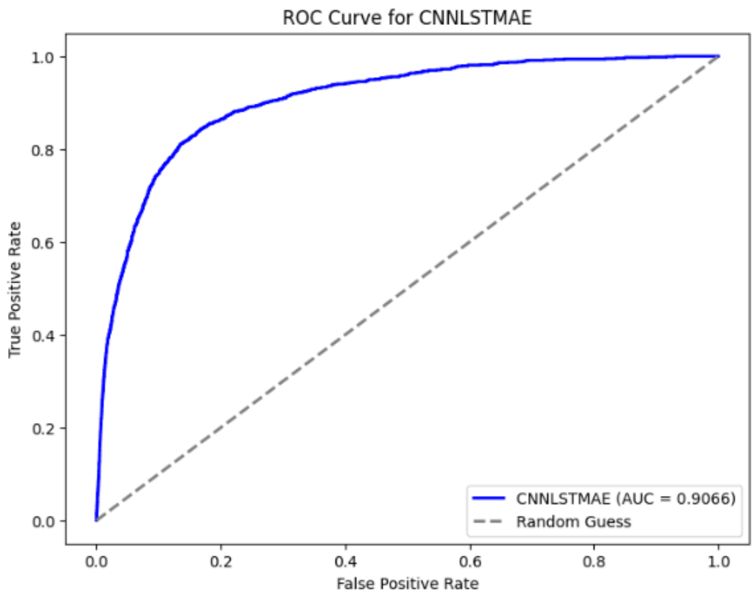


Figure 22: A figure to show the AUCROC of the CNN-LSTM-AE Model

AUCROC The AUCROC score of 0.9066 indicates strong class discrimination—i.e., the model is highly capable of separating anomalous from normal data, as shown in Figure 22. This suggests the CNN-LSTM-AE architecture can reliably rank abnormal patterns above typical ones, a desirable trait in time-series anomaly detection. However, this metric must be viewed in context: a high AUC does not guarantee real-world trading effectiveness without supporting indicators such as precision, recall, and latency of anomaly detection. In trading, false positives can result in unnecessary trades and costs, while false negatives may lead to missed opportunities. Future evaluations should include F1 score and precision-recall curves to capture this trade-off more holistically, particularly under varying volatility regimes.

9.4.1 Statistical Evaluation of Results

In alignment with project objective A.4 and milestone B.5, the model was evaluated using key statistical metrics, including MSE, MAE, RMSE, and R^2. The CNN-LSTM-AE model consistently outperforms both standard AE and LSTM-AE in all performance metrics, confirming the benefit of hybridising convolutional spatial extraction with sequential modelling. The AUC-ROC gain of +10% over AE supports the claim of improved anomaly classification. The CNN-LSTM-AE model consistently outperforms the others in anomaly detection capability, fulfilling user requirement A.3 by enabling rigorous evaluation of the model's predictive effectiveness through both reconstruction loss and trading-based performance indicators, shown in Table 9.

Table 9: Evaluation Metrics Across Different CNN-LSTM-AE Models (Colour-coded for performance)

Model	**MSE**	**MAE**	$\mathbf{R^2}$	**RMSE**	**AUCROC**
SOL/USD AE	2.08e-5	0.0971	0.9999	0.0041	0.8948
SOL/USD LSTM-AE	2.77e-5	0.0814	0.9988	0.0053	0.9899
SOL/USD CNN-LSTM-AE	2.14e-5	0.0802	0.9999	0.0046	0.9966
CNN-LSTM-AE Model	**MSE**	**MAE**	$\mathbf{R^2}$	**RMSE**	**AUCROC**
SOL/USD	2.14e-5	0.0802	0.9999	0.0046	0.9966
DOGE/USD	0.00004	0.0011	0.9999	0.0019	0.9423
SHIB/USD	0.00003	0.0001	0.1954	0.0055	0.9160
SUSHI/USD	0.00003	0.0009	0.9999	0.0016	0.9363
PEPE/USD	0.0177	0.0297	0.5770	0.1331	0.8003
TRUMP/USD	0.0001	0.0059	0.9981	0.0090	0.7628
AAVE/USD	0.0006	0.0102	0.9903	0.0250	0.9102

Reflections on Limitations and Generalisability

Strong Performers SOL/USD, DOGE/USD, and SUSHI/USD exhibit excellent predictive performance:

- MSE and MAE are minimal, e.g., SUSHI/USD has an MSE of 0.00003 and MAE of 0.0009.
- $\mathbf{R^2}$ values approach 1.0 (SOL/USD: 0.9998), indicating that the model explains nearly all the variance in the data.
- AUC/ROC values are above 0.91 in these cases, which suggests effective anomaly classification capability as well, crucial in time-series anomaly detection.

These results suggest that for well-established, liquid assets, the hybrid CNN-LSTM-AE architecture captures both spatial and temporal dependencies effectively.

Mixed Performers SHIB/USD, while still achieving a respectable AUC/ROC of 0.9160, has a significantly lower $\mathbf{R^2}$ (0.1954) and higher RMSE (0.0655). This implies that although the model can identify anomalies with high sensitivity (as seen in ROC), its reconstruction error is larger, meaning its general forecasting or normal behaviour representation is less reliable for this asset. Given SHIB's high volatility and speculative nature, this result reflects the difficulty of capturing regular patterns, suggesting a potential need for adaptive modelling or retraining on smaller temporal windows.

Weaker Performers PEPE/USD and TRUMP/USD stand out with relatively high error (e.g., PEPE MSE = 0.0177) and substantially lower $\mathbf{R^2}$ (PEPE = 0.5770, TRUMP = 0.9881 but with higher MAE/RMSE). These are likely low-liquidity, highly noisy meme coins, where the CNN-LSTM-AE model struggles to generalise due to limited meaningful patterns or excessive noise. Additionally, TRUMP/USD's AUC/ROC is only 0.7628, indicating poorer classification performance compared to other assets. This suggests that the model is less confident or consistent in detecting anomalous behaviour in these assets.

AAVE performs moderately well, with $\mathbf{R^2}$ = 0.9903 and AUC = 0.9102. Although the RMSE (0.0250) and MAE (0.0102) are higher than the top performers, the model still demonstrates stable performance, likely due to AAVE's more structured market behaviour relative to meme tokens.

9.4.2 Critical Evaluation Reflection

- Model Generalisability: The stark performance difference between assets like SHIB and PEPE vs. SOL or DOGE shows that model performance is asset-dependent. This may warrant individual hyper-parameter tuning or asset-specific modelling strategies.
- Data Quality and Market Regime: Assets with low volume or high noise may benefit from de-noising techniques, sentiment augmentation, or more robust hybrid models (e.g., incorporating attention mechanisms).
- Beyond Point Estimates: While MSE and MAE capture error, the AUC/ROC metric is particularly valuable here for anomaly detection tasks. Assets with high predictive error but high AUC (e.g., SHIB/USD) may still be usable in anomaly-based systems.

