Financial Time-series Analysis for High-Frequency Trading

# Kumsetty Nikhil Venkat - 181IT224

*Dept. of Information Technology NITK Surathkal* Mangaluru, India

[nikhilv](mailto:nikhilvenkat26@gmail.com)[enkat26@gmail.com](mailto:enkat26@gmail.com)

# Amith Bhat - 181IT105

*Dept. of Information Technology NITK Surathkal* Mangaluru, India

[amithbhat01@gmail.com](mailto:amithbhat01@gmail.com)

Arya Sharma - 181IT207 *Dept. of Information Technology NITK Surathkal*

Mangaluru, India [arya.181it207@nitk.edu.in](mailto:arya.181it207@nitk.edu.in)

***Abstract*—The noisy and volatile nature of the stock market makes financial time-series analysis a highly challenging topic in the field of deep learning. This has often been attributed to the noisy and volatile nature of the stock market. Especially, in the field of High-Frequency Trading (HFT), prediction is a very challenging task because the automated inference system requires both precision and speed. In this project, we have used an unconventional construction method for predicting time series data, with positive results. These technologies are trained and tested with the benchmark LOB FI-2010 dataset, and the corresponding results are compared and analyzed using a variety of methods.**

***Index Terms*—Temporal attention , shallow neural networks, financial time-series, prediction horizon, attention mask, limit order book.**

1. INTRODUCTION

Time-series analysis and prediction has been a widely researched problem in the preceding decades. Time-series analysis has been applied to problems in fields as varied as natural language processing, finance and economics, meteorology, human behaviour analysis and a range of of other fields. Moreover,the complex dynamics of financial markets results in highly non-stationary and noisy observed data. Hence, this represents a very limited perspective of the actual price generating process.

Over the decades, several types of mathematical features have been proposed to extract meaningful and useful data from the stock market time-series data and systems. some of the features include the Auto-Regressive Integrated Moving Average (ARIMA) [4], [3]. However, these type of features are often made with several base assumptions, leading to misalignment in future observations. Therefore, machine learning models such as Logistic Regression and Support Vector Machines were used, which give better results than ARIMA in various situations.

Although the above-mentioned machine learning models perform well, they are not specifically designed to integrate

temporal data such as financial data into time series data. However, Recurrent Neural Networks (RNN) is specifically designed to extract temporal data from the sequential data. RNNs began to gain popularity in many different application areas [1], [5], [6] recently due to improved computer and computation hardware. Deep neural networks work directly on the introduction of raw data instead of hand-made objects. As a result, irrelevant data is automatically removed, improving the overall system performance.

While RNNs in general are quite efficient at work, trained structures are often difficult to interpret. Also, while not focusing on architecture that often improves performance and comprehension, it also includes higher computer costs across the model. This precludes the implementation of the model in many financial forecasting scenarios where the ability to quickly train the system and make predictions with large degrees of continuous input data plays an important role. Therefore, RNNs are not usually used for financial forecasting.

Therefore, in this project, we have implemented a new type of multivariate time-series data layer construction. The proposed structure is designed to use the concept of bilinear layers by introducing attention mechanism to the temporal mode. The LOB benchmark dataset, the FI-2010 database, was used in this project.

1. LITERATURE SURVEY
2. *Related Work*

In this project, we have taken [20] as a basic paper - using and developing the model proposed by its authors.

In financial time-series analysis, portfolio trading models were derived using Deep Belief Networks and Auto Encoders in [7], [8]. A 3 hidden layer MLP (Multi-Layer Perceptron) modelling the joint distribution of bid and ask prices was also used to study the spatial relations between LOB levels in [9].

Many deep neural networks for financial time-series analysis were proposed within a complex forecasting pipeline to account for the noisy and volatile nature of the market. In context of high-frequency LOB data (which is what we use as a dataset), the authors proposed to normalize the LOB states by the preceding days’ statistics. These normalized LOB states are then fed into a CNN [10] or an LSTM network [11]. Finally, the authors in [21] proposed a DeepLOB model which made heavy use of multiple CNN layers, before wrapping them in LSTM layers for reinforced learning.

