

# Financial Time Series Prediction with Multiple Variables using Transformers Neural Networks and Comparison with LSTM

Tarun Vemasani

Department of Computer Science and Information System  
Texas A&M University - Commerce  
Commerce, Texas, USA  
tvemasani@leomail.tamuc.edu

Latha Sree Kelari

Department of Computer Science and Information System  
Texas A&M University - Commerce  
Commerce, Texas, USA  
lkelari@leomail.tamuc.edu

## ABSTRACT

The Yahoo-finance dataset indicates the different company's aspect of the stocks, news, quotes etc. for every possible change with respect to time. We implement a solution analysis with statistical approach indicating the different tickers of the companies for every instant of time. We have considered 3 cases of time as (1 day, 1hr, 5 mins) for dataset to be compared with monthly and yearly wise based on the requirement for training aspect.

We indicate the maximum ticker counts as M, for which M-1 tickers are trained to predict the last ticker depending upon the type of Algorithms chosen. Our proposed work Transformer Neural Net indicating the DL model for the prediction of Stocks or Prices for the last tickers. We compare the same with the LSTM and other machine learning algorithms also.

## INTRODUCTION

A financial exchange serves two fundamental purposes. The first is to promote the process for the organizations' means of trade. The second is to identify and address the environment in which exchange can take place in an appropriate manner. Because of the market's obvious weakness and extremely unpredictable nature—in which shares, and values may rise or fall sharply—contributing to and profiting from it has never been simple. The factual percentage of profit dispersion for a particular securities or market file is instability. Typically, the riskier the security, the higher the degree of unpredictability. The volatility of the real costs of fundamental stocks is referred to as recorded instability or "known unpredictability." They have demonstrated to be most testing yet fulfilling and advantageous for venture. There are different examinations on the conduct of the market.

Recently, several considerations have been given to certain subordinates, like destiny and choices. For the purposes of the board, expecting these subordinates is of little consequence. In addition, accurate forecasting of the market's direction can help

financial professionals make significant gains with a small amount of capital. It is possible to think of a financial exchange forecast as a challenging time-dependent expectation. There are numerous variables that are persuasive on the monetary markets, including political occasions, cataclysmic events, financial conditions, etc. Despite the multifaceted nature of the developments in market costs, market conduct isn't totally arbitrary. Rather, it is represented by a very nonlinear dynamical framework. In the context of deep learning, a transformer is a model that adopts the process of self-attention and differentially weights the relevance of each component of the incoming data. Natural language processing (NLP) and computer vision are the two primary areas of application for this technology (CV).

Similar to recurrent neural networks (RNNs), transformers are designed to analyse sequential input data, such as natural language. Tasks like text summarizing and translation may be handled by transformers. In contrast to RNNs, however, transformers process the entire input at once. Thanks to the attention mechanism, context information may be obtained anywhere in the input sequence. The transformer need not go over each word in turn if the input data is a sentence written in natural language, for example. Higher parallelization is possible than with RNNs thanks to this, which reduces the length of time required for training.

Transformers were initially introduced by a team of researchers at Google Brain in 2017, and they are now swiftly replacing RNN models like extended short-term memory as the model of choice for difficulties with natural language processing (LSTM). The improved training parallelization now allows for training on larger datasets. This led to the creation of pretrained systems that can be customized for specific tasks. Examples include BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre-trained Transformer), which were trained using sizable language datasets like the Wikipedia Corpus and Common Crawl.

We indicate prediction of the single ticker chosen from the Y-finance dataset, realizing the different aspects of the open, close, volume, High. We realize the prediction features with 9 tickers

indicating the tenth ticker to be predicted based on the all values from 9 tickers. We estimate a solution to realize such effective changes and prediction weights on the model create the specific change in prediction and Testing values. Finally, an overall comparative study with different time instances is verified based on the type of the algorithm chosen and designed.