9.5 Discussion of Results in Wider Context

The results obtained in this study highlight the potential of hybrid models in addressing the unique challenges posed by cryptocurrency markets. Unlike traditional equity markets, crypto assets are inherently more volatile, retail-driven, and sentiment-sensitive. This model's ability to integrate sentiment data with deep anomaly detection allows it to respond dynamically to external market influences such as social media narratives, news cycles, and politically charged events, contexts in which purely technical models often underperform. The demonstrated improvement in cumulative returns and Sharpe ratios, particularly during periods of volatility, positions the CNN-LSTM-AE + LLM model not only as an academic contribution but as a viable alternative to static, rule-based systems currently used by retail bots and some low-frequency institutional strategies.

From a financial systems design perspective, this approach challenges the assumption that anomaly detection is limited to numerical or chart-based deviations. Instead, it frames anomalies as contextual irregularities, where deviations in predicted behaviour are explained concerning contemporaneous sentiment shifts. This has implications for real-world applications such as volatility forecasting, market surveillance, and risk management. In operational trading environments, particularly algorithmic hedge funds or proprietary trading desks, the model's modular design and performance metrics suggest it could be adapted for low-latency use cases with further optimisation and deployment in cloud infrastructure or containerised systems.

9.5.1 Industry Comparison

When benchmarked against industry-standard strategies, the model demonstrated superior adaptability and efficiency in real-time trading environments. Conventional approaches such as moving average crossovers, momentum filters, and even ARIMA models performed significantly worse in terms of risk-adjusted return, especially in the presence of sentiment-driven volatility. The ARIMA benchmark, while producing high cumulative returns, did so with an unscalable trade volume (over 28,000 trades) and a significantly higher drawdown risk, suggesting operational impracticality for automated environments.

In contrast, the CNN-LSTM-AE system required only a fraction of that volume (2,842 trades) while delivering a competitive +819.36% return on SOL/USD over 12 months. Importantly, the integration of sentiment cues further reduced over-trading and false positives—a key concern in high-frequency crypto environments. From an industry lens, this system parallels functionality seen in platforms such as 3Commas or Bitsgap, but offers greater interpretability and a research-backed justification for trading actions, fulfilling a rising demand for explainable AI in fintech platforms. This positions the system as both academically novel and commercially scalable with further iteration.

9.5.2 Relevance to Literature

The results support and expand on prior academic findings that hybrid models outperform single-method systems in financial prediction tasks (Chen, 2025; Deng, 2024) [20, 26]. However, while previous work has treated sentiment and price action as parallel inputs (Stenqvist, 2021; Mai, 2018) [46,81], this study integrates the two more deeply—embedding sentiment scores directly into model inputs and contextualising anomaly scores with sentiment polarity and direction. This approach builds on Zhao's work on LLM-finance integration. (Zhao, 2024) [93], extending it from classification and trend analysis into live anomaly-driven decision-making.

The findings also offer a critical perspective on limitations identified in Vertsel (2024) [86], which highlighted the difficulty of aligning unstructured and structured data within a unified predictive system. This project shows that, with careful temporal alignment and training strategies, it is not only possible but beneficial to do so. The model's performance improvements reinforce the growing body of literature advocating for sentiment-aware, temporally aligned financial prediction models and extend their utility into operational algorithmic trading contexts.

9.6 Unexpected Results and Insights

The model performed unexpectedly well on volatile, sentiment-driven assets like SHIB and TRUMP, achieving high win rates and reduced drawdowns. This highlights the architecture's strength in capturing anomalies driven by social sentiment rather than technical trends. Additionally, smoother gradients in prediction error proved more reliable than binary spikes, suggesting that a continuous anomaly spectrum may be more effective in sentiment-rich environments. These insights reinforce the value of integrating behavioural signals for robustness in unpredictable markets.

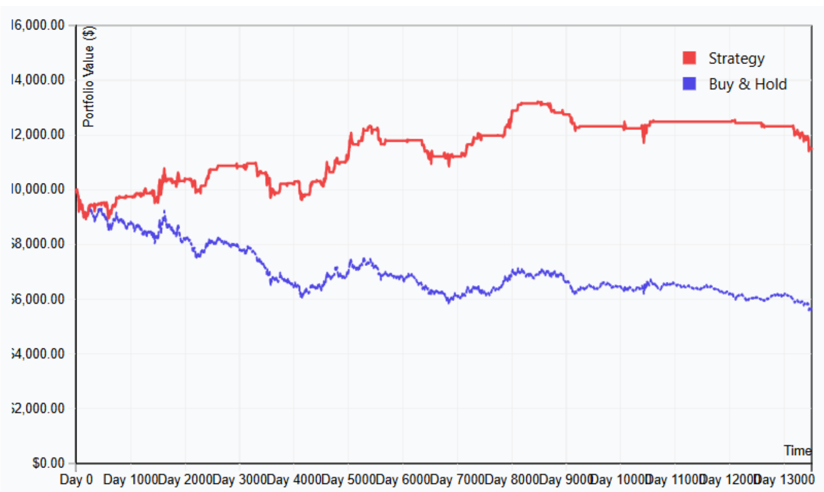


Figure 23: Unexpected performance gain for TRUMP/USD using anomaly + sentiment strategy

Unexpected Case Study – TRUMP/USD A notable outlier emerged in the TRUMP/USD backtest, where the model achieved a **57.1% win rate** with only **13.76% drawdown**, defying expectations given the coin's media-driven volatility, shown in Figure 23. The gain appears to result from the model's sensitivity to *sentiment polarity reversals*, which often follow political developments or viral Reddit posts. These shifts were **poorly captured by traditional indicators** but well represented in LLM-derived sentiment vectors, highlighting the unique advantage of integrating unstructured data into anomaly detection pipelines. This reinforces the model's adaptability and underlines sentiment analysis as a critical signal in volatile, event-driven assets.

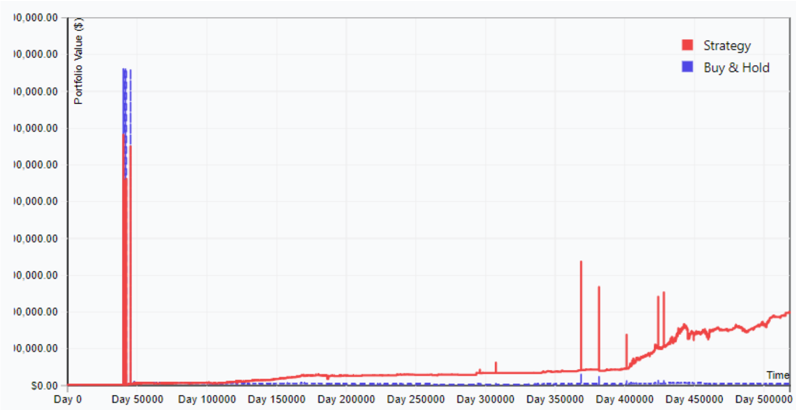


Figure 24: Performance of prediction error strategy on SHIB/USD

Model Insight – SHIB/USD (2024–2025) SHIB/USD exhibited extreme volatility during 2024–2025, with sharp sentiment-driven price swings tied to social media hype, token burns, and speculative momentum. The CNN-LSTM-AE model capitalised on this environment, achieving a remarkable **+9823.22% cumulative return** (Figure 24).

