Enhancing Financial Forecasting through Transformer Models Using Social Media and News Insights

MSc Research Project
MSc in Data Analytics

Saiharsha Gurijala
Student ID: 23194341

School of Computing
National College of Ireland

Supervisor: Aaloka Anant

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Saiharsha Gurijala ………………………………………………………………………………………………
Student ID:	23194341 ………………………………………………………………………………………………
Programme:	MSc in Data Analytics ……………………………………………… **Year:** 2024 …………………………..
Module:	MSc Research Project ………………………………………………………………………………………………
Supervisor:	Aaloka Anant ………………………………………………………………………………………………
Submission Due Date:	12/12/2024 ………………………………………………………………………………………………
Project Title:	Enhancing Financial Forecasting through Transformer Models Using Social Media and News Insights ………………………………………………………………………………………………
Word Count:	6583 …………………………………… **Page Count** 22 …………………………………………………..

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Saiharsha Gurijala ………………………………………………………………………………………………
Date:	12/12/2024 ………………………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing Financial Forecasting through Transformer Models Using Social Media and News Insights

Saiharsha Gurijala
23194341

Abstract

Predicting the stock market correctly is challenging as it is dependent on corporate fundamentals, macroeconomic factors, and market sentiment. This study presents a new model combining Temporal Fusion Transformers (TFT) and sentiment analysis for stock price prediction. Inference is conducted based on the historical data observed for many structured features along with the time-sequenced sentiment obtained from discussions on Reddit Financial Sub-reddit and financial news, which is fetched to perform Finbert, a specialized NLP for financial data points. The prediction between the two Long Short-Term Memory (LSTM) models used here shows that the TFT model has the superpower to radically higher Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score than a traditional LSTM model. Furthermore, the model's interpretability capabilities, including attention layers and feature importance assessments, enhance understanding of the underlying drivers of stock price forecasts. These results underscore the importance of sentiment-driven features in terms of prediction accuracy, which renders their practical significance apparent. This research presents a scalable, interpretable approach to financial forecasting using advanced transformer-based architectures and unconventional data. This framework is a critical step toward connecting academic research with practical applications, offering tools for algorithmic trading and decision-making. This means their work is not only important for upcoming advances in AI in finance, but also for the use of AI more generally, as we try to find ways to use systems that are accurate but also interpretable to help us guide the muddle of complex modern finance.

Keywords: *Temporal Fusion Transformers (TFT), Long Short-Term Memory (LSTM), Financial Forecasting, FinBERT, Natural Language Processing.*

1 Introduction

Financial markets are fundamentally unpredictable, influenced by an intricate web of fundamental drivers like corporate earnings and macroeconomic indicators and non-fundamental drivers like public confidence and exogenous events. Forecasting the stock price accurately is most important for well-informed investment strategy, risk management as well as algorithmic trading and is an active area of research. Conventional techniques, such as ARIMA and regression-based models, tend to rely extensively on historical data and lack the capability of modelling the nonlinear relationships and abrupt changes that are often observed in financial markets (Zhang & Zheng, 2021). Machine learning and deep learning have enhanced forecasting accuracy by detecting latent interactions in financial data; however, these approaches are often limited by a lack of natural merging of different data sources– the inclusion of unstructured (typically sentiment) across multivariate time-series models is a common limitation (Yang & Kostadinov, 2020).

However, recent developments of architectures including Temporal Fusion Transformers (TFT) provide a way to overcome these limitations. By leveraging attention mechanisms and feature importance analyses (Lim et al., 2021), TFT, a market-leading DL model, is able to capture complex temporal dependencies but still provides more interpretability. Additionally, sentiment data from Reddit and financial news articles, analysed using FinBERT, an NLP model designed for financial content, improves the predictions by combining standard financial metrics with alternative data, making the predictions more practical and useful.

1.1 Motivation

Classic forecast models are built on more traditional datasets and therefore provide a reasonably complete picture of the world but neglects the impact of alternative data including public sentiment. Web sites like Reddit mirror investor sentiment, while news coverage provides expert narratives to describe market activity. Good things come from these data sources that can add to the accuracy and interpretability in forecasting. By combining sentiment insight with advanced ML approaches like TFT, this research aspires to uncover latent patterns, improve prediction accuracy, and furnish implementable algorithmic trade signals.

This need is driven by the increasing demand for adaptive, transparent, and comprehensible forecasting models that leverage alternative data to overcome the limitations of traditional methods (Huynh & Nakamori, 2019). This study demonstrates how incorporating sentiment-based features into structured financial data models can enhance prediction accuracy and facilitate better decision-making.

1.2 Research Question

How can transformer-based models, combined with alternative data like social media sentiment and news analytics, improve the accuracy of real-time financial time series predictions and enhance adaptive algorithmic trading strategies?

1.3 Research Objectives

The key objectives of this research are:

- **Enhancing Predictive Model:** Evaluate the effectiveness of Temporal Fusion Transformers in improving the accuracy and interpretability of financial forecasts.
- **Integration of Alternative Data:** Demonstrate the value of combining structured financial data with unstructured sentiment information to enrich forecasting processes.
- **Practical Applications:** Bridge the gap between academic research and real-world algorithmic trading, offering a scalable framework for adaptive decision-making.

1.4 Overview of the Proposed Approach

This study combines structured financial data with unstructured sentiment data obtained from posts on Reddit and articles on financial news. We perform sentiment analysis on these using

FinBERT, a financial domain-specific trained language model, to extract actionable insights. To model multivariate time-series data with complex temporal dependencies and multiscale patterns, we use the Temporal Fusion Transformer. Performance of the proposed model is compared with LSTM networks. To enable real-time predictions, visualizations, and trading strategy simulations, a web application based on Flask library has been developed.

1.5 Structure of the Thesis

The thesis is organized into sections that systematically address the research objectives. In Section 2, we provide a literature review of relevant literature, covering aspects of financial forecasting, sentiment analysis, transformer-based models and pointing out research gaps. Section 3 provides the methodology, including data collection and preprocessing, model development and evaluation strategies. In Section 4, we present the design and implementation of the proposed framework where its architecture and algorithms are discussed. Section 5 assesses the performance of the model using comparative analysis and interpretability metrics. In this final section, we draw together the main findings of this thesis, talk through the implications, and suggest pathways forward for financial forecasting & algorithmic trading research.

2 Related Work

This segment covers the evolution of the financial forecasting exercise, emphasising on important methodologies, challenges and opportunities. The overview includes classical methods, the role of sentiment analysis, developments in deep learning, and new transformer models, specifically the Temporal Fusion Transformer (TFT). They also talk about incorporating alternative data and the role of explainability in financial models. Finally, this review highlights the knowledge gaps and potential areas for future work.

