What goes up…

A deep learning approach for predicting compound returns of the S&P 500 index

Justin Hanson
University of Texas
Case Studies in Machine Learning
AI 395T, Fall 2024

Github repository
https://github.com/hansonjw/UT_AI395T_what_goes_up

Abstract

Financial analysis and market prediction are intense areas of analytical study. However, asset price prediction and modeling remains challenging, especially in the stock market. Many professional investors would say it is next to impossible to beat average returns over the long term for investments such as stocks and stock indexes. Many studies have been under taken to build Deep Learning models to predict stock market prices, but surprisingly, many of these studies are (1) make predictions over short time horizon's such as the next day, and (2) are focused on looking for patterns in the historical price movements movements (*technical analysis*). This paper takes a different approach by applying a Deep Learning to model longer term forecasting (5 years out) and incorporating macro economic data. In this approach a 'simple' LSTM model more accurately predicted the price of the S&P 500 prices 5 years out based on reading 3 years of feature data than a baseline exponential curve fit based projection.

Investing

Investing is typically described as the spending of money now, with the expectation of gaining more money at a future date. While this concept makes sense, it does open lots of interesting ideas and areas of study. When we assert an idea such as 'gaining future money' what kinds of assumptions are we making? Is this future money guaranteed (risk)? Is the timeline of when the future money assured and/or estimable? How much future money or value should one expect for committing a sum of money now? How would one assess and compare available investment options?

Irving Fisher, an economist of the later 19th and early 20th century was one of the first to begin to formulate methods such as *present value* and estimating future cashflow streams for asset valuation (Reamer & Downing, 2016, pg 231). He was later followed by people like Benjamin Graham who wrote *"Security Analysis"* which became a standard textbook for many very successful investors such as Warren Buffett. These people, while rigorous with the craft of investing, were very clear minded of the limitations of predicting future prices and valuation as a whole. As Benjamin Graham wrote:

> *"The soundness of a security purchase is determined by future developments and not by past history or statistics. But the future cannot be analyzed; we can seek only to anticipate it intelligently and prepare for it prudently. Here the past comes in - through the back door, as it were - because long experience tells us that investment anticipations, like other business anticipations, cannot be sound or dependable unless they are closely related to past performance."*
> (Graham & Dodd, 1951, pg 3)

Technical Analysis is generally a practice of investing which focuses on looking for "signals" in the historical price changes of a stock or index to anticipate future prices. While this practice is controversial, especially for individual stocks, perhaps there is more information for indexes, which are aggregate measure of broad stock market prices. In contrast, *Fundamental Analysis* is method of analysis in which other information such as financial data, etc. is incorporated into anticipation of future prices.

With this background in mind, this paper implements a simple Deep Learning model to make predictions on a broad based stock market index. The results are then compared to a "baseline" method which employs an exponential curve fit and extrapolates out over the prediction range. For a first cut, and for simplicity a prediction range of 5 years is chose for the analysis.

The Benchmark

The S&P 500 is a generally considered a major benchmark for measuring investment performance and is used widely as a broad spectrum economic indicator, especially in the United Stats of America (S&P Global). Many seasoned investors consider the S&P 500 performance difficult, if not impossible to beat through active investing over the long term.

Warren Buffett, perhaps the most well know investor, documented the results of a bet in which he wagered $100M (winnings going to charity) against a professional active investor - one who selects individual investments rather than depending on a passive investment strategy such as selecting an index. The premise of the bet was that the professional investor could not beat the S&P 500 returns over a 10 year period. Warren Buffett won the bet and beat a the returns of a professional active investor with a simple bet on the S&P 500 (Buffet 2018). This is an astounding result and illustrates the general acceptance of the S&P 500 as a broad stock market indicator.

It is important to mention John "Jack" C. Bogle who founded Vanguard funds and is generally considered one of the pioneers of passive index investing and helped establish indexes for investors to invest in. The S&P 500 is chosen for this study.

Compound Growth

Compound growth is typically described by text books for finance with a strictly mathematical formulation. The formulation, or model, is a hypothetical security (a bond) which has the following key general parameters:

- A fixed rate of income (annual, monthly, etc.) - also called a coupon
- A duration or term to maturity (30 years, 10 years, etc.)
- A principle which is the amount of money loaned out and returned at the end of the term

The term "hypothetical" is used as this strict formulation is not typically available as a realistic investment option in any pragmatic sense. Investors are constantly comparing relative values of assets, assessing market conditions, assessing interest rates - which change over time, etc.