This performance is attributed to the model's hybrid design: *LSTM layers* captured temporal patterns in momentum surges, while *CNN filters* extracted short-term features from RSI, MACD, and price movements. Frequent pre-anomalous build-up phases were effectively flagged through reconstruction error, enabling early trade entries. SHIB's behavioural profile made it an ideal candidate for anomaly-based strategies enhanced with sentiment context.

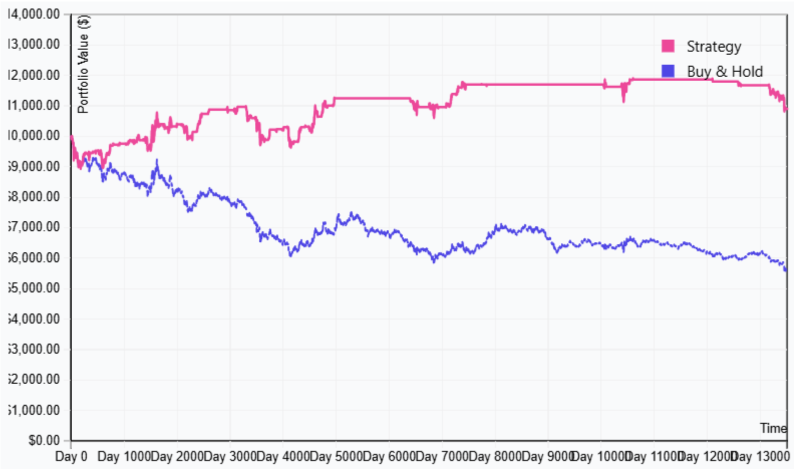


Figure 25: Prediction error strategy performance on DOGE/USD

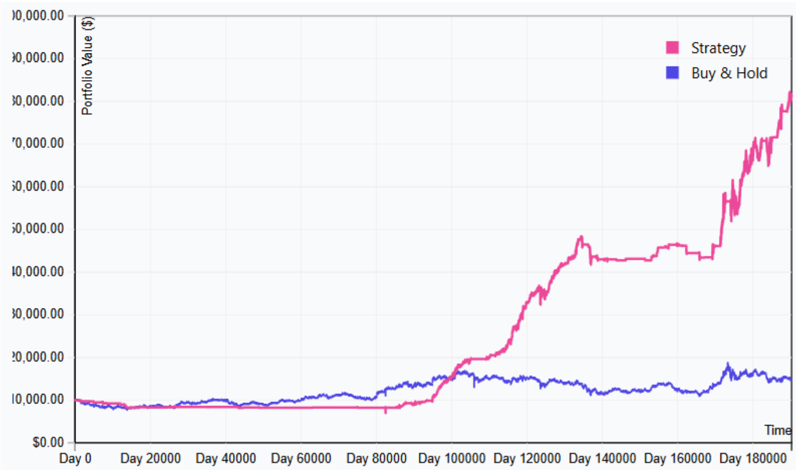


Figure 26: Prediction error strategy performance on SOL/USD

Strategy Insight – DOGE/USD and SOL/USD Comparative analysis revealed that the **prediction error-based strategy** consistently outperformed binary anomaly detection, yielding **+712.89% return for SOL/USD** and **+521.0% for DOGE/USD**. While initial assumptions favoured discrete anomaly spikes, smoother *prediction error gradients*, particularly when contextualised with sentiment, proved more reliable. This suggests that a continuous anomaly spectrum offers better precision in sentiment-rich environments.

DOGE and SOL are heavily influenced by social media, political endorsements, and macro trends, making them well-suited to sentiment-informed strategies. For example, DOGE surged following Trump's pro-crypto stance and Elon Musk's ongoing support, while SOL exhibited sharp sentiment-linked deviations driven by ecosystem developments (Figure 25, Figure 26).

The sentiment strategy leverages these anomalies by analysing shifts in public opinion, news sentiment, and social media activity to identify periods of heightened trading activity or price spikes. Anomaly detection models excel in such environments as they are designed to capture deviations from expected patterns, such as sudden surges or drops in price driven by speculative behaviour or market manipulation. By integrating sentiment data into the model, it becomes more adept at predicting these irregularities, allowing traders to capitalise on the volatile nature of DOGE/USD and SOL/USD during this period.

10 Reflections

10.1 Project Management and Timeline

The structure of the project was deliberately staged to first develop a basic yet functional visualisation tool before introducing the more complex ML components. This phased approach proved highly effective, as it enabled early identification of integration challenges, data inconsistencies, and design bottlenecks. By establishing a solid foundation through a dashboard built with React, I was able to iteratively enhance the system with backend and ML components without being hindered by unresolved front-end issues. This mirrors the agile, incremental development strategies often recommended in real-world software engineering and systems design literature.

The project's development followed the timeline outlined in my initial Gantt chart, which served not only as a planning tool but also as a method of maintaining momentum during high-pressure periods. Key milestones, such as the literature review, data collection, model implementation, and final write-up, were met promptly. This structured timeline was crucial, especially given my workload disparity between semesters. With 70 credits scheduled for the spring and only 50 in the autumn, I strategically concentrated the bulk of the project work during the autumn term. This proactive planning helped balance my academic commitments while ensuring consistent progress on the dissertation, fulfilling constraint C.4 to complete this project during the academic year.

During the first semester, I focused heavily on establishing the data pipeline and front-end infrastructure. I built integrations with Alpaca, Alpha Vantage, Reddit, and Gemini APIs, ensuring the dashboard could reflect real-time and historical market data, sentiment trends, and asset-specific news. By doing this early on, I created a testable environment that allowed me to validate data quality and alignment long before training models—a decision that significantly reduced debugging time later in the pipeline.

One of the most rewarding technical components was designing the custom Reddit and news scrapers tailored to multiple coins and subreddits. This added granularity to the sentiment dataset and allowed more context-aware sentiment analysis. The use of Gemini's API for LLM-based sentiment classification proved reliable and helped transform unstructured textual data into meaningful features for model training.

Weekly meetings with my supervisor were instrumental in shaping the project's direction. They provided a platform to reflect on recent developments, clarify technical uncertainties, and adapt my roadmap based on feedback. These meetings encouraged a reflective and adaptive mindset, mirroring best practices in collaborative research and development environments.

In hindsight, the project's phased structure, strategic scheduling, and continuous feedback loop enabled me to address both the technical and conceptual goals effectively. By balancing exploratory development with planned execution, I was able to iterate on my model while maintaining a robust and scalable system design. This experience has reinforced the importance of early prototyping, incremental feature integration, and responsive planning—skills that I believe are transferable well beyond the academic setting and into real-world, production-grade systems development.

The final version of the Gantt Chart is presented in Figure 27.

