IPO Forecasting with Temporal Fusion Transformers: A Comparative Study with Deep Learning Models

# Aditi Sinha Bharath J Ramalingam B Mahima

*2109508@reva.edu.in 2109742@reva.edu.in 2110583@reva.edu.in*

*REVA University, Bengaluru REVA University, Bengaluru REVA University, Bengaluru*

Rayan Rafi Sheikh Dr. Shantala Devi Patil

*2108750@reva.edu.in shantaladevi.patil@reva.edu.in*

*REVA University, Bengaluru REVA University, Bengaluru*

# *Abstract—* *Initial Public Offerings (IPOs) are essential for private companies seeking access to public markets and capital. Predicting IPO outcomes, such as pricing and performance, has become increasingly data-driven with advancements in machine learning (ML) and deep learning (DL). ML models like decision trees and random forests excel at analyzing structured financial data, while DL models, including recurrent neural networks (RNNs) and convolutional neural networks (CNNs), identify complex patterns and temporal dependencies. Recent developments in predictive modeling have introduced advanced techniques such as Lasso regression for feature selection, gated recurrent units (GRUs) for improved time-series analysis, and attention mechanisms that enhance model interpretability. These methodologies improve the accuracy of IPO predictions, enabling investors and stakeholders to make more informed decisions. One of the most promising approaches is the use of Temporal Fusion Transformers (TFTs), which integrate multiple data sources and model dependencies across time, significantly refining IPO analysis. By leveraging these powerful tools, financial experts can enhance forecasting accuracy and navigate market uncertainties more effectively. By reviewing methodologies and highlighting challenges, this study demonstrates how we can reshape IPO analysis with the use of TFTs, enabling more accurate data-driven decision-making in financial markets and address the issues that were observed in the previous models. Continuous refinement of ML and DL methodologies and adequate understanding of the financial market, will help in building a more suitable model.*

# *Keywords —IPO prediction, machine learning, deep learning, neural networks, RNN, CNN, financial analysis*

# I. Introduction

Initial Public Offerings (IPOs) serve as a milestone for private companies, and that gives them access to public markets, raise substantial capital, and enhance their market visibility. Accurate prediction of IPO outcomes, such as pricing, first-day returns, and long-term performance, is essential for investors, regulators, and company stakeholders to make informed decisions. Although traditional financial metrics and qualitative evaluations have long been used for IPO analysis, the increasing complexity of market dynamics requires the adoption of advanced data-driven approaches.

Machine learning (ML) and deep learning (DL) models have emerged as powerful tools in financial prediction tasks, leveraging their ability to analyze large datasets and extract meaningful insights. ML models such as decision trees, support vector machines, random forests, and ensemble techniques have been extensively employed for IPO prediction. They excel in handling structured data and identifying influential features such as earnings per share, market conditions, and industry-specific factors.

Deep learning models, including recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformer-based architectures, have pushed the boundaries of prediction accuracy. These models can capture non-linear relationships, temporal dependencies, and intricate patterns in financial data. For instance, RNNs are particularly useful in analyzing sequential data like market trends, while CNNs are adept at extracting features from graphical or tabular data representations. In addition, hybrid models that combine the ML and DL methods are gaining traction because of their ability to balance interpretability and predictive performance.

However, existing ML/DL models face limitations in IPO prediction due to their inability to capture long-term dependencies, sensitivity to market volatility, and lack of interpretability. RNNs and GRUs struggle with vanishing gradients, CNNs are not inherently suited for sequential data, and traditional transformers are computationally expensive. Addressing these gaps, this study introduces Temporal Fusion Transformers (TFTs), which combine multi- horizon forecasting capabilities, gated residual networks, and attention mechanisms to improve IPO performance prediction.

Temporal Fusion Transformers (TFTs) offer a powerful solution by incorporating multi-horizon forecasting, selective feature importance weighting, and self-attention mechanisms. Unlike traditional deep learning approaches, TFTs provide uncertainty quantification, making them well-suited for financial prediction tasks like IPO forecasting.

