

**Name:** Rayavarapu Navanith

**Email address:** navaneethrayavarapu@gmail.com

**Contact number:** +91-9160134919, +91-7893830313

**Anydesk address:** 900 921 949

**Date:** 19<sup>th</sup> Jun 2020

**Self Case Study -1:** \*\*\*Implementation of An-Attention based LSTM Model for Financial Time Series Prediction\*\*\*

---

## Overview

\*\*\* Write an overview of the case study that you are working on. **(MINIMUM 200 words)** \*\*\*

### 1. Problem Description :

Forecasting financial time-series data is always a challenging task because of its dynamics. The key characteristics present in financial time-series data are a) volatility, b) non-stationarity, c) periodicity, d) non-linearity and e) long-term dependencies. By using traditional statistical models like ARIMA and GARCH, can capture volatility and periodicity, and also can interpret the model but it fails to capture non-stationarity and non-linearity in the series. By using neural networks like RNN, can capture non-stationarity and non-linearity series but fail to capture long-term dependencies in the series. And also the problem with neural networks is its interpretability, they fail to describe the role of each input feature. In this project, I'm trying to implement the research paper titled "[AT-LSTM: An Attention Based LSTM Model for Financial Time Series Prediction](#)". This model takes care of all the characteristics in financial time series.

## 2. Dataset Description :

I'm trying to use the Nasdaq composite historical data from January 2, 1991 to June 20, 2020. Data can be downloaded from [yahoo.finance](https://finance.yahoo.com/) website. This data contains daily stock prices of Open, Close, High and Low.

---

## Research-Papers/Solutions/Architectures/Kernels

\*\*\* Mention the urls of existing research-papers/solutions/kernels on your problem statement and in your own words write a detailed summary for each one of them. If needed you can include images or explain with your own diagrams. **it is mandatory to write a brief description about that paper. Without understanding of the resource please don't mention it\*\*\***

1. <https://iopscience.iop.org/article/10.1088/1757-899X/569/5/052037>

I'm implementing the above research paper. This research paper's main objective is to use Attention based sequences instead of raw sequences and then feed the AT based sequences to the LSTM.

The following diagram explains the objective of above problem :

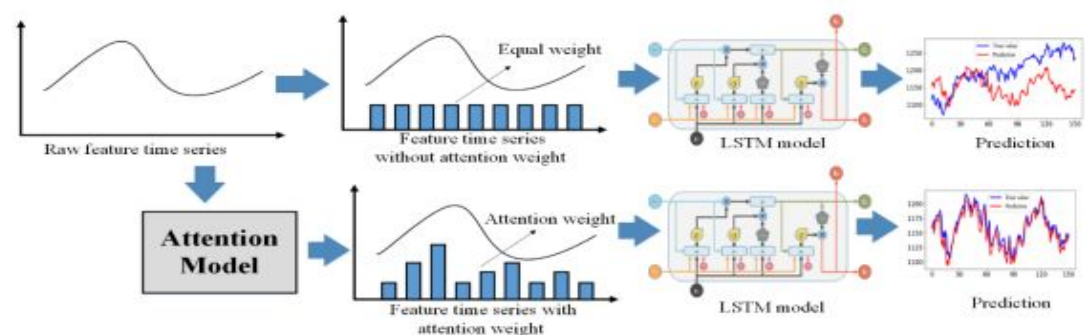


Figure 1. Comparison of pipelines in traditional LSTM-based model and AT-LSTM. The traditional LSTM-based financial time series prediction framework directly takes the raw feature sequences as input, with all input feature treated equally. AT-LSTM gives important features a higher weight and reduce the weight of redundant features to guarantee good prediction accuracy.

2. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227222>

Above research paper uses the same attention mechanism but instead of converting raw sequences to AT based sequences it uses the context vector(encoder) to train the model And also this model uses wavelet transform denoising in the initial layers.