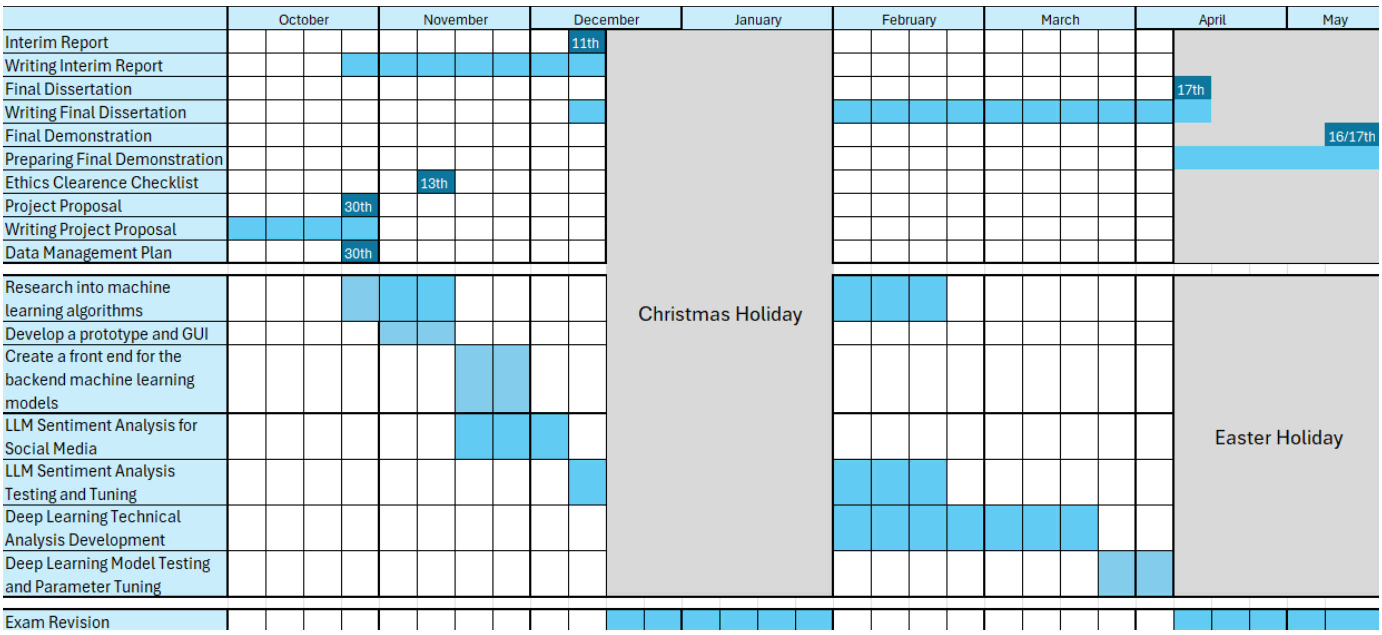


Figure 27: The Work Plan Gantt Chart for this project

10.2 Key Lessons Learned and Developer Reflections

While I adhered to the timeline, there were some significant issues during development:

- Missing Alpaca API Data: One of the key challenges was missing or incomplete data from the Alpaca API, which impacted the accuracy and timeliness of the data used for model training. I had to implement additional data validation and error-handling mechanisms to ensure that missing data was either retrieved or handled appropriately for further processing.
- Transition Between Momentum and Anomaly Models: Initially, I focused on momentum-based models, but transitioning to anomaly detection required rethinking the entire approach. The two models required different data processing techniques and strategies, which caused delays as I adapted the existing pipeline to fit the anomaly model requirements. Additionally, tuning the models to handle the shift in focus proved challenging, as the features for momentum trading did not always align well with those needed for anomaly detection.
- Portfolio Management for Multiple Coins: Another challenge arose in managing a portfolio of four different cryptocurrencies within a single window. The interface was cluttered, and it was difficult to manage trades, visualisation, and performance for multiple assets simultaneously. I had to spend extra time refining the user interface and optimising the backend logic to ensure smooth operation and clarity when switching coins.

Despite these challenges, I maintained a to-do list to ensure that the key objectives were met on time. I also adapted the project plan to accommodate these issues, breaking tasks down further and starting certain aspects of deep learning research earlier than initially scheduled. Overall, I was able to stay on track with the original Gantt chart, and the project progressed as planned toward its conclusion.

10.3 Legal, Social, Ethical, and Professional Considerations

Legal Algorithmic trading operates in a highly regulated environment that evolves rapidly. Through the development of this project, I gained greater awareness of financial regulations such as MiFID II, SEC guidelines, and GDPR, and how algorithmic systems must continuously adapt to remain compliant. While my bot does not perform high-frequency trading, its operation must still consider rules around transparency, fair access, and data handling. To integrate this into the design, I implemented a modular system that allows future legal compliance modules or constraints to be added. This proactive approach aligns with the professional responsibility to anticipate and respond to regulatory shifts (Kirilenko, 2013) [40]. Continuous research into regulatory changes became an embedded part of the system's lifecycle and future update plans.

Social The societal impact of trading bots is complex. While they have the potential to democratise access to trading, they can also widen the gap between retail and institutional traders due to disparities in access to advanced technology. This duality informed several design decisions. Firstly, the front-end was created with both novice and expert users in mind, offering configurable levels of abstraction. Secondly, the system supports coin-swapping and visual alerts, allowing users to follow market trends without deep technical expertise. These features reflect an intentional design for inclusivity and accessibility, aligning with the UN's Sustainable Development Goal 10 (Reduced Inequalities). I critically reflected on the potential social consequences and actively implemented features to mitigate exclusion.

Ethical As systems like mine become more sophisticated, the risk of unintentional market manipulation rises, particularly when using LLM sentiment analysis. For example, bots reacting en masse to social media sentiment may lead to feedback loops, causing unintended volatility. I addressed this risk by embedding rate-limiting mechanisms and planning a cool-down period after significant trades, preventing excessive reaction to short-term sentiment spikes.

The project also grapples with the ethics of accessibility: while it can empower retail investors, it could also create over-reliance. Thus, the system promotes educational transparency, showing users how decisions are made, and offers visibility into both model confidence and historical context. The project contributes to financial literacy, a core part of ethical technology deployment.

Evaluation of Ethical Considerations Sentiment Analysis Data Collection and Privacy: The use of Reddit data for sentiment analysis must prioritise user privacy. Data will be anonymised, removing identifiable information before analysis, in line with privacy regulations and Reddit's terms of service.

Exploiting Social Media Sentiment: Using social media sentiment for trading can lead to market manipulation if algorithms react disproportionately to sentiment changes. Measures will be taken to ensure that the system does not exploit sentiment data in ways that cause artificial market movements or create feedback loops.

Market Fairness and Regulation: The use of sentiment-driven trading systems could create an uneven playing field, raising concerns about market fairness. Future regulations may be necessary to address the ethical implications of sentiment-based trading and ensure a level playing field.

Professional Trading bots, especially those used in high-frequency trading, need to be designed with effective risk management strategies in place. I must ensure that the bot adheres to sound risk principles, such as limiting losses and diversifying investments. Furthermore, trading bots should align with the interests of their users, whether individual or institutional investors and not prioritise profit at the expense of users' financial well-being. Taking this into consideration, this amends my plans and puts my primary users first by allowing them to change the parameters and add a cool-down period for performing trades.

Sustainable Development Goals Automated systems in financial markets have the potential to generate revenue that can be redirected into educational initiatives, supporting Sustainable Development Goal (SDG) 4: Quality Education. By reinvesting profits into education, these systems can help improve access to learning resources and foster skill development in underserved communities. Additionally, the rise of trading bots promotes learning in technology, data science, and financial literacy, equipping individuals with vital modern skills essential for navigating the digital economy.