2.1 Traditional Financial Forecasting Methods

Methods such as autoregressive integrated moving average (ARIMA) and exponential smoothing are based on preceding history, and also assume linear relationships and stationarity. They prove effective in short-term predictions but fall short in modelling nonlinear dependencies and external factors such as market sentiment (Zhang & Zheng, 2021). Some of these limitations were overcome with the introduction of statistical machine learning models like Support Vector Machines (SVM) and Random Forests, for example, which handled nonlinearity and high-dimensional data. But such models are not temporal and not adaptive in dynamic financial environments (Huynh & Nakamori, 2019).

2.2 Sentiment Analysis in Financial Markets

Sentiment analysis has proved to be an impactful way of capturing market behaviour on the back of unstructured data from social media and financial news. Public sentiment's influence over short-term stock price fluctuations is well documented (Bollen et al., 2011), making it a useful input for predictive models. Models specifically fine-tuned on financial text, such as FinBERT, have also been developed, which provide improved accuracy of the sentiment classification compared to generalised models. These advancements, however, do not yet include challenges in fusing sentiment data with structured financial metrics to create reliable forecasting models (Yang & Kostadinov, 2020).

2.3 Deep Learning in Financial Forecasting

The most popular one among those, the novel LSTM networks adapted for multivariate time series with their strong sequential dependency capturing capabilities and forecast performance reshaped the landscape of financial forecasting. LSTM performs better than competitive baselines on stock price prediction tasks (Fischer & Krauss, 2018). Newer architectures, such as Gated Recurrent Units (GRUs) and attention mechanisms, have improved both the temporal modelling and the computational efficiency immensely. But these models are not scalable, interpretable and primarily do not capture the integration of heterogeneous data sourced.

2.4 Transformer Models in Financial Forecasting

Transformers were first introduced for tasks in the field of natural language processing (NLP), where they have achieved state-of-the-art performance on a variety of sequential data tasks. This allows them to attend to the most salient features, which helps with predictive accuracy and interpretability (Vaswani et al., 2017). Recent research Transformers have been adapted for their utilization in financial forecasting use-cases, improving the ability to appropriately handle complex temporal dependencies in multivariate time-series. However, since the implementation of Transformers in finance is still an emerging field, there is still much room for in depth work within the area (Lim et al., 2021).

2.5 Temporal Fusion Transformer (TFT) and Applications

Temporal Fusion Transformers applies attention mechanisms and gating networks to attend to relevant features and relevant times and formulates time-series forecasting as a supervised learning problem. This characteristic makes it particularly appropriate for financial forecasting, significantly outperforming standard models like LSTMs and GRUs concerning predictive power and interpretability (Lim et al., 2021). Since TFT has not limited itself into stock price prediction, it has also been also considered in energy and healthcare facets. And, in finance, its feature importance score helps increase the interpretable decisions which is very important when making decisions.

2.6 Integration of Alternative Data in Forecasting

Alternative data has served to fill the gap. for example, sentiment extracted from Reddit and financial news articles have been used to complement quantitative metrics with qualitative insights in financial forecasting. They offer significant insights into the behaviours of markets, enabling predictions that are more realistic and contextual (Huynh & Nakamori, 2019).

Recent studies, which highlight, through the application of sentiment analysis with advanced ML techniques, the transformative power of this approach in understanding society, are summarized in Table 1. Although some progress has been made, the harmonization of disparate data sources remains a challenge in building scalable forecasting frameworks.

Authors	Data Source	Sentiment Data and Applications	Model Used	Key Findings	Future Work
Bollen et al. (2011)	Twitter	Social media sentiment for short-term stock predictions	ARIMA with sentiment	Sentiment improves short-term stock price prediction	Explore long-term prediction capabilities
Devlin et al. (2019)	Financial news	Improved sentiment classification for financial text	FinBERT	Challenges with complex jargon	Refine pre-training for financial-specific terminology
Yang & Kostadinov (2020)	Financial news	Domain-specific sentiment for analyzing company-level performance	FinBERT + LSTM	Domain-specific sentiment enhances model performance	Test with additional unstructured data sources
Lim et al. (2021)	Multivariate time series	News and time-series sentiment to improve temporal dependency modelling	Temporal Fusion Transformer (TFT)	TFT outperforms LSTM in accuracy and interpretability	Extend to real-time applications and trading setups
Chae et al. (2023)	Financial news, Reddit	Effective domain-specific sentiment analysis for stock forecasting	FinBERT + LSTM	Limited handling of unstructured data	Expand to incorporate multiple unstructured data sources.

Table 1: Comparative Analysis of Recent Research Studies

The studies summarised above collectively indicate the potential of alternative data and sentiment in predictive modelling. Augmenting traditional numeric data by integrating both structured and unstructured alternative data like news, social media sentiment, helps to construct new predictive frameworks improves both accuracy and interpretability.

2.7 Gaps in Current Research

However, despite this enormous progress, financial forecasting is still devoid of many gaps. First, despite high accuracy for making predictions obtained by LSTM and TFT models, their interpretability remains questionable (Huynh & Nakamori, 2019), making them less applicable for real one-time decisions. Second, other data sources like including social media and news sentiment, have yet to find their place in the structured world of financial metrics. Indeed, the current studies are either oriented to one type of data & they miss two or more independent but complementary datasets for holistic forecasting (Yang and Kostadinov, 2020). As financial markets evolve, they present challenges for common data streams in real time, & more research is needed if financial models are to be applied under these conditions.

Conclusion

The review charts the transformation of financial forecasting from basic econometric modules to advanced deep learning architectures like TFT. Social media or news sentiment are alternatives data sources allowing one to understand predictive process in a more accurate and interpretable way. Model interpretability, data integration, and real-world applicability remain problem areas. There are critical gaps to move the field forward. Against this backdrop, our work aims to fill the existing gap with a novel approach leveraging TFT and domain-specific sentiment analysis for risk and return forecasting that is accurate, interpretable, and adaptable to academia and practice.

3 Research Methodology

This section of this study documents the structured methodology used and consists of the research design, data collection, preprocessing, model development, evaluation, deployment, and ethical considerations. This work follows the research objectives by combining state-of-the-art machine-learning techniques with structured financial data and other alternative sources, like sentiment analysis, to obtain robust and interpretable forecasting results in the financial domain. The methodology is summarized in Fig 1, which illustrates the sequential flow of the research process, from data collection to deployment and evaluation.