This study proposes a novel TFT-based framework for IPO performance prediction and evaluates its effectiveness against traditional ML/DL approaches. The contributions of this study are as follows:

1. We introduce the application of Temporal Fusion Trans- formers (TFTs) for IPO performance prediction, demonstrating their effectiveness in capturing multivariate time-series dependencies.
2. We compare TFTs against traditional ML/DL approaches, highlighting their advantages in predictive accuracy and interpretability.
3. We provide an in-depth analysis of feature importance in IPO forecasting, offering insights for financial decision- makers.

II**.** LiteratureSurvey

* 1. *Introduction*

Accurately predicting an Initial Public Offering (IPO) performance is a critical challenge in financial markets. There have been a lot of previously tried and tested methods, which include traditional models, such as GARCH and regression-based approaches, which rely on historical data but struggle with market volatility, investor sentiment, and long-term dependencies. Machine learning (ML) techniques, including Random Forests, Gradient Boosting, and Support Vector Machines (SVMs), have improved prediction accuracy by analyzing complex financial indicators. Deep learning models like RNNs, LSTMs, and CNNs further enhance forecasting by capturing sequential dependencies.

Recent advances integrate sentiment analysis and hybrid models, combining fundamental and market-based indicators. However, these methods still face key limitations, including:

* + 1. Short-term focus, lacking multi-horizon forecasting ca- pabilities.
    2. Limited interpretability, making it difficult to understand key financial drivers.
    3. Inability to dynamically adapt to changing market condi- tions.
  1. *Traditional Financial Models*

Guzman et al. (2010) [1], explored advancements in predicting IPO pricing and performance using machine learning techniques, focusing on SBIR awards to identify significant variation in startups’ exit orientation.

Meghna et al. (2019) [2], examined IPO Pricing using statistical techniques and a GARCH model to evaluate long term IPO performance in India. Their analysis found no significant impact of dividend policy on IPO prices, but IPO underpricing was evident in short term trends.

Rui Ge et al. (2023) [3] used a Lasso based regression model, demonstrating that Retained profits per Share (REPS) are the most important predictor of IPO pricing in the Chi- nese market, which achieved a 70% accuracy in IPO Pricing forecast.

These models analyzed structured data, such as financial ratios, company fundamentals, and market conditions in partic- ular. However, they failed to capture non-linear dependencies in IPO Performance and did not adapt to external market shifts such as investor sentiment and macroeconomic trends.

* 1. *Machine Learning Based Approaches*

Quintana et al. (2017) [4] demonstrated that Random Forest outperformed traditional regression models for IPO prediction, achieving a higher predictive accuracy ( 65.21%) due to its robust ability to handle high-dimensional financial data, and their resilience against outliers, as compared to IBK, multi- layer perceptrons, and radial basis neural networks. They were particularly effective in ranking IPOs based on their potential initial returns and reducing under-pricing errors compared to observed outcomes.

Pravinkumar et al. (2023) [5] compared various ML al- gorithms, such as AdaBoost, Random Forest, Logistic Re- gression, Artificial Neural Networks, and Support Vector Ma- chines, coming to the conclusion that Artificial Neural Net- works were the most effective, achieving a 68.11% accuracy and highlighting the importance of specific factors.

These Machine Learning Models improved IPO forecasting by handling structured financial data and ranking important features, which led to improved classification accuracy. This came with a lot more feature engineering required to set up, as well as struggles with accurately taking into account time- series dependencies.

* 1. *Deep Learning Approaches*

Murugan (2018) [6] explored Recurrent Neural Networks (RNNs) for financial prediction, highlighting their effective- ness in processing sequential data, such as time-series data, spoken language, and sound. Having an intrinsic feedback loop that allows retention of temporal context across sequences led them to be very effective at tasks like language modeling, stock market prediction, etc.

Calafiore et al. (2019) [3] developed a ”Neuro-Survival” model, integrating Neural Networks with a Survival Analysis framework to examine the timing of the IPO within a specified timeframe, to provide accurate estimations of the time-to-IPO probability and successfully forecast the likelihood of IPO occurrences.

Ni (2022) [8] applied CNNs and LSTMs to IPO prospectus analysis, achieving higher accuracy by combining text-based financial data with historical stock trends. LSTM have loops and a hidden state that enable the network to deal with sequences and store historical data in the hidden state. Two sets of weights are used in RNNs: one for the inputs and one for the hidden state vector. The network learns weights for the inputs and the hidden state during training. Both the hidden state, which is dependent on earlier inputs, and the current input are used to determine the output when it is implemented.   
  