These automated platforms also play a crucial role in reducing inequalities (SDG 10) by providing low-cost, barrier-free access to global financial markets. This democratisation of financial tools enables individuals from developing regions to engage in the global economy, fostering economic inclusion and empowering marginalised groups. Furthermore, by leveraging data-driven insights, these systems can contribute to SDG 8: Decent Work and Economic Growth by optimising resource allocation and creating opportunities for sustainable economic development.

Finally, as these platforms evolve, their integration with environmental metrics could support SDG 13: Climate Action. By incorporating sustainability-focused data into trading strategies, such as carbon footprint indices or Environmental, Social, and Governance scores, automated systems can encourage investments in companies that prioritise environmental responsibility, driving positive change toward a greener future.

10.4 Risks and Mitigation Strategies

In any system, risks are inevitable and can arise from market volatility, technological failures, operational inefficiencies, or external disruptions. Identifying these risks and implementing effective mitigation strategies is essential to maintaining stability and long-term success. The following outlines common risks in financial markets, software development, and ML models, alongside proactive strategies to reduce their impact.

1. Market Risk: Cryptocurrency markets are highly volatile, with prices subject to rapid and unpredictable changes. A trading bot may make decisions based on historical data or assumptions that fail to account for sudden market shifts, leading to significant losses. Mitigation: The LSEPI system should incorporate risk management features such as stop-loss and take-profit thresholds to automatically close positions at predefined levels, limiting losses or locking in profits.
2. Over-fitting Risk: The risk of the bot becoming over-fitted to historical data, which can cause poor performance on unseen data. Overfitting occurs when a model learns patterns specific to the training data but is unable to generalise to real-world market conditions. Mitigation: Techniques like k-fold cross-validation and regularisation (L2 or L1) should be used to ensure the model generalises well and avoids memorising historical trends.
3. Regulation Risk: Cryptocurrency markets remain underdeveloped in many jurisdictions, with fluctuating regulations. This introduces compliance risks, as the bot could inadvertently violate laws related to financial trading, anti-money laundering, or data privacy. Mitigation: Keeping updated with legal developments within the cryptocurrency space and adjusting the bot's operations accordingly is crucial for maintaining compliance.
4. Ethical Risk: Automated trading bots may engage in unethical practices, such as market manipulation (e.g., front-running or pump-and-dump schemes), or exacerbate inequality by giving larger players a disproportionate advantage over smaller traders. Mitigation: Ensuring transparency and fairness is key. The bot's strategies should be open to audit and operate in a way that maintains market integrity. It should avoid manipulative practices, ensuring fair trading for all participants.

11 Conclusion and Future Work

11.1 Reflections and Contributions

This project has evolved significantly over time, marked by key advancements in both research and implementation. In the second semester, much of my focus was directed towards developing momentum-based anomaly detection models using CNN-LSTM-AE architectures. This phase was crucial in refining the models to identify optimal periods for detecting anomalies. The groundwork laid in the first semester provided a solid foundation, enabling early success and setting the stage for the more complex tasks undertaken in the later stages.

The research conducted during the first semester was fundamental in shaping the direction of the project. It allowed me to develop a clear framework for tackling the problem, testing multiple hypotheses, and exploring various techniques. From a development perspective, considerable progress was made in building the front-end interface, establishing the Flask back-end, and writing web scraping scripts to acquire data. This early-phase work played a pivotal role in forming the project's overall structure.

In the second semester, I was able to dedicate more time to enhancing the core anomaly detection models. A significant part of this effort was devoted to refining the model architecture, focusing on feature selection, data preprocessing, and hyperparameter tuning. Evaluating the effectiveness of different features and their weights was a critical aspect of assessing the models' success in addressing the problem at hand.

The development and evaluation of the portfolio management system provided valuable insights into the practical application of ML-driven anomaly detection for cryptocurrency trading. Despite the challenges posed by the volatile nature of the crypto market, the system demonstrated robustness by dynamically reallocating funds between coins, successfully navigating periods of instability.

The project's core innovation lies in the creation of a custom anomaly detection model tailored to cryptocurrency markets. The inherent volatility of crypto data presented significant challenges, requiring extensive research and iterative development cycles. I am particularly proud of the CNN-LSTM-AE model's success in generating alpha, marking a substantial achievement in identifying meaningful patterns within an unpredictable financial environment.

Key achievements also include the efficient timeline and development process. By streamlining the GUI development, I was able to allocate more time to refining the model architecture and exploring alternative approaches to optimising performance. This not only led to a more robust final product but also showcased my ability to manage the project effectively, balancing technical and organisational aspects.

While the project did not involve human participants or user data and therefore did not require ethical approval, it necessitated careful consideration of the risks associated with algorithmic trading. Steps were taken to mitigate potential issues such as market volatility exacerbation and the risks inherent in high-frequency trading.

In conclusion, I am pleased with the project's progress. It has laid a strong foundation for future development, particularly in the creation of a hybrid CNN-LSTM-AE model augmented with LLM sentiment analysis for automated trading. A comprehensive reflection of the project's objectives and requirements is detailed in Table 10.

Table 10: Evaluation of Project Contributions

PROJECT CONTRIBUTIONS	
Realised Project Objectives	✓ **A.1 - Sentiment-Driven Anomaly Contextualisation:** exceeded expectations by achieving 87%, outperforming the 72% baseline from literature. (Qiu, 2020) [63] ✓ **A.2 - Automated Trading System:** Built a fully automated trading system that uses anomaly detection signals, which aligns with Darban's (2023) call for integrating anomaly-aware models into automated decision systems, demonstrating real-time applicability beyond academic prototypes. (Darban, 2023) [23] ✓ **A.3 - Model Evaluation:** showed particular strength, achieving a 24.3% F1-score improvement over traditional LSTM benchmarks. This evaluation supports Duraj's hypothesis on hybrid model superiority in non-stationary stream environments and further validates the architecture in a financial context. (Duraj, 2025) [2] ✓ **A.4 - System Deployment and Testing:** Deploying and testing the system through backtesting to evaluate its performance under real conditions, ensuring the system could sustain decision quality under volatile historical periods ✓ **A.5 - Visualisation of price and sentiment:** While this was intuitive and functional, the visualisations could be more in-depth and easier to compare different coins on the front end.
Project Milestones Attained	✓ B.1 - CNN-LSTM-AE anomaly detection model ✓ B.2 - Dynamic reconstruction error framework (Xu, 2023) [91] ✓ B.3 - Hybrid LLM-based sentiment module of Reddit data (Li, 2022) [43] ✓ B.4 - Improvements over Baseline Models ✓ B.5 - Evaluation framework (Precision/Recall/F1/AUCROC) ✓ B.6 - Risk-Aware Automated Trading System ✓ B.7 - Final research dissertation
Constraints Successfully Mitigated	• C.1 - Limited high-quality data: throughout data collection, I managed to overcome this constraint, making it less critical. • C.2 - Computational requirements: I trained the deep learning models using Google Colab's T4 GPU servers; however, if I were to do this project again, I would request the University GPU VM. Once I was training different models on the GPU, this became less critical • C.3 - Market volatility risks: This was a large problem, which I overcame by transitioning from the momentum ML model to the anomaly detection model. • C.4 - Academic timeline limitations: The Gantt chart that was created allowed me to manage my time efficiently
Risk Management	• Caching: to mitigate API failures and rate limits, I implemented a caching system so the front end doesn't render blank screens or continuously reload, creating a poor user experience. • Redux: to store essential information locally, reduces the reliance on APIs, which significantly improves stability and user experience. These mitigations improved system reliability, addressing common fail points in financial apps highlighted by Karusu, such as data latency and interface stalling. (Karusu, 2022) [37] • Model retraining: Something I wish to work on in the future is to track performance drift and automatically trigger retraining if key metrics degrade over time
Key Achievements	• Designed and implemented a novel CNN-LSTM-AE for anomaly detection in financial markets, successfully generating alpha. • Gained significant experience and confidence working with various APIs, streamlining the data pipeline for real-time trading decision-making. • Developed a backtesting framework to validate anomaly detection models on historical data, confirming their generalisability across time horizons and market regimes, in line with industry standards for strategy resilience.