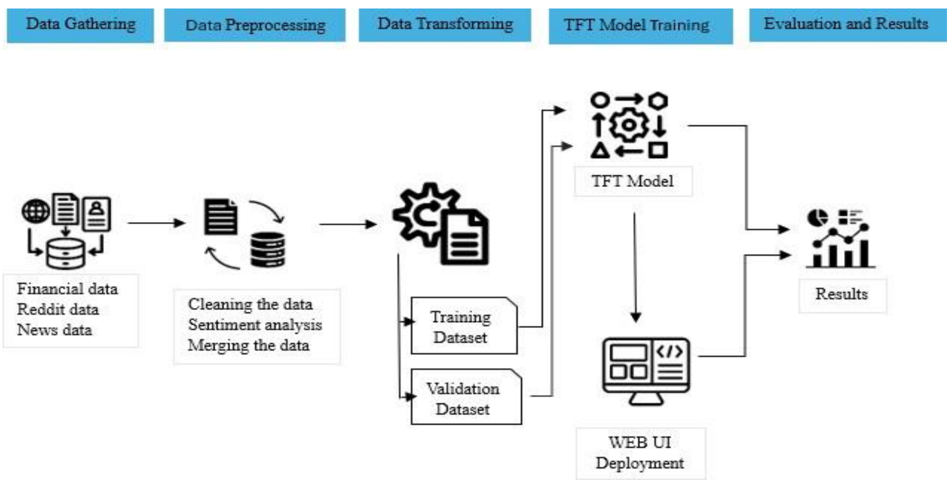

Fig. 1. Research Methodology

3.1 Research Design

This study follows a quantitative research design by utilizing the TFT for multivariate time-series forecasting. We chose this model because it has the capability to combine different types of datasets including history of the stock, static covariates (such as stock identifiers) and dynamic covariates like sentiment scores, volume of the stock. The research design

provides another dimension to cut through to stock price forecasting by integrating structured financial metrics and unstructured sentiment data from Reddit and financial news.
For performance benchmarking, a LSTM model was created and set as the baseline for comparison regarding the accuracy of predictions as well as the interpretability of the model. This design focusses on applicability by employing a hybrid framework to capture the inherent dynamics between financial and sentiment data. The results are visualized in an interactive web application where users can play with predictions, gain insights and simulate trading lessons, providing a connection between academia and practice.

3.2 Data Collection

Data was sourced from multiple platforms to ensure a analysis of market dynamics:

1. Financial Data: Stock price data for Tesla and Apple, including the open price, the close price, the daily high and low prices as well as the volume of trades. Data extracted from **Yahoo finance** covering **2018-2024** would provide enough history to model long-term distributions while still capturing short-term noise on each stock's performance. Such data lends itself to that of alternative datasets, and provides a structured foundation for how financials relate to external factors such as sentiment.

2. Reddit Data: Reddit data was sourced from Kaggle datasets and dynamically collected using the PRAW library to capture stock-related discussions:

- **Historical Data (2018-2022):** Reddit posts discussing financial markets and stocks were obtained from a Kaggle dataset, which provided comprehensive historical data on stock-related discussions. This dataset included key features such as post titles, content, and timestamps, offering a detailed view of past investor sentiment.
- **WallStreetBets Data (2022 - November 2024):** Posts from the WallStreetBets subreddit were gathered from another Kaggle dataset. These posts focus on influential discussions often tied to significant market movements, providing insights into trends.
- **Dynamic Data (December 2024 - Present):** The latest data is dynamically collected using the **PRAW library**, capturing up-to-date posts discussing Tesla, Apple, and other financial topics in real-time.

The collected data, including titles and text content, was analysed using sentiment analysis tools like FinBERT. This sentiment analysis enabled the identification of retail investor behaviour, opinions, and trends, offering valuable insights into market dynamics over time.

3. Financial News Data: Sources of data for financial news articles are based on the Global Database of Events, Language, and Tone **(GDELT),** which is a widely recognized polarity database with extensive coverage of news articles from reputable sources. Articles range from 2018 to November 2024 in the dataset, and further news data from December 2024 is collected dynamically using the **News API**. As there would be several articles on Tesla and Apple, these articles were filtered to only selective topics on these companies, such as company events, macroeconomic trends, and overall analysis. Sentiments were moted out of the Title and abstract down the analysis through the FinBERT Analysis, where the sentiment scores provide insight into the narratives written by professionals on market movements and investors interpretation of sentiment. The dataset allows for insight into how structured financial metric models can be aligned with unstructured sentiment data from the retail investor perspective, providing a multi-dimensional view of the relationship between semi-structured news angles and performance outcomes as actionable information. Such a dual vision may also contribute to a more nuanced appreciation for market dynamics, leading to improved predictive activity in the quest for better financial forecasting models.

3.3 Data Preprocessing

Clean, transform and integrate raw data for modelling. Financial data was cleaned by removing duplicates, and missing values were filled using linear interpolation and the forward-fill method. The research first assessed sentiment movement through FinBERT sentiment analysis by using it to extract sentiment scores (positive, negative, and neutral) from both Reddit posts and financial news headlines. Scores were aggregated on business days to match sentiment data with stock market activity. Last, the structured financial data and the unstructured sentiment data were combined to form a complete full dataset while ensuring the time alignment for the analysis and predictive task.

3.4 Model Development

Two models were selected to explore predictive capabilities:

- **Temporal Fusion Transformer (TFT):** The Temporal Fusion Transformer (TFT) was selected for its capability to process multivariate time-series features, and at the same time providing interpretability through attention mechanisms and importance of contribution analysis. The model contextualizes structured financial information with fluid sentiment from text data revealing latent temporal dependencies. And advanced features like temporal attention weights and dynamic feature selection flexibility lotted the model align rapidly with changes in market dynamics.
- **Long Short-Term Memory (LSTM):** As a baseline for comparison, the Long Short-Term Memory (LSTM) model was used. The LSTM model is specifically created to learn temporal dependencies so our data on stock prices provided per sample was scaled and had a sequence length of 90 days which allowed for the model to learn about past c trends and associations. It assessed the predictive power and interpretability of an advanced model by benchmarking the results of the TFT against the LSTM.

3.5 Evaluation Metrics

The models are going to evaluated using a combination of practical & statistical metrics:

- **Statistical Metrics:** Mean Absolute Error, Root Mean Squared Error and R^2 Score
- **Economic Feasibility:** The models were assessed for economic viability using simulated trading strategies with a 30-day horizon. Predictions were employed to signal decisions whether to buy or sell assets and cumulative returns were calculated based on these signals. The percent profit, determined against a simple initial investment, gave crude indications of potential financial returns. The analysis also accounted for return volatility to make sure that profitability metrics reflected real-uplift, market dynamics.

3.6 Ethical Considerations

The study followed ethical guidelines, including keeping Reddit user data anonymous and following privacy and data protection laws. Model predictions were for research, and not financial trading. All the datasets used were obtained from publicly available sources, in accordance with their terms of use. It demonstrates that the research is conducted with proper academic methodology and also in accordance with ethical considerations of usage of tech and data for financial research.

Summary

In this paper, we introduce a new method that utilizes the state-of-the-art ML algorithms with robust data preprocessing to provide accurate and interpretable prediction for financial data. By integrating structured finance data with unstructured sentiment data obtained from Reddit and news, the study presents an enhanced framework for market trend analysis. Comparing with the LSTM model as the baseline, we evaluate the interpretability and dynamic feature selection performance of the TFT model. The focus of the methodology is on practical relevance, with the use of statistical evaluation metrics and real-world simulated orders. We are combining best of the two worlds with this strategy like real time financial insight with decade-long extensive academic literature about informed trading and decision making.