1. *Motivation*

Most of the deep learning models in the market for High Frequency Trading(HFT) at the moment consist of non- specific classical models such as SVM, CNNs or recurrent structures such as LSTMs. Other custom-made DL models are very deep in nature or are not efficient, both of which are huge disadvantages in the highly competitive field of HFT.

1. *Problem Statement*

To create a custom-built deep learning model with two hidden layers to classify whether the stock price will increase, decrease or remain stationary over some prediction horizon. This is done by leveraging the idea of bilinear projection and incorporating an attention mechanism in the temporal mode, for maximum efficiency.

1. *Objectives*
   * To create a model with an extremely shallow layered architecture so as to maximise the efficiency of the model, with respect to time, hence building practicable HFT models which can be used in the real world.
   * The model must be very accurate, giving results comparable to other state-of-the art models.
2. METHODOLOGY

In this section, we examine our proposed structure in the problem of predicting medium price movements based on large LOB high-level databases. Before specifying in the test settings and numerical results, we first define the data and predictive function.

*A. Temporal Attention Augmented Bilinear Layer*

The studied model used only for a certain period of time in the past to predict the future value in a given horizontal study sequence. To learn the value of each time in the proposed BL, we suggest that the overridden Bilinear Layer (TABL) map input X ∈R*D×T* to the output Y ∈R*Dt×T t* as follows:

where *αij* and *eij* mean something in it (i, j) of A and E, respectively , Ⓢ it means duplication of wisdom operator, and Φ is a non-linear map defined as Eq. 2 *W*1 ∈ *RD×T* , W



Fig. 1. - (Eq. 5.)

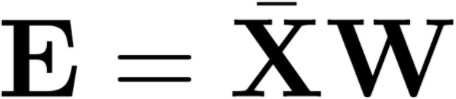


Fig. 2. - (Eq. 6.)

R*T ×T t* , *W*2 *RD×T* , B R*Dt×T t* and *λ* are a proposed bilinear layer of Temporary Extension. An additional temporary Bilinear Layer models depend on different W1 and W2 variants for the inclusion of intermediate attention step W and *λ*.

∈ ∈ ∈

To proceed with the Temporal Augmentation bilinear layer we have 5 steps, which are described in detail as follows:

* In Eq. 5, *w*1 is used to change the representation each time *Xct*, where t = 1,. . . , T of X in the new feature space *RDt* . This model relies on the X mode while keeping the temporary setting inactive.
* The second step aims to learn how important temporary conditions are to each other. This is achieved by reading a structured matrix W with dense elements centered on 1 / T.

Let’s explain *X*¯*t* R*Dt* and *et* R*Dt* the t column X and

∈ ∈

E respectively. From Eq. 6, we see that *et* is a weighted combination of T-temporal positions in the space of *RDt* , i.e., T-columns of X, which have periods of time always equal to

1 / T since the diagonals of W are set to 1 / T. Therefore, the element *eij* in E adds the corresponding value of the *x*¯*ij* item to the other - *x*¯*ik*, where k ƒ= *j.*

* Normalize the values of E using the softmax function in Eq. 7, the proposed layer pushes many objects to close to zero while keeping prices some of them positive. This process produces the attention mask A.
* Attention mask A found in step three is used to eliminate the effect of non-essentials on *RDt* . Instead of using the hard-earned approach, the readable scale *λ* in Eq. 8 allows the model to learn a soft attention span. In the first stage of the learning process, the learning features extracted from the previous layer can be noisy and non-discriminatory, so hard attention can mislead the model to insignificant information. In contrast, soft attention may allow the model to learn discriminatory features at the beginning of the phase. Here we must know that it is compulsory to sleep in it width [0, 1], i.e. 0 *λ* 1.
* Similar to Bilinear Layer, the final step of the proposed layout estimates the *w*2 interim map, excluding high-level representation after a change of bias and linearity.

In general, the introduction of a focused approach in the

Thank you for using [www.freepdfconvert.com](http://www.freepdfconvert.com/) service!

Only two pages are converted. Please Sign Up to convert all pages. <https://www.freepdfconvert.com/membership>