

# Predict Stock Price Change with Market News

**MOTIVATION** Stock price prediction is an important area in the finance industry. It can be used for managing investment or mitigating various types of financial risks and managing portfolios, valuing assets etc. Looking at various trends in stock prices over the years, we can forecast the future stock price. Such model can be used by investors/traders to predict how the market would react.

Along with that, one of the factors contributing to the stock price change is news. News such as announcements of acquisitions, earnings or major mergers may affect the stock prices and may determine company's earnings affecting people's decision to buy or sell the stock. And we would like to see the effect of market news on stock price change.

**PROBLEM STATEMENT** To predict stock price for Apple Inc. based on financial/business news using time series analysis. To analyze if the market news affects in stock price change.

## QUESTION TO BE ANSWERED

- 1) Can we accurately predict the increase or decrease in the stock price of Apple Inc.?
- 2) Does the market news affect the change in stock price?

❖ **FOR ACCURATELY PREDICTING STOCK PRICE:**

## **DATASET SOURCE Stock Prices**

**Data :** Source : [Yahoo Finance](#) Date

Range : 01/01/2000 - 12/31/2017

## **Features**

:

<b>Column Name</b>	<b>Description</b>	<b>Type</b>
--------------------	--------------------	-------------

Date	Date	DateTime
------	------	----------

Open	Opening price of Stock on the given Date	Float
------	--	-------

Close	Closing price of Stock on the given Date	Float
-------	--	-------

**APPROACH** Our stocks dataset includes stock values for Apple Inc. for a time span of 2000 - 2017. Therefore we will be using time series forecasting to predict the future stock values for Apple Inc. We will also analyze the dependence of future values on present and past values.

- We **plot time series** test and check for any **trend, seasonality or anomalies** in the time series data in order to process it further for applying time series models like ARIMA
- We **verify if the time series is stationary** by checking with Durbin Watson test to assess the presence of autocorrelation and by using the Augmented Dickey-Fuller Test. If it is not stationary, we can use detrending (removing trend) or differencing.
- With the help of **ACF and PACF plots**, we will get the parameters p (number of autoregressive terms), q (number of moving average terms) and d

(number of non-seasonal differences needed for stationarity). Using these parameters we can then build an **ARIMA** (AutoRegressive Integrated Moving Average) model.

- Now we can **predict the stock price** using these ARIMA model.

**EVALUATION** We can also show a plot of the actual price change and the predicted price change to visualise the performance of different models. We will evaluate the performance of ARIMA model using Mean Absolute Error and Mean Forecast Error. A value near 0 means it is a pretty good fit and is performing well.

## ❖ FINDING CORRELATION BETWEEN NEWS AND STOCK PRICE

**DATASET SOURCE Market News Data** : We will be extracting news data using [NYTimes API](#) to query for news articles on Apple Inc from date range 01/01/2000 - 12/31/2017.

### Features

:

Column Name	Description	Type
-------------	-------------	------

Date Publish	Date	DateTime
--------------	------	----------

News Snippet	Snippet of the news article	String
--------------	-----------------------------	--------

News Type	Types of News	String
-----------	---------------	--------

Although the API gives us a lot of attributes, we would be considering the date, snippet and news type in the article, since the snippet in the article would contain the major information required to analyse the sentiment/mood of the news article.

**APPROACH** Our market news dataset consists of articles pertaining to Apple Inc. from 2000 - 2017. ● For **cleaning data**, we will use the news type attribute to filter news in finance, business and technology ensuring we have gathered information pertaining to Apple Inc.

- For **missing values**, If we don't have news on a particular day about Apple, we will replace the null value with an empty string, so that it wouldn't affect sentiment analysis.
- Our first task would be **analysing the sentiments of the article**. We will be using the Natural Language Processing toolkit and sentiment intensity analyser from the Vader library within the toolkit to extract the polarity and intensity of the input. With the output compound providing an intensity from -1 (negative) to 1 (positive) and the negative and positive coefficient as well.
- On the basis of the results obtained, we **perform Pearson's correlation** coefficient, which would give us a value between +1(positive correlation) and -1(negative correlation).
- The result would tell us if there is actually a correlation between the two.

## **ASSUMPTIONS**

- We will also take into consideration that the effect of news is instant and valid on the same day, as many stocks news arrives fairly often and the market responds quickly. Therefore we will try to see what effect a news story will have on stock price on the same day.
- We will assume that only news stories mentioning 'Apple Inc' will affect its stock prices since it would be a very broad problem if we take into consideration competitors action that indirectly affects Apple's stock prices.