4 Design Specification

This section illustrates the architecture design and implementation structure of proposed financial forecasting framework by incorporating the TFT. It brings together structured financial metrics and unstructured sentiment data in manner that produces a model that is simultaneously both strong and scalable, as well as interpretable with respect to actionable insights. Such a framework could provide accurate forecasts using the latest in time-series modelling techniques, while also providing models ready for deployment. This section is divided into critical sub-sections, each capable of being read independently from one another to demonstrate the flexibility and breadth of the framework.

4.1 Overview of the TFT Model Framework

This work is primarily built upon the TFT which has shown stellar performance on stock forecasting attributed to its novel architecture that allows the dynamic integration of static covariates (stock identifiers), dynamic covariates (sentiment scores and volume, open values of the stock), and historical time-series (stock prices). The model is interpretable via feature importance analysis and temporal attention, but it also captures complex temporal relationships. Architecture binding variable selection networks (VSN), gated residual networks (GRN), temporal attention mechanism and position-wise feed-forward layers to adaptively jointly learn the representation of financial and sentiment data. Such attributes ensure magnificent predictability performance and interpretability, yielding probabilistic forecasts and offering robust representations of market mechanics.

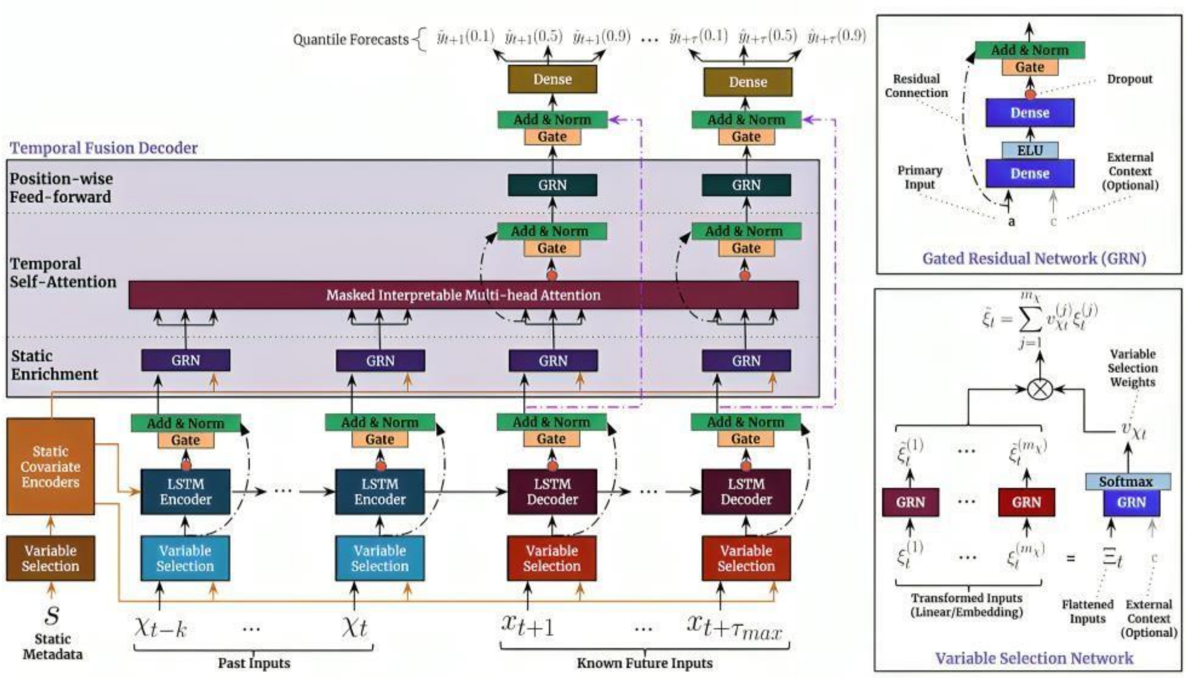

Fig 2: Temporal Fusion Transformer Architecture (Lim et al., 2021).

The **architectural design of the framework** is modular, allowing uniform integration of diverse data sources and adaptive forecasting capabilities. Key components are:

- **Temporal Fusion Transformer (TFT):** The TFT is introduced for modelling intricate, multivariate time-series data with both static/dynamic covariates. Its feature selection methods and self-attention mechanisms are adaptable to the most important features and time steps, allowing it to learn multi-order dependencies.
- **Static Covariate Encoders:** This part of the model encodes metadata (for example: stock identifiers) to improve the model/harmonizer contours. Static features complement the temporal dynamics to help the model to learn long-term trends and relationships.
- **Variable Selection Networks (VSNs):** VSNs use Gated Residual Networks (GRNs) to dynamically select and prioritize the features that will have the largest impact at each time step. This allows the model to focus on relevant inputs, such as trading volume and sentiment data, and to filter out noise.
- **Temporal Attention Mechanisms:** The Integrators employed a masked multi-head self-attention mechanism that emphasizes influential time periods so that the model captures both short- and long-term patterns in temporal financial and sentiment data.
- **Quantile Forecasting Outputs:** TFT generates probabilistic forecasts where P10, P50, and P90 quantiles represent a range of potential outcomes. This allows stakeholders to make risk-aware decisions under different levels of prediction uncertainty.

4.2 Key Features of Temporal Fusion Transformer

The TFT's advanced capabilities are designed to financial forecasting by the following features:

- **Adaptive Feature Selection:** VSN performs feature selection dynamically, permitting the model to efficiently adapt to dynamic market conditions.

- **Temporal Attention Mechanisms:** The self-attention layers help intercept which time periods are most influential to the predictions and improve the model at conversing with important trends and events occurring in the market.
- **Static Enrichment:** Static covariates like stock identifiers ensure that the model retains long-term forecasting accuracy and contextual understanding.
- **Probabilistic Predictions:** Quantile-based predictions provide the full picture of uncertainty in forecasts, aiding decisions in high-stake financial applications.

4.3 Temporal Fusion Transformer (TFT) Model Architecture

The architecture introduced in Fig 3, upon which we base this research framework, is the Temporal Fusion Transformer (TFT). This integrates structured data representations on financial data, unstructured representations of sentiment data, even temporally, to give precise estimates of what a stock price will be, which are interpretable and probabilistic.

Overview: The TFT model is specifically designed to process multivariate time-series data by dynamically adjusting to temporal dependencies and integrating diverse features. Refer to Fig. 3 for detailed insights into its architecture.