Setting aside practical considerations for now, the hypothetical model we can help us develop an intuition of "compound growth", or "compound interest" which is a term many academic presentations of the topic use. My preference is to utilize the term "compound growth" as a point of distinction which will hopefully become clear.

Say for the sake of example, an investor loans out $1,000 (buys a bond) at an annual coupon rate of 5% and term of 10 years. At the end of the 10 years the investor will receive back the original

$1,000 as well as get to keep all of the coupon payments along the way which are $50 a year over the term of the bond. In total the investor made $500 over 10 years.

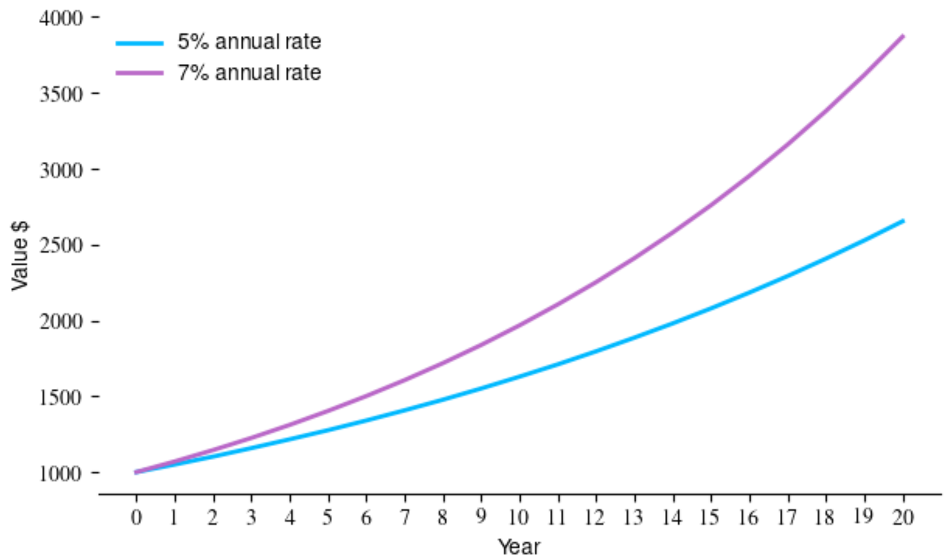

Figure 1- Hypothetical value of $1000 invested with compound returns

Now if we factor in another hypothetical assumption that each coupon payment can also earn coupon payments of 5% for the remainder of the original 10 year loan. Following the example, at the end of 2 years the investor would have received $50 + $50 + $2.50 in coupon payments. Following this pattern out for the full ten years, the investor will receive a total sum of $629, which is not insignificantly more than the original $500. In fact, if the annual rate of return is 7% the investor can expect to nearly double the original principle amount ($1000) over ten years. This quick example highlights the importance of compounding returns, or specifically in the context of loans and interest we can invoke the term "compound interest".

When this hypothetical example is plotted out you see the distinct and characteristic curve of compound growth. Compound growth is the signal investors are typically searching for. Figure 1 shows the returns of a 10 year $1000 investment at 5% and 7% interest rates.

The Signal In The Noise

How does this simple hypothetical model apply to stocks? Are there any intuitions we can derive from the basic model and how it might apply to an asset such as the S&P 500 index? Stocks typically don't have obligated, or contractual cash coupon payments, and there is no term, or even a guarantee that an investor will receive back their initial investment. The investment can go to zero which is why stock are generally considered "risky". There reason the compound growth model is import is because the S&P 500 index does exhibit the kind of compound growth illustrated with the basic model. Figure 2 plots the long term price of the S&P 500 going back to 1929 both on a linear scale (blue) and a logarithmic scale (yellow).