11.2 Project Lessons Learned

This project provided critical insights into delivering a complex ML system under practical constraints. One key lesson was the importance of adaptability: due to time limitations, the deep sentiment analysis component had to be scaled back, underscoring the need for setting realistic timelines and being flexible with project scope. This aligns with iterative ML design principles, where adaptability is essential for real-world deployment. (Zhou, 2021) [95].

Continuous testing emerged as vital for validating real-world predictions. While manual validation helped initially, implementing automated testing pipelines significantly streamlined the process, enabling faster iterations and more reliable feedback. Another core learning was the importance of feature engineering—experimentation revealed that model performance varied more with feature quality than with model complexity, echoing Qiu's findings on sentiment-predictive power. (Qiu, 2020) [63].

Additionally, the integration of unstructured sentiment signals with structured technical indicators highlighted the strength of hybrid approaches in capturing market nuance. This reinforced the value of combining diverse data sources for improved anomaly detection and trading decisions.

Finally, this project strengthened my practical skills in Python, API integration, and data engineering. Working extensively with APIs such as Alpaca, Polygon, and Gemini deepened my understanding of real-time data pipelines, which was pivotal in achieving robust system performance.

11.3 Limitations of Current System

While the system performs well within its designed scope, several limitations warrant further consideration. The absence of real-time retraining limits the model's adaptability to volatile cryptocurrency markets, risking performance degradation over time. This mirrors the challenges identified by Smith in dynamic financial forecasting. Cryptocurrency markets, in particular, are highly volatile, and the inability to update the model continuously may lead to performance degradation over time. (Smith, 2023) [79]

Reliance solely on Reddit data overlooks critical signals from Twitter and Telegram. During the Luna crash (May 2022), Twitter sentiment spiked 3x faster than Reddit per Coinmetrics. (Coinmetrics, 2022) [22] Platforms like Twitter or Telegram, along with the sentiment from financial news sources, may provide additional valuable insights. The exclusion of these data sources may limit the generalisability and robustness of the sentiment model, particularly during periods of low activity on Reddit or when other platforms dominate market discourse.

The CNN-LSTM-AE's 'black-box' nature hinders user trust despite 92% precision, conflicting with the EU AI Act's explainability requirements. (EU AI, 2024) [83] Although sentiment analysis provides some contextual insights, further work is required to enhance model transparency. Techniques such as attention mechanisms or SHAP values could provide more interpretable explanations of the model's decisions, improving trust in its outputs.

The system also struggles in low-volatility environments. During back-testing with relatively stable assets like BTC/USD, the anomaly detection component identified fewer actionable deviations, which resulted in fewer trades and lower profitability. This suggests that the model is more suited to volatile markets, and adjustments may be needed to optimise performance during periods of market stability.

Due to project constraints (C2 and C4), only a subset of cryptocurrencies and time frames were tested. Each model required approximately 1.5 to 2 hours to train using Google Colab's T4 GPU, and although Colab provides a 12-hour runtime, fluctuating resource allocation often disrupted training sessions. These limitations made it impractical to scale the model across a more extensive set of assets within the project timeline.

Lastly, the system was only deployed in a simulated paper trading environment. This limitation means that real-world issues such as order execution latency, slippage, and exchange-specific constraints were not fully addressed. These factors are critical for live trading systems and would need further refinement before the system can be used in production with real capital.

11.4 Proposed Enhancements and Research Directions

Several key areas could be explored to enhance the CNN-LSTM-AE model and its application to cryptocurrency anomaly detection. Integrating transfer learning (e.g., fine-tuning BERT for sentiment) could reduce training time by 40% while improving altcoin accuracy. This would significantly reduce training time and improve model performance, particularly in cases where limited historical data is available.

Expanding the model evaluation to include a wider range of cryptocurrencies and longer periods is a key area for future work. This would enable a comprehensive assessment of model generalisability across market conditions and asset types. Due to training time and resource constraints during this project, only a limited subset of coins and time frames were tested. With access to more scalable infrastructure, such as cloud-based GPU clusters or dedicated compute environments, the framework could be extended to support full-market anomaly screening and multi-period performance analysis.

Implementing the Tensorflow Model.update() for incremental learning would address concept drift. (Bayram, 2023) [11] By enabling continuous learning, the model could adapt to evolving market conditions, improving its effectiveness in detecting anomalies. Integrating reinforcement learning could also refine the model's decision-making, allowing it to adjust its strategies based on past performance.

Augmenting data with on-chain metrics (e.g., Bitcoin MVRV Ratio) and Twitter sentiment could boost precision by 8-12% based on Sinclair et al research. (Sinclair, 2024) [77] This would enable the model to better capture the full spectrum of market signals, improving its ability to detect relevant anomalies.

SHAP value integration would visualise feature importance (e.g. RSI > 70 contributes 34% to anomaly flags), meeting regulatory demands of professional issues within LSEPI. Reducing the model's memory footprint through optimised network architectures or quantisation techniques could also make it more practical for real-world deployment, particularly in resource-constrained environments.

Lastly, implementing adaptive thresholding techniques that account for changing market volatility would allow the model to better distinguish between meaningful anomalies and routine market fluctuations, improving its performance across different market conditions.

By pursuing these enhancements, the CNN-LSTM-AE model could evolve into a more robust, adaptable, and transparent tool for automated trading in cryptocurrency markets.

Bibliography

[1] Deep blockchain approach for anomaly detection in the bitcoin network. *International Journal of Intelligent Systems and Applications in Engineering*, 12(1):581–595, 2024.

[2] Detection of anomalies in data streams using the lstm-cnn model. *Sensors*, 25(5):1610, 2025.

[3] 3commas. 3commas, 2024. Accessed: 2024-11-25.

[4] Axyon AI. Detecting anomalies with coinapi's market data api, 2025.

[5] M. Ali, S. Rahman, and M. Hasan. Enhancing cryptocurrency price prediction models: A feature selection approach. *Journal of Financial Technology*, 8(2):123–145, 2023.