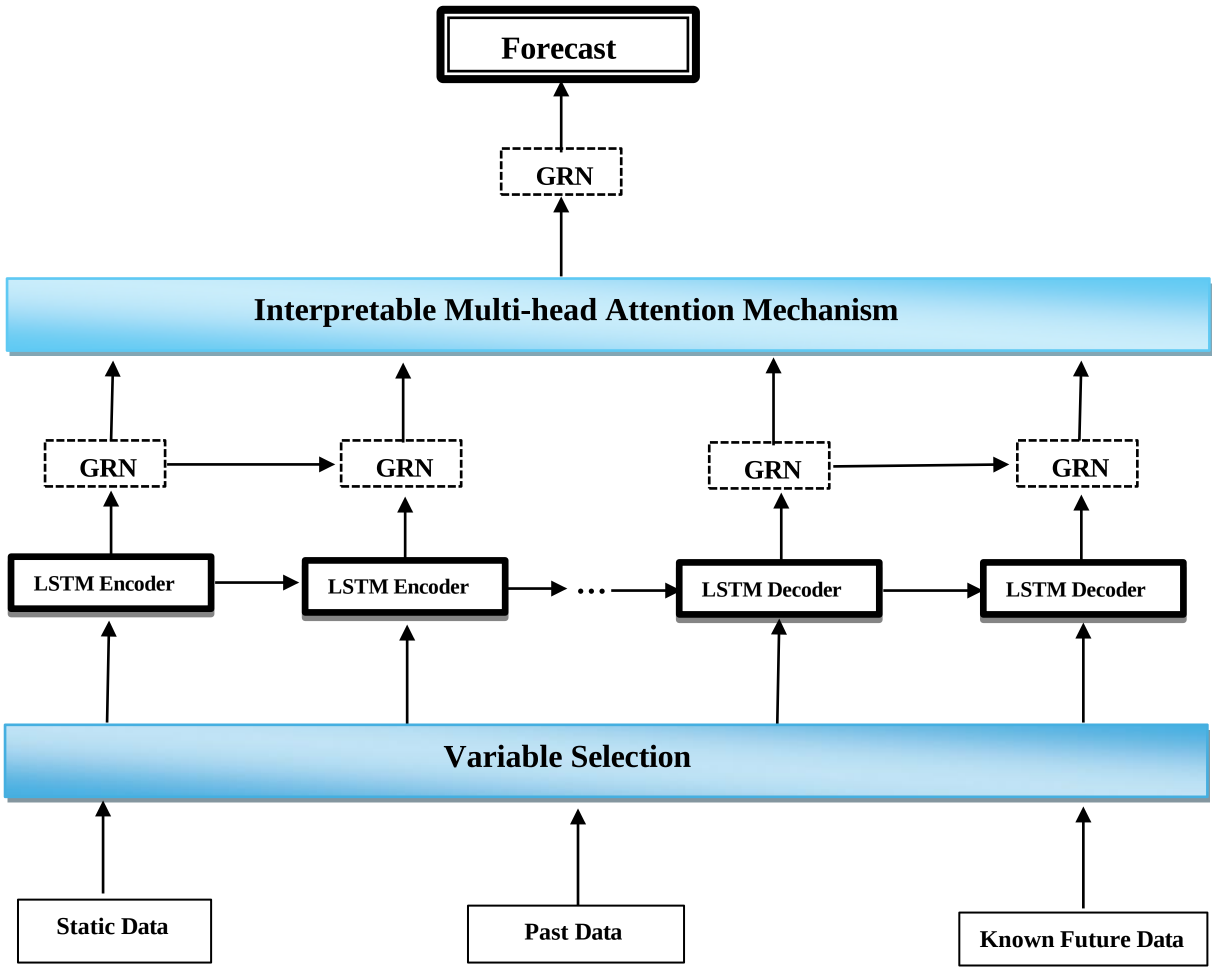

Fig 3: Temporal Fusion Transformer Architecture

Its unique architecture is modular & scalable, and adopts state-of-the-art mechanisms like VSNs, GRNs, and Temporal Attention Mechanisms for strong forecasting & interpretability.

Key Components of the Architecture:

Input Layers:

- **Static Data:** Encodes metadata like stock identifiers using techniques such as NaNLabelEncoder. These static covariates add contextual information to the model.
- **Past Data:** Includes historical stock prices, trading volumes, and sentiment scores aggregated on business days. Lagged features are engineered to capture historical trends.
- **Known Future Data:** Temporal features like day of the week and covariates such as predicted trading volume are incorporated to enhance forecasting accuracy.

Variable Selection Networks (VSNs): An effective attention mechanism approach is to implement Variable Selection Networks (VSNs) to allow the models to dynamically select important features at every time step, such as the sentiment signals and trading volume which have higher impacts and avoid useless feature inputs. This way VSNs filter the noise and the less significant variables, making the model more adaptable to market conditions variations thus enhancing the robustness and accuracy of financial prediction.

LSTM Encoder: Extracts the temporal dynamics and long-term dependency from past 180 days historical data. The encoder allows capturing the sequential trends, which is essential to understanding price behaviour and market behaviour.

Temporal Attention Mechanism: Identifies key time periods within the historical data that considerably influences the prediction. Such as spikes of sentiment or anomalies in trading volumes. This allows for interpretable results by pinpointing events or trends responsible for movements in stock prices.

LSTM Decoder: Predicts the representations learned by the encoder using future information that is also known. This element creates accurate predictions by merging historical insights with prospective trends.

Gated Residual Networks (GRNs): Enhance feature representation and transformation throughout the architecture. GRNs ensure efficient learning of dependencies between features.

Interpretable Multi-Head Attention: Provides interpretability by assigning weights to different features and time periods, showing which inputs contribute most to predictions. This helps in understanding the importance of sentiment data, trading volume, or historical price trends in the forecasting process.

Output Layer: Produces probabilistic forecasts for stock prices, represented by quantiles such as P10, P50, and P90. This allows stakeholders to assess potential risks and uncertainties.

With consideration of these design goals, TFT architecture aligns closely with the scope of this research regarding robustness, interpretability and scalability. This allows the model to make accurate predictions while avoiding information overload by focusing on only the relevant features, hence adapting to the dynamic nature of the finance domain. It is highly interpretable and provides transparency by offering attention mechanisms and feature importance scores that highlight factors that have a high influence on stock prices being predicted. The modular design of TFT enables easy integration of new dataset and features, therefore offering scalability in the long term and flexibility for changing financial forecasting use cases.

5 Implementation

We describe in this section the implementation and integration of the financial forecasting framework that we are building that uses the Temporal Fusion Transformer (TFT) and sentiment analysis to generate accurate, interpretable and actionable stock price forecasts. The section also includes a web-based interface for real-time interaction with and visualization of the model.

5.1 Data Preparation and Integration

The first step was to prepare structured financial data to be coupled with unstructured sentiment data into a single dataset. This step also assured that all relevant data were temporally aligned and pre-processed suitably for feeding into TFT model.

Sentiment Analysis:

- Sentiment scores (positive, neutral, and negative) were generated using FinBERT, a financial-domain-specific transformer model.
- Sentiment analysis was conducted separately for Reddit post titles and content, as well as financial news article titles.
- For each business day, the maximum sentiment score (positive, neutral, or negative) was computed to represent the dominant sentiment of the day, capturing significant sentiment signals for each timestamp.
- The aggregated sentiment data was merged with structured financial data based on the timestamp to ensure temporal alignment.