 The ability of simple RNNs to learn longer-term dependencies is constrained in practice. RNNs are frequently trained via backpropagation, which can cause either an explosive or disappearing gradient issue. In applications where the network must learn long-term links, these issues limit effectiveness by making the network weights either extremely small or extremely high.

Deep Learning Models, particularly LSTMs and CNNs, are effective in capturing nonlinear relationships and sequential dependencies. RNNs have difficulties during training, espe- cially when it comes to controlling gradient flow during back- propagation, which can result in problems like vanishing or inflating gradients. They also lack multi-horizon forecasting, and require large datasets for training.

* 1. *Sentiment Analysis and NLP*

Ly & Nguyen (2020) [9] examined the predictive power of sentiment analysis applied to IPO prospectuses such as Form S-1, which offer valuable insights into the investment offerings and the company’s prospects. Evaluating their performance across various timeframes showed a 9.6% improvement in classification accuracy over random baseline models.

Emidi & Gala´n (2022) [10] used Topic Modeling and Logistic Regression to analyze IPO prospectus sentiment. Their analysis reveals that variables like IPO price, ”SPAC,” ”commerce - retail,” and ”tech - entertainment” are crucial, while ”insurance - healthcare” and ”digital solutions” reduce accuracy of the model. This led to the conclusion that market sentiment influences IPO pricing and performance more than fundamental metrics alone.

Natural Language Processing models leverage investor sen- timent, textual data, and financial reports to a great extent, improving the IPO predictions to reflext in-market trends more accurately. These NLP models do require extensive pre processing and often lack contextual accuracy without additional financial indicators.

* 1. *Hybrid Approaches*

Baba & Sevil (2020) [11] found that ensemble models, such as Gradient Boosting and Stacked Neural Networks, improved IPO classification accuracy.

Neghab & Cevik (2023) [12] introduced Explainable AI (XAI) techniques like SHAP values to enhance interpretability in IPO price forecasting. They found that the offer price is the most significant predictor of IPO under-pricing, followed by equity retained and assets. It also reveals that under-pricing tends to be more pronounced in technology-based sectors, and during IPO surges, higher dispersion in firm quality leads to greater under-pricing. Tree-based models, particularly decision trees, outperformed other ML techniques in both regression and classification tasks. These findings underscore the importance of considering industry sector and market conditions when evaluating IPO pricing.

Bastı et al. (2015) [13] analyzed Turkish IPOs and found that market sentiment plays a crucial role in IPO performance, supporting the cyclical behaviour theory, where IPOs in bull markets tend to have higher initial returns. Their approach, combining decision trees (DT) with SVMs provide more re- liable insights by combining results from different algorithms through data fusion and sensitivity analysis.

Hybrid models improve feature selection, prediction accu- racy, and explainability in IPO forecasting, however these methods still rely on predefined feature sets, struggle with multi-horizon dependencies, and lack dynamic weighting mechanisms.

* 1. *Research Gap*

Significant advancements have been made in the field of predicting IPO outcomes, and there still exist gaps that mostly arise due to key limitations such as traditional models’ lack of ability to capture complex relationships, ML and DL models’ struggle with interoperability and long-term dependencies. When targeted with Sentiment Analysis methods, they require external financial context for a more accurate analysis, while Hybrid approaches lack dynamic feature weighting and they fail to provide uncertainty estimates in predictions..

This section compares the performance of different models for IPO prediction.

|  |  |
| --- | --- |
| **Model** | **MAE** |
| Gradient Boosting Classifier | 18.286 |
| Logistic Regression | 27.374 |
| Random Forest Classifier | 21.592 |
| Ada boost Classifier | 17.398 |
| SVM | 20.984 |
| ANN | 22.749 |

# III. Temporal Fusion Transformers

# 

Temporal Fusion Transformers (TFTs) offer a new approach that addresses these limitations by:

* + 1. Capturing long-term dependencies without vanishing gradients.
    2. Dynamically weigh feature importance using attention mechanisms.
    3. Provide uncertainty estimations, increasing reliability for financial forecasting.