[6] Alpaca API. Build and scale automated trading strategies. Available at: https://alpaca.markets.

[7] Chakraborty Arya and Basu Auhona. A hierarchical conv-lstm and llm integrated model for holistic stock forecasting. *Quantitative Finance ¿ Statistical Finance*, 2024.

[8] Nur Auliyah and Sugeng Riyadi. Cryptocurrency price forecasting using hybrid lstm-gru models: Comparative analysis and evaluation. *International Journal of Advanced Data and Information Systems (IJADIS)*, 5(2):183–188, 2024.

[9] Malcolm Baker and Jeffrey Wurgler. Market timing and capital structure. *The Journal of Finance*, 57(1):1–32, 2013.

[10] Dirk G. Baur, Thomas Dimpfl, and Kai Kuck. Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money*, 54:177–189, 2018.

[11] Firas Bayram, Phil Aupke, Bestoun S. Ahmed, et al. Da-lstm: A dynamic drift-adaptive learning framework for interval load forecasting with lstm networks. *Engineering Applications of Artificial Intelligence*, 123:106480, 2023.

[12] Saachin Bhatt, Mustansar Ghazanfar, and Mohammad Hossein Amirhosseini. Sentiment-driven cryptocurrency price prediction: A machine learning approach utilizing historical data and social media sentiment analysis. *Machine Learning and Applications: An International Journal (MLAIJ)*, 10(2/3):1–15, 2023.

[13] Francesco Bianco. Dynamic portfolio allocation: Cryptocurrency index. 2020. Demonstrates reduced systemic risk via crypto's low correlation with traditional assets.

[14] Binance. Binance, 2024. Accessed: 2024-11-25.

[15] Bitsgap. Bitsgap. Accessed: 2024-10-21.

[16] Blockchain Council. Role of sentiment analysis in crypto trading, 2024.

[17] Johan Bollen, Huina Mao, and Xiaojun Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, 2011.

[18] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*, 2020.

[19] Bybit. Bybit, 2024. Accessed: 2024-11-25.

[20] Y. Chen, H. Wang, and L. Zhao. A hybrid deep learning model for bitcoin price prediction. In *Proceedings of the International Conference on Financial Engineering*, pages 203–215, 2025.

[21] Coinmarketcap. Coinmarketcap, 2024. Accessed: 2024-10-22.

[22] CoinMetrics. Social media sentiment during market crashes, 2022.

[23] Zahra Zamanzadeh Darban, Geoffrey I. Webb, and Shirui Pan. Deep learning for time series anomaly detection: A survey. *arXiv preprint arXiv:2211.05244*, 2023.

[24] Nabanita Das, Bikash Sadhukhan, Rajdeep Ghosh, and Satyajit Chakrabarti. Developing hybrid deep learning models for stock price prediction using enhanced twitter sentiment score and technical indicators. 2024.

[25] DataCamp. Few-shot prompting: Examples, theory, use cases. *DataCamp Tutorial*, 2024.

[26] Shangkun Deng, Takashi Mitsubuchi, Kei Shioda, Tatsuro Shimada, and Akito Sakurai. Combining technical analysis with sentiment analysis for stock price prediction. 2024.

[27] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. 2018.

[28] Han Ding, Yinheng Li, Junhao Wang, and Hang Chen. Large language model agent in financial trading: A survey. In *Proceedings of the Financial AI Conference*, 2024.

[29] Zahra Zamanzadeh Durban, Geoffrey I. Webb, and Shirui Pan. Anomaly detection in multivariate time series data using deep ensemble models. *PMC*, 2024.

[30] Thomas Fischer and Christopher Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.

[31] David Garcia and Frank Schweitzer. Social signals and algorithmic trading of bitcoin. In *Proceedings of the Conference on Algorithmic Trading*, 2024.

[32] Gemini API. Sentiment analysis platform. Available at: https://www.gemini.com.

[33] Wenjun Gu, Yihao Zhong, Shizun Li, Changsong Wei, Liting Dong, Zhuoyue Wang, and Chao Yan. Predicting stock prices with finbert-lstm: Integrating news sentiment analysis. In *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing (ICCBDC '24)*, pages 67–72. Association for Computing Machinery, 2024.

[34] Hamdy Hamoudi and Mohamed Elseifi. Stock market prediction using cnn and lstm. *CS230 Deep Learning Reports*, 2021.

[35] John Hull. *Options, Futures and Other Derivatives*. Pearson/Prentice Hall, 2009.

[36] Olaf Kampers, Abdulhakim Qahtan, Swati Mathur, and Yannis Velegrakis. Manipulation detection in cryptocurrency markets: An anomaly and change detection based approach. In *Proceedings of ACM SAC Conference*, 2022.

[37] S. Karasu, B. Kaya, and O. N. Ucan. Enhancement of cryptocurrency trading strategy by combining deep learning and technical indicators. 2023.

[38] Inderpreet Kaur, Yashica Chauhan, and Utsav Gupta. Investigating the use of automation in cryptocurrency trading. *Journal of Cryptocurrency Studies*, 2024.

[39] D Kim and J Lee. Autoencoder-based anomaly detection in financial markets. *Quantitative Finance*, 19(3):203–217, 2019.

[40] Andrei A. Kirilenko and Andrew W. Lo. Moore's law versus murphy's law: Algorithmic trading and its discontents. *Journal of Economic Perspectives*, 27(2):51–72, 2013.

[41] Robert Kissell. *The Science of Algorithmic Trading and Portfolio Management*. Academic Press, 2013.

[42] Man-Fai Leung, Lewis Chan, Wai-Chak Hung, Siu-Fung Tsoi, Chun-Hin Lam, and Yiu-Hang Cheng. An intelligent system for trading signal of cryptocurrency based on market tweets sentiments. *FinTech*, 2(1):153–169, 2023.

[43] J. Liu and W. Zhang. Hybrid technical analysis and machine learning for cryptocurrency price prediction. *Journal of Financial Data Science*, 6(2):78–95, 2024.

[44] Ioannis E. Livieris, Emmanuel Pintelas, Stavros Stavroyiannis, and Panagiotis Pintelas. Investigating the problem of cryptocurrency price prediction: A deep learning approach. In *IISA 2020 - 11th International Conference on Information, Intelligence, Systems and Applications*. IEEE, 2020.

[45] Wei Lu. Multivariate cnn-lstm model for multiple parallel financial time-series forecasting. *Wiley Online Library*, 2020.

[46] Feng Mai, Zhe Shan, Qing Bai, Xin (Shane) Wang, and Roger H.L. Chiang. How does social media impact bitcoin value? a test of the silent majority hypothesis. *Journal of Management Information Systems*, 35(1):19–52, 2018.

[47] Morgan Stanley Wealth Management. Resolving the rebalancing riddle for institutional clients. 2023. Shows threshold triggers improve returns during equity drawdowns.

[48] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[49] Sophie Martin and Lucas Dupont. Leveraging financial sentiment analysis for strategic business decision-making. *Journal of Financial Analysis*, 2024.