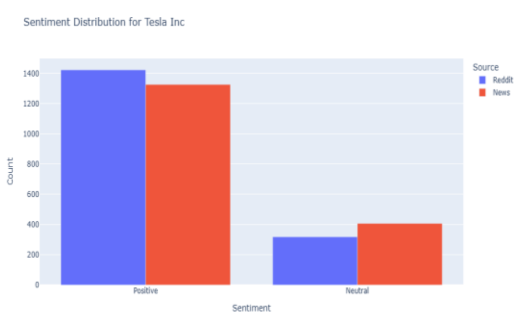

Fig 4. Sentiment Distribution for Tesla

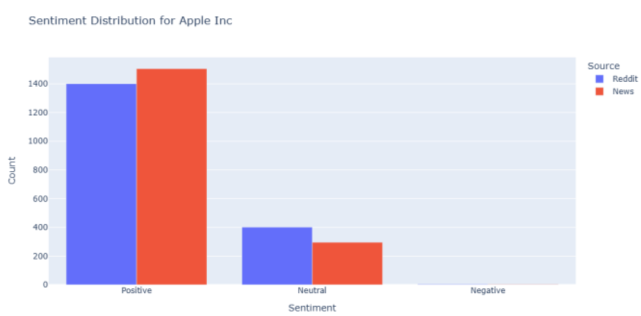

Fig 5. Sentiment Distribution for Apple

The resulting dataset captured a combination of historical stock metrics and sentiment trends, enabling rich inputs for the predictive modelling process.

5.2 Model Development and Training

The key part of this implementation had been the development and training of the TFT model on stock price data of Tesla (TSLA) and Apple (AAPL). Incorporating the advantages of the LSTM model, which is capable of learning the long-term dependencies of sequential data, and provides interpretable predictions, further solidified its choice for a multivariate time-series framework for financial forecasting.

5.2.1 Model Configuration

Dataset Construction:

The final dataset for training included:

- **Static features:** Stock identifiers, such as group_id, which uniquely represent the stock.
- **Time-varying known features:** Opening prices, trading volume, and engineered temporal features such as the day of the week.
- **Time-varying unknown features:** Closing prices, Reddit sentiment, and news sentiment.
- **Lagged features** and temporal variables were engineered to capture historical patterns and dependencies, such as lagged closing prices for 7, 15, 30, 60, and 90 days.
- **Missing values** were handled using forward-filling and linear interpolation techniques.

TimeSeriesDataSet Creation: The ‘TimeSeriesDataSet’ class from PyTorch Forecasting was used to prepare datasets for training and validation:

- The maximum encoder length was set to 180 days to analyse extensive historical trends.
- The prediction length was set to 30 days for short-term forecasting.
- Sentiment features were encoded using ‘NaNLabelEncoder’, and the target variable (closing price) was normalized using ‘GroupNormalizer’ with a softplus transformation.

5.2.2 Hyperparameter Tuning

Optimization with Optuna: Parameters such as learning rate, hidden size, attention head size, and dropout rate were optimized using Optuna, aiming to minimize the validation loss.

Best Hyperparameters: Got the best parameters with the optimization of Optuna. Refer the Table 2 for more information:

Table 2: TFT Model Best Hyperparameters for Tesla, Apple

Parameters	TSLA	AAPL
Learning rate	0.0007693367817925835	0.000760235330401722
Hidden size	24	41
Attention head size	3	3
Dropout	0.19608882013889972	0.1383927412069838

5.2.3 Model Training

Training Configuration:

- The Tesla model was trained for **150 epochs**, and the Apple model was trained for **94 epochs**. Early stopping with a patience of 15 epochs was applied to avoid overfitting.
- The Tesla model included **85.7K trainable parameters**, while the Apple model had **158K trainable parameters**, reflecting the complexity and variability in their respective datasets.

Dynamic Feature Selection: The TFT model utilized attention mechanisms to dynamically select relevant features and time periods, focusing on critical inputs like sentiment scores and significant price movements.

Evaluation and Metrics: Validation losses achieved were **4.263** for Tesla and **1.099** for Apple. Visualization of predicted vs. actual stock prices highlighted the models' ability to capture trends accurately.

5.3 Tools and Technologies

The implementation utilized a range of tools and technologies to achieve its objectives:

Machine Learning and Forecasting:

- **PyTorch Lightning:** Used for training and managing the TFT model. This framework provided a robust platform for scalable deep learning experimentation.
- **PyTorch Forecasting:** Facilitated the creation of TimeSeriesDataSet, enabling efficient handling of multivariate time-series data & dynamic feature selection for the TFT model.
- **Optuna:** Applied for hyperparameter tuning to optimize the learning rate, hidden size, dropout rate, and attention head size for the TFT model.
- **TensorFlow/Keras:** Used for developing the baseline LSTM model to benchmark against the TFT model.

Natural Language Processing (NLP):

- **Transformers:** The FinBERT model, a domain-specific transformer for financial text, was used for sentiment analysis of Reddit posts and financial news articles.
- **Hugging Face Pipeline:** Enabled sentiment extraction in a streamlined manner using FinBERT.

Data Processing and Visualization:

- **Flask:** A lightweight web framework used to build the web application for user interaction.
- **Ngrok:** Allowed secure and public access to the locally hosted web application during development and testing.

Hardware: NVIDIA GPU for accelerating model training.

5.4 LSTM Model Implementation

A Long Short-Term Memory (LSTM) model was implemented as a baseline for comparison with the TFT model. The following outlines the LSTM model's architecture and integration:

Model Architecture:

- **First LSTM Layer:** 100 units with return sequences enabled.
- **Dropout Layer:** 25% dropout rate to prevent overfitting.

- **Second LSTM Layer:** 50 units for final feature extraction.
- **Dense Layers:** Includes a 25-unit dense layer for intermediate representation and a single-unit output layer for stock price predictions.

Integration into the Framework:

- The LSTM model was implemented using TensorFlow/Keras and trained on Tesla and Apple datasets.
- It was incorporated into the web application, allowing users to select the LSTM model for predictions and compare its performance with the TFT model.

By incorporating the LSTM model, the framework provided a robust baseline for evaluating the TFT model's performance, emphasizing its strengths in interpretability and predictive accuracy.

5.5 Web Application Deployment

A Flask-based web app was developed to allow users to interact with the trained models. The key features include:

- **Stock Predictions:** Users can input a stock ticker (TSLA or AAPL) and select a model (TFT or LSTM) to view predicted vs. actual stock prices graph and metrics.
- **Model Comparison:** Users can compare the performance of the TFT and LSTM models based on metrics like MAE, RMSE, and R^2 score.
- **Trading Strategy Simulation:** A feature simulates trading strategies based on TFT predictions, providing buy/sell signals, cumulative returns, and profit percentages.