*A. Methodology*

This study adopts a structured approach to developing an IPO forecasting model, proposing a Temporal Fusion Transformer (TFT) based approach. TFT’s can utilize specific features such as multi-horizon time-series forecasting, attention-based feature selection, and uncertainty estimation. All of these can be leveraged to improve the prediction accuracy.

The methodology follows a structured approach, covering data collection, feature engineering, model development, training, evaluation, and deployment.

* + - 1. *Data Collection and Preprocessing:* The dataset comprises historical IPO data, financial indicators, market conditions, and sentiment analysis, sourced from licensed financial data providers such as Bloomberg, Yahoo Finance, web-scraped reports, and API-accessible market indices. Sentiment data can be acquired by web scraping financial news forums and social media platforms. Macroeconomic Data Indicators, such as GDP growth, and interest rates, can be acquired from providers such as the Federal Reserve and IMF.

Data preprocessing involves cleaning raw inputs by handling missing values. Time-series gaps can be handled with forward and backward imputation. Numerical attributes can be normalized with min-max scaling. Categorical variables such as sector, region and IPO type can be encoded with one-hot encoding to ensure consistency and usability. A structured data pipeline, implemented with Apache Airflow, automates these preprocessing steps for efficiency and reproducibility.

* + - 1. *Feature Engineering:* Feature Engineering is critical for ensuring the model is able to predict with high accuracy. The dataset integrates both static and dynamic features.

Static features include company specific fundamentals, like revenue growth, earnings per share (EPS), price-to-earnings (P/E) ratio, leverage ratios, and Industry and sector classifications (e.g. technology, healthcare, energy).

Dynamic features include Market volatility indicators, such as the VIX index, and the previous day’s stock movements. Macroeconomic indicators, such as interest rates, GDP growth and inflation rate are also taken into

# Table 1

# Model Comparison

consideration. The combination of these features allows the model to capture complex temporal patterns in IPO performance.

* + - 1. *Model Architecture:* Traditional deep learning models such as long short-term memory (LSTM) networks and gated recurrent units (GRUs) struggle with long-term dependencies and feature importance weighting. The TFT model overcomes these limitations through multi horizon forecasting, which predicts IPO performance across multiple timeframes and interpretability compared to conventional deep learning architectures.

Unlike Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), the TFT model leverages attention mechanisms to dynamically assign importance to different features at each time step. Additionaly, overfitting risks are mitigated with the help of Gated Residual Networks, which refine feature selection and make our model more stable. Variable Selection Networks automatically select the most relevant features. Tensorflow was used to implement the model.

Unlike recurrent models, TFT employs temporal self- attention and gating mechanisms to filter only the most relevant past IPO data for forecasting.

* + - 1. *Training and Hyperparameter Optimization:* The dataset is divided into training (80%), validation (10%), and test (10%) subsets. Training minimizes the Mean Squared Error (MSE) loss function using the Adam optimizer with adaptive learning rate scheduling. Hyperparameter tuning is conducted via grid search and random search to optimize parameters such as learning rate, batch size, and attention dropout rates. Regularization techniques, including dropout layers and batch normalization, further enhance generalization and prevent overfitting.

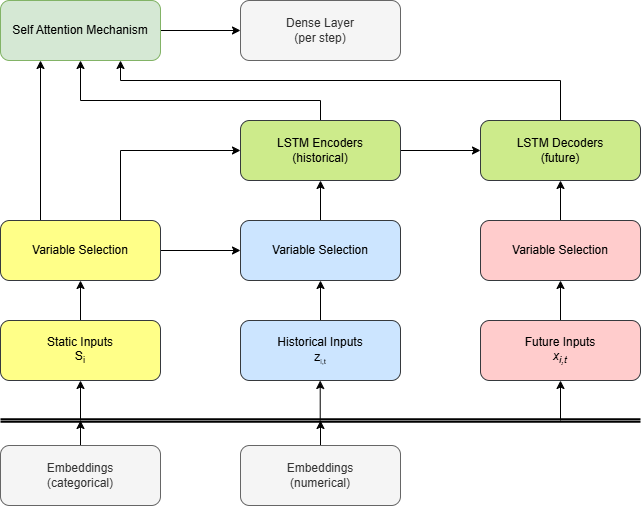


Fig. 1. Model Architecture

* + - 1. *Evaluation Metrics and Benchmarking:* Model performance is assessed using Mean Absolute Error (MAE), which quantifies average deviation. Mean absolute error