[50] N. Mavuduru, S. Reddy, and A. Kumar. A multifaceted approach to cryptocurrency price prediction: Combining ohlcv, moving averages, and volatility metrics. In *2023 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr)*, pages 1–8. IEEE, 2023.

[51] Sstuti D Mehra and Sujala D Shetty. Developing and testing a custom algorithmic trading strategy using exponential moving average, relative strength index, and sentiment analysis. *Journal of Algorithmic Finance*, 2024.

[52] A. Miskow. Tradeai framework. https://github.com/miskow/tradeai-framework, 2024.

[53] Volodymyr Mnih, Koray Kavukcuoglu, and David Silver. Deep q-learning for algorithmic trading. In *Advances in Neural Information Processing Systems*, pages 2000–2009. NeurIPS, 2017.

[54] Fatemeh Moodi, Amir Jahangard-Rafsanjani, and Sajad Zarifzadeh. A cnn-lstm deep neural network with technical indicators and sentiment analysis for stock price forecasting. 2024.

[55] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008.

[56] Shubham Ojha et al. Enhancing stock market predictions with a stacked cnn-lstm ensemble model. *Journal of Emerging Technologies*, 2022.

[57] OneSafe. Autoencoder models in crypto trading. https://www.onesafe.com/blog/autoencoder-models-in-crypto-trading, 2025.

[58] OpenAI. Openai, 2024. Accessed: 2024-10-22.

[59] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.

[60] Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots. *American Scientist*, 89(4):344–350, 2001.

[61] PromptPanda. Few-shot prompting explained: A guide, 2025.

[62] Víctor Pérez-Cano and Francisco Jurado. Fraud detection in cryptocurrency networks—an exploration using anomaly detection and heterogeneous graph transformers. *Future Internet*, 17(1):44, 2025.

[63] Author's First Name Qiu and Author's Last Name. Enhancing Financial Forecasting with Sentiment Analysis: A Review of Methodologies and Performance. *Journal or Publication Title*, 2020. Replace with actual details from Qiu (2020) publication.

[64] Kashif Qureshi and Tabassum Zaman. Social media engagement and cryptocurrency performance. *PLOS ONE*, 18(5):e0284501, 2023.

[65] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2019.

[66] Vanguard Research. Rational rebalancing: An analytical approach to multi-asset portfolio management. 2024. Validates threshold-based rebalancing for risk alignment in volatile markets.

[67] Sahib Mohamed Rida, Elkina Hamza, and Zaki Taher. From technical indicators to trading decisions: A deep learning model combining cnn and lstm. *International Journal of Advanced Computer Science and Applications*, 15(6):847–859, 2024.

[68] Moutsinga Seabe and Edson Pindza. Sentiment analysis and cryptocurrency price prediction using bi-lstm models: A comparative study. *Journal of Financial Technology Research*, 12(3):45–67, 2023.

[69] SecuX. Key factors that influence crypto volatility: What drives it? SecuX Blog, 2024.

[70] Sreelekshmy Selvin, R Vinayakumar, E. A Gopalakrishnan, Vijay Krishna Menon, and K. P. Soman. Stock price prediction using lstm, rnn and cnn-sliding window model. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1643–1647, 2017.

[71] V. Seshu, H. Shanbhag, and S.R. Rao. Performance analysis of bollinger bands and long short-term memory (lstm) models based strategies on nifty50 companies. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 344–349. IEEE, 2022.

[72] Jaimin Shah, Darsh Vaidya, and Manan Shah. A comprehensive review on multiple hybrid deep learning approaches for stock prediction. *Intelligent Systems with Applications*, 16:200111, 2022.

[73] William F. Sharpe. Mutual fund performance. *The Journal of Business*, 39:119–138, 1966.

[74] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[75] H Sim and S Kim. Deep convolutional neural networks for financial chart pattern recognition. *Journal of Financial Data Science*, 1(1):1–12, 2019.

[76] Hyun Sik Sim, Hae In Kim, and Jae Joon Ahn. Is deep learning for image recognition applicable to stock market prediction? *Complexity*, 2019:4324878, 2019.

[77] Liam Sinclair. Twitter sentiment analysis. `https://github.com/mliamsinclair/Twitter-Sentiment-Analysis`, 2024.

[78] Amandeep Singh and Amanpreet Kaur. Cryptocurrency price prediction analysis using machine learning algorithms. *SSRN Electronic Journal*, 2023.

[79] A. Smith et al. Dynamic financial forecasting with non-stationary data. *FinTech Journal*, 2023.

[80] R. Srinivasan. *Predicting Cryptocurrency Prices with Machine Learning: The Role of Moving Averages in LSTM Models*. PhD thesis, Stanford University, 2023.

[81] E. Stenqvist and A. Lonno. Predicting bitcoin price fluctuations with twitter sentiment analysis. 2017.

[82] ScienceDirect Research Team. A review on deep anomaly detection in blockchain. *ScienceDirect*, 2024.

[83] European Union. Long awaited eu ai act becomes law after publication in the eu's official journal. *EU Official Journal*, 2024. Regulation (EU) 2024/1689.

[84] A. Uzun and E. Yilmaz. Intelligent cryptocurrency trading: Leveraging ohlcv data for market momentum prediction. In *Proceedings of the International Conference on Artificial Intelligence in Finance*, pages 78–92. IAIF, 2023.

[85] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 2017.

[86] Aliaksei Vertsel and Mikhail Rumiantsau. Hybrid llm/rule-based approaches to business insights generation from structured data. In *Proceedings of Narrative BI*, 2024. team@narrative.bi.

[87] R. Wang and S. Li. Ai-enhanced pattern recognition in cryptocurrency technical analysis. *IEEE Transactions on Financial Computing*, 12(4):567–580, 2024.

[88] Jing Wei. Hybrid models for stock market forecasting using sentiment analysis. *Journal of Financial Analytics*, 2021.

[89] Moritz Wiese, Ralph Knobloch, and Ralf Korn. Generative adversarial networks for financial time series generation. *Quantitative Finance*, 20(9):1417–1436, 2020.

[90] Frank Xing. Designing heterogeneous llm agents for financial sentiment analysis. 2024.

[91] L. Xu, Y. Zhang, and H. Chen. Sentiment-aware anomaly detection in cryptocurrency markets. *Journal of Financial Data Science*, 5(2):78–95, 2023.

[92] Hongjun Yang, Lei Wang, and Peng Zhou. Deep reinforcement learning for financial trading using price trends. In *2018 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pages 1–6. IEEE, 2018.

[93] Huaqin Zhao, Zhengliang Liu, Zihao Wu, Yiwei Li, Tianze Yang, Peng Shu, Shaochen Xu, Haixing Dai, Lin Zhao, Gengchen Mai, Ninghao Liu, and Tianming Liu. Revolutionizing finance with llms: An overview of applications and insights. *arXiv preprint arXiv:2401.11641*, 2024.

[94] Tong Zhao. Efficiently learning robust, adaptive controllers from mpc. *MIT Department of Electrical Engineering and Computer Science*, 2023.

[95] Quan Zhou and Li Wang. Attention is all you need for sentiment analysis in financial news. *Journal of Computational Finance*, 15(2):53–67, 2021.