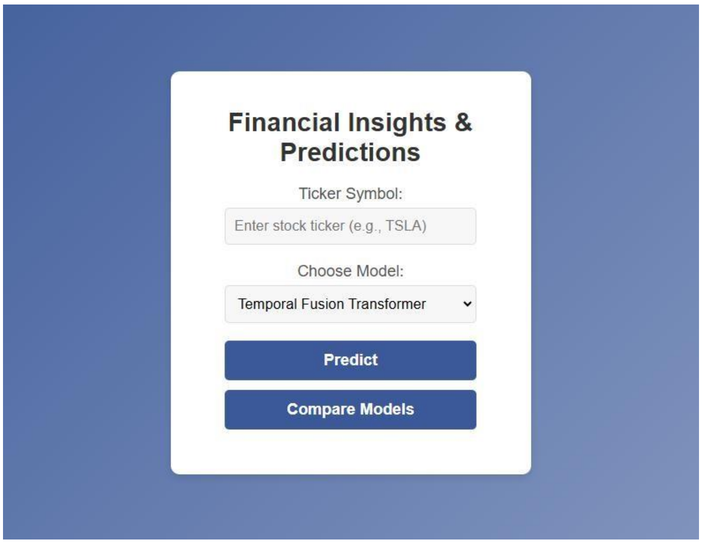

Fig 6. Home Page of Financial Predictions Website

This user-friendly interface (UI) bridges the gap between advanced ML models & practical financial decision-making, showcasing the integration of advanced transformer-based architectures with sentiment analysis to deliver interpretable, and user-centric financial forecasting solutions.

6 Results and Discussion

The following sections discuss the execution of the framework and outlines the advantages and disadvantages of the results. The analysis is based on statistical results and visualizations that serve as proof of the model performing well for financial forecasting. The results are evaluated against the research objectives, complemented by the graphs and metrics that illustrates the effectiveness of the framework.

6.1 Temporal Fusion Transformer (TFT) for Tesla (TSLA)

The TFT model was trained and validated for Tesla stock price predictions using a dataset combining financial metrics and sentiment scores. The model delivered the following performance metrics:

- **Mean Absolute Error (MAE):** 8.04
- **Root Mean Squared Error (RMSE):** 10.47
- **R^2 Score:** 0.95

Predictions vs. Actual Values Plot: The graph shows the alignment of predicted & actual stock prices during validation, highlighting the model's capability to accurately capture price trends.

Fig. 7. Tesla TFT Model Plot

Feature Importance: The feature importance analysis for Tesla indicates that static features like group_id and encoder features such as volume had the highest influence on predictions, followed by decoder features like reddit_sentiment.

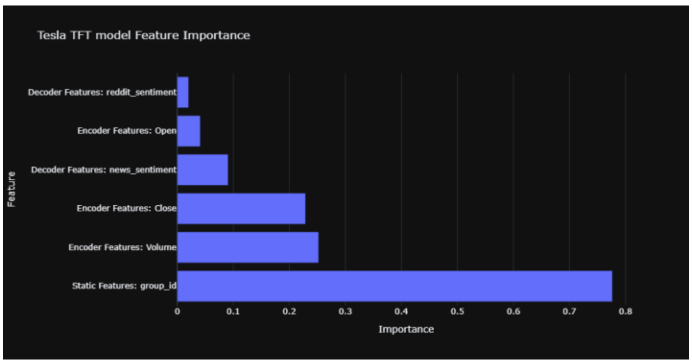

Fig. 8. Tesla TFT Model Feature Importance

6.2 Temporal Fusion Transformer (TFT) for Apple (AAPL)

The TFT model was evaluated for Apple stock price predictions, yielding the following metrics:

- **Mean Absolute Error (MAE):** 2.07
- **Root Mean Squared Error (RMSE):** 2.71
- **R² Score:** 0.63

Predictions vs. Actual Values Plot: This plot illustrates the model's ability to capture short-term trends in Apple stock prices, albeit with slightly reduced precision compared to Tesla.

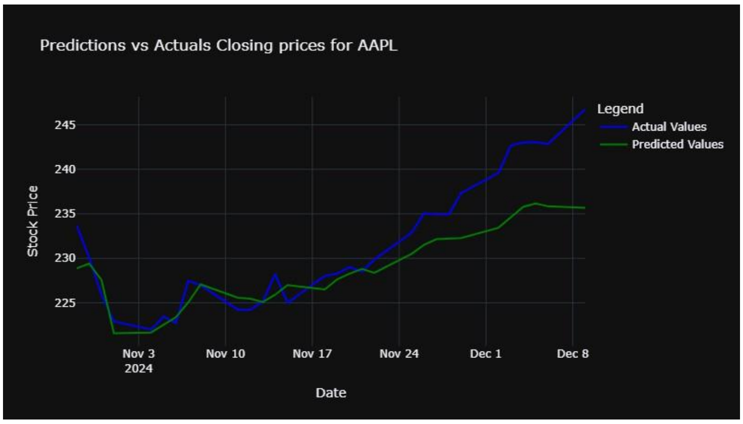

Fig. 9. Apple TFT Model Plot

- **Feature Importance:**

The analysis for Apple highlights reddit_sentiment as the most impactful feature, followed by news_sentiment and volume. This reinforces the significance of alternative data in stock price prediction.

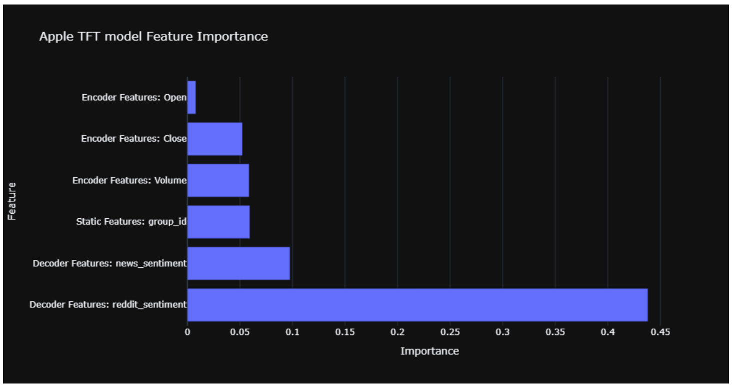

Fig. 10. Apple TFT Model Feature Importance

6.3 Trading Strategy Simulation

In order to evaluate the practical utility of the TFT model, we performed trading strategy simulations for Tesla and Apple stocks. These models were tested in different simulations, demonstrating their predictive accuracy as well as its significance for more profitable future trades.