(MAE) calculates the average size of mistakes or inaccuracies between a dataset’s actual values and its anticipated values. The calculation involves calculating the mean of the absolute deviations between each prediction and the associated actual result. Benchmark comparisons against traditional models, including Random Forest regressors and GRU-based deep learning models, provide insights into the TFT model’s effectiveness.

IV. Conclusion

This study introduces a Temporal Fusion Transformer (TFT)-based approach for IPO forecasting, addressing key limitations of existing machine learning (ML) and deep learning (DL) models. The results of this study revealed that the Temporal Fusion Transformer (TFT) model achieved a Mean Absolute Error (MAE) value of 13.420. The TFT model’s better accuracy and effectiveness in forecasting tasks are demonstrated by its comparatively low error, which highlights its superior performance over other predictive models.

Traditional ML approaches, such as Random Forest and Gradient Boosting, have demonstrated moderate success in IPO prediction but struggle with feature selection, dynamic dependencies, and multi-horizon forecasting. Similarly, LSTMs and GRUs, while effective for time-series modeling, suffer from vanishing gradients, lack of interpretability, and poor performance in long-term trend forecasting.

The TFT model offers improvements by leveraging:

(1) Multi-horizon forecasting, capturing dependencies over extended time periods. (2) Attention-based feature selection, dynamically weighting relevant financial indicators. (3) Gated residual networks (GRN), improving feature stability and reducing overfitting risks.

V. FUTURE WORK

Several areas for future research can further improve TFT-based IPO prediction models:

Incorporating market sentiment analysis from financial news, analyst reports, and social media can enhances pre- diction reliability, and this can be achieved by using NLP techniques such as BERT-based embeddings and attention based transformers. These may improve model performance by capturing subjective qualitative factors that may influence IPO success.

Hybrid approaches, that combine TFT with graph neural networks to capture relationships may give greater insights and higher accuracy.

Future work can also focus on improving model interpretability using SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) to provide clearer insights for financial analysts.

By addressing these areas, future research can further refine TFT-based financial forecasting, ensuring higher accuracy, adaptability, and practical usability in financial markets.

REFERENCES

1. Jorge Guzman: Strategic Orientation in the Market for Startup Exits (2010)
2. S. Meghna, N. Suresh and J.C. Usha: A Study on Impact of Dividend Policy on Initial Public Offering Price Performance (2019)
3. Rui Gea, Lu Jiangb: IPO Pricing Research Based on Lasso-Fully Con- nected Neural Networks (2023)
4. David Quintana , Yago Sa´ez and Pedro Isasi: Random Forest Prediction of IPO Under-pricing (2017)
5. Pravinkumar M Sonsare1, Ashtavinayak Pande, Sudhanshu Kumar : A Comparative Analysis of Machine Learning Algorithms for IPO Under- performance Prediction (2023)
6. Pushparaja Murugan: Learning The Sequential Temporal Information with Recurrent Neural Networks (2018)
7. Giuseppe C. Calafiore, Marisa H. Morales, Vittorio Tiozzo : Survival and Neural Models for Private Equity Exit Prediction (2019)
8. Senhao Ni: Predicting IPO Performance from Prospectus Sentiment (2022)
9. Howard Ly, Khanh Nguyen: Do Words Matter: Predicting IPO Performance from Prospectus Sentiment (2020)
10. Christian Emidi, Sebastian Gala´n: Prospectus Content as Predictor of IPO Outcome: A topic model approach (2022)
11. Boubekeur Baba, Guven Sevil: Predicting IPO initial returns using random forest (2020)
12. Davood Pirayesh Neghab, Mucahit Cevik: Identifying the Factors Influ- encing IPO Underpricing using Explainable Machine Learning Techniques (2023)
13. Eyup Bastı, Cemil Kuzey, Dursun Delen: Analysing initial public offerings’ short-term performance using decision trees and SVMs (2015)