3. <https://www.analyticsvidhya.com/blog/2019/12/6-powerful-feature-engineering-techniques-time-series/>

The above blog suggests how to generate features from time-series data. There are five different features we can generate using time-components. They are

- a. Date related features
- b. Time-Based features
- c. Lag features
- d. Rolling window feature
- e. Expanding window feature

---

## First Cut Approach

\*\*\* Explain in steps about how you want to approach this problem and the initial experiments that you want to do. (**MINIMUM 200 words**) \*\*\*

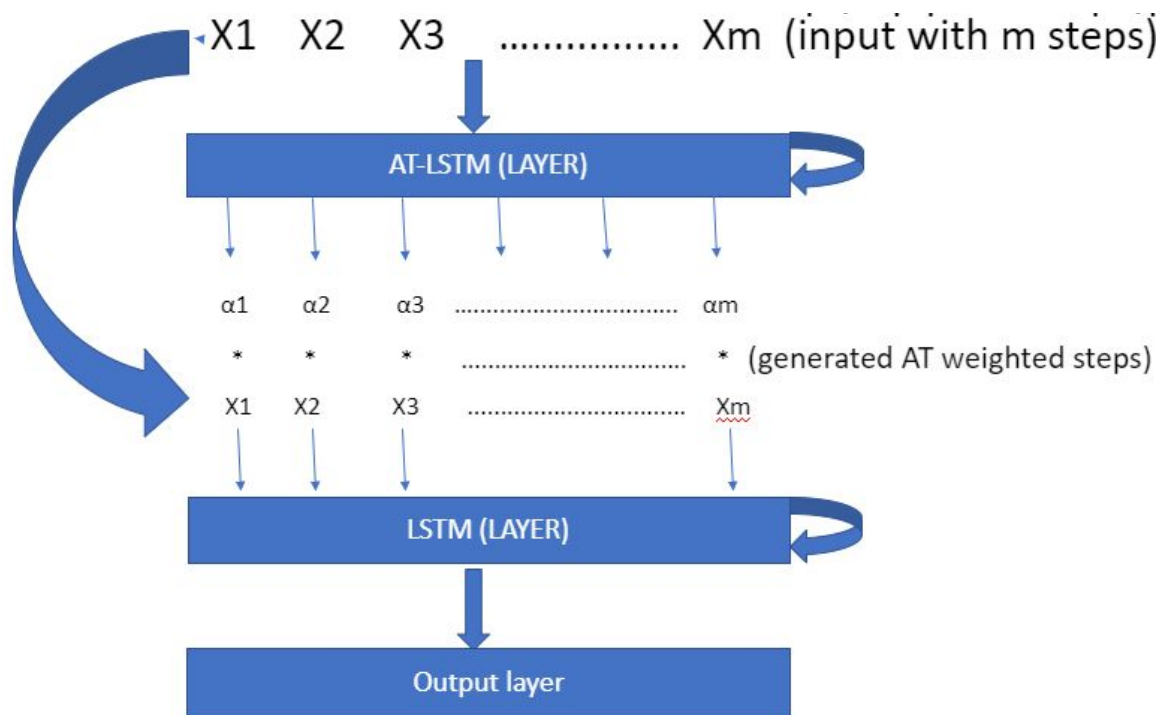
### Target variable Description :

1. There are three different types of outputs we can generate from the given dataset.
  - a) Next Period closing price Direction (Classification problem)
  - b) Net change in the closing price (regression problem)
  - c) predicting closing price (regression problem)

### Input feature Description :

2. There are two types of features that can generate from this time-series.
  - a) One is univariate feature (closing price of previous time steps)
  - b) Another one is multivariate features (using remaining features contained in the dataset ex: open price, closing price, volume, high, low)

My approach for this problem is to implement the research paper. According to the research paper, the target variable is closing price and it's multivariate features are open price, closing price, volume, high, low from previous time steps. Following diagram is the basic model architecture.



Where,

->  $[X_1, X_2, \dots, X_m]$  are input sequences where  $m$  indicates the time step

-> Each  $X_m$  contains  $n$  features like closing price, opening price, High, low, volume.

-> Using the AT-LSTM layer, generate attention weight for each sequence input. Such that  $[a_1 * X_1, a_2 * X_2, \dots, a_3 * X_m]$  becomes a new attention generated sequence.

-> This sequence would be input to the LSTM layer.

-> The final unit ( $h_t$ ) from the previous LSTM layer would be fed to the output layer.

-> As it is a regression problem The evaluation metric is MAPE.

---