Tesla Trading Simulation: The simulation for Tesla utilized the TFT model's predictions to generate buy and sell signals based on stock price movements. The Graph of the simulation as follows:

Fig. 11. Tesla TFT Model Trading Strategy Simulation Results

Apple Trading Simulation: Similarly, the Apple trading simulation employed the TFT model's predictions to assess its performance. Results include:

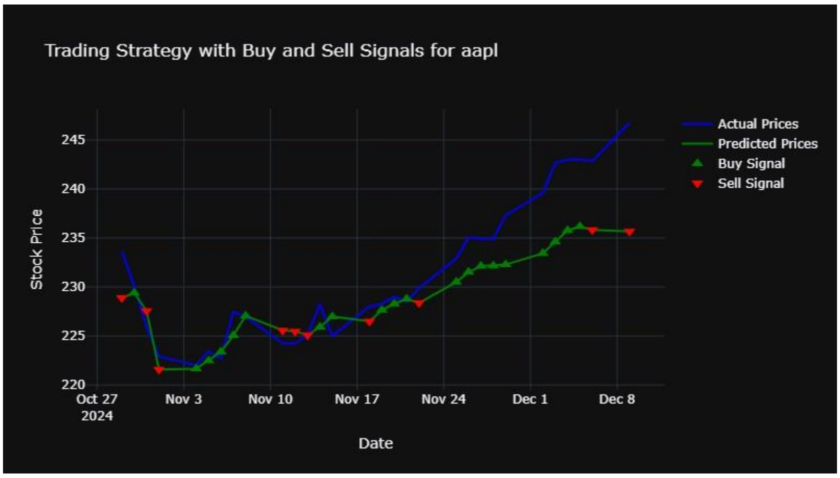

Fig. 12. Apple TFT Model Trading Strategy Simulation Results

Summary

These simulations illustrate its capacity to convert predictions into actionable investment strategies, and demonstrate the model's potential utility for financial decision-making. And while Tesla scored big profits, Apple's results suggest that now may not be the right time to buy the stock. As usual, it is better to wait and see rather than following trends and investing blindly.

6.4 Model Comparison

To benchmark the efficiency of the TFT model, its performance is compared with a baseline LSTM model. Key metrics for Tesla and Apple stock predictions are presented below:

Table 3: Model Metrics Comparison

	TSLA		AAPL	
Metrics	**TFT Model**	**LSTM Model**	**TFT Model**	**LSTM Model**
MAE	8.04	12.56	2.07	4.99
RMSE	10.47	18.14	2.71	5.50
R^2 Score	0.95	0.85	0.63	0.54

6.5 Discussion

The results highlight the advantages of using the TFT model for financial forecasting:

- **Accuracy:** The TFT model outperformed baseline approaches in predictive accuracy, as evidenced by lower error metrics and higher R^2 scores.
- **Feature Insights:** Sentiment data, especially from Reddit and financial news, proved critical for enhancing prediction quality. The importance of structured financial data like stock volume was also evident.
- **Practical Application:** The trading simulations demonstrated that the framework could offer actionable insights for investment strategies.

Overall, the evaluation confirms that the framework achieves its objectives by integrating advanced transformer-based architectures with sentiment data, contributing to financial decision-making tools for academic and real-world applications.

7 Conclusion and Future Work

The successful construction and validation of a temporal fusion transformer framework for financial forecasting, in combination with sentiment analysis, was achieved using alternative data sources including Reddit and financial news, thus showing the potential impact of such alternative methods. The submitted model integrates historical stock prices, sentiment scores, trading volume, and technical indicators to generate precise and interpretable forecasting results. These results demonstrate, particularly with TFT's superior performance over traditional models such as LSTM, making it better suited for capturing complex temporal dependencies and integrating heterogeneous data sources.

The key findings show that the top-performing TFT consistently outperform the LSTM baseline across all datasets, as demonstrated by the reported MAE, RMSE, and R^2 values. Model performance was significantly enhanced by sentiment data for Reddit discussions and financial news, illustrating how alternative data teaches us something about market behaviour we might not have captured otherwise. The divide and conquer simulations of various trading strategies also confirmed the practical utility of the framework and its valuable insights into investors and other financial stakeholders. This empowers financial institutions with cutting-edge predictive capabilities while ensuring that the underlying models remain transparent and interpretable, fostering trust in decision-making processes.

Future work could extend the scope of this framework in several ways. Sentiment analysis can be applied to many of these analyses, and expanding these into areas such as Twitter, Facebook and LinkedIn can give a broader view of public sentiment and investor behaviour, which may improve the accuracy of the model. The use of various financial news sources, integrated with more regional or global ones as well as more specialized channels, could allow for a larger sentiment dataset that would enhance prediction accuracy even more. Such capabilities for real-time data ingestion and processing usage will further foster ongoing updates to the model for adaptation to financial markets with drastic pace. Typically real-time data is second to second and tick to price tick. In addition, using alternate data like institutional reports, macroeconomic indicators, or high-frequency trading data may provide insights into market fundamentals. Further improving the architecture is another direction, advanced transformer-based architectures or hybrid architectures or architectures combined with reinforcement learning can be good approaches.

This research highlights the power of transformer-based architectures in financial forecasting. The framework promotes better predictive modelling and decision-making by combining structured financial metrics with unstructured sentiment data effectively, addressing critical challenges in the relatively new field. Technological advancements in this area can potentially lead to adaptive, interpretable, and scalable tools that bridges the gap between academic research and industry utility, setting the stage for novel approaches to financial analytics.

References

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1): 1-8.

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint, arXiv:1810.04805.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint, arXiv:1702.08608.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2): 654-669.

Huynh, T. Q., & Nakamori, Y. (2019). A two-stage feature selection method for enhancing the interpretability of financial models. Expert Systems with Applications, 39(11): 9687-9697.

Lim, B., Arik, S. O., Loeff, N., & Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4): 1748-1758.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is All You Need. Advances in Neural Information Processing Systems, 30: 5998-6008.

Yang, S., & Kostadinov, K. (2020). Improving financial forecasts through sentiment analysis of financial news using deep learning. Finance Research Letters, 35: 101310.

Zhang, Y., & Zheng, Y. (2021). Exploring financial markets through deep learning: A review of applications and advancements. Computational Economics, 57(1): 99-122.

Chae, J., Lee, J., & Kim, S. (2023). Enhancing stock market predictions with domain-specific sentiment analysis and transformers. Journal of Financial Data Science, 5(2): 45-61.

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Lin, W. (2021). Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. Proceedings of the AAAI Conference on Artificial Intelligence, 35(12): 11106-11115.

Reddit Data (2018-2022): Kaggle Dataset - [Reddit Stock-Related Posts Dataset] (https://www.kaggle.com/datasets/injek0626/reddit-stock-related-posts?select=posts.csv).

Reddit Data (2022-2024): Kaggle Notebook - [WallStreetBets Data Collection] (https://www.kaggle.com/code/gpreda/collect-wallstreetbets-data/notebook).

Financial News Data (2018-2024): GDELT Dataset - [GDELT Database] (https://www.gdeltproject.org/).

News API. (n.d.). Retrieve Live and Historical News Data. Retrieved from [https://newsapi.org] (https://newsapi.org)

Stock Data (2018-2024): Yahoo Finance API - Stock data for Tesla (TSLA) and Apple (AAPL) accessed via the 'yfinance' Python library.

FinBERT Model: Pre-trained Transformer model for financial sentiment analysis, available on [Hugging Face] (https://huggingface.co/).

Optuna: Framework for hyperparameter optimization, used for tuning Temporal Fusion Transformer models. Documentation available at [Optuna Official] (https://optuna.org/).

PyTorch Forecasting: Time-series forecasting library used for Temporal Fusion Transformer implementation. Documentation available at [PyTorch Forecasting] (https://pytorch-forecasting.readthedocs.io/).