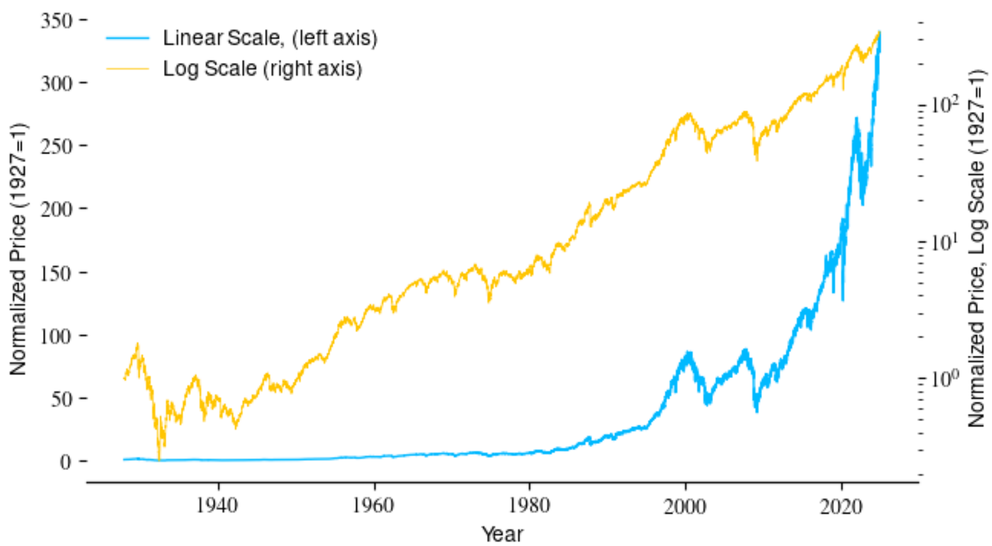

Figure 2 - S&P 500 long term growth, compound returns

If stocks don't have the same parameters/structure (term, coupon, principle), what is the mechanism that is driving the kind of compound growth typically associated with the basic hypothetical model?

Valuation, finance, security analysis

In a very basic sense businesses have expenses and revenues. At the end of a period of time, or say the life of the business, the revenues less any expenses incurred are the 'profit' of the business enterprise. This profit belongs to the shareholders. Now, shareholders can liquidate the business, or they can choose to direct the "profit" back into the enterprise for expansion, reinvention into new markets, etc.

Now this is overly simplistic, but we must start here in order to draw connections back to the basic hypothetical model. It worth thinking about how difficult this all is to predict, especially given long time horizons (years, decades) in which modern businesses operate. However, this idea is not a feature solely for modern businesses. It's an old idea as noted by Adam Smith:

> *"Profit is so very fluctuating, that the person who carries on a particular trade, cannot always tell you himself what is the average of his annual profit. It is affected, not only by every variation of price in the commodities which he deals in, but by the good or bad fortune both of his rivals and of*

his customers, and by a thousand other incidents, to which goods, when carried either by sea or by land, or even when stored in a warehouse, are liable. It varies, therefore, not only from year to year, but from day to day, and almost from hour to hour." - Adam Smith (Smith, 1776, pg 51)

It's very much worth contemplating what exactly Adam Smith is point out and noting how relevant it still is to day, especially with modern security analysis.

The endeavor

Up to this point we have laid a foundation of financial concepts that are core to contextualizing financial analysis, asset pricing, and prediction. Can we build machine learning model to help with asset price predictions?

The problem definition:

Will a "simple" deep learning model, which reads 3 years of past price and macro economic data demonstrate evidence of accurately predicting the S&P 500 price 5 years out into the future? Will this macro economic data and a machine learning model provide any better performance than traditional estimates of future returns based on historical price performance?

Technical Formulation

The model built for this exploration is a deep learning model which uses two LSTM layers and a dense layer to predict the price of the S&P 500 index. LSTM layers developed by Hochreiter, and Schmidhuber were selected due to their effectiveness and application for time series RNN models (Hochreiter, Schmidhuber (1997)).

This paper only serves as a starting point and is not intended to present a conclusive model for predicting the 5 year S&P 500 performance. Rather through this exploration the hoped for outcome is to discover evidence and motivation to justify future exploration.

A brief review of literature

A number of papers have been published employing various machine learning and deep learning techniques to model and predict stock market prices. These models have been build for both individual stocks, as well as for indices such as the Dow Jones Industrial Average and the S&P 500. Table 1 contains a summary of 5 selected papers in which Deep Learning models were applied to financial data in a similar approach for this paper.

It is somewhat surprising to note most of these studies focused on short term predictions - typically focused on the next day's price. Also noted in these studies is a lack of f*undamental analysis*, or analysis that include other data such as macroeconomic features, or features beyond historical pricing. One exception is Bao, et al. (2017) who employed a model that seemed to leverage vast amounts of data including news, financial reports, etc. Another item of note is that all the papers discussed the LSTM models and referenced the original work of Hochreiter & Schmidhuber (1997). The lone exception is Won, et al (2022) which focused their work on machine learning techniques including decision trees, and did not utilize deep learning.

Reference / Source	Individual Stock or Index	Fundamental or Technical	ML/DL Model Implemented	Prediction Timeframe
Faajijou, et al. (2021)	Idexes (S&P 500, Dow, Nikkei, etc.)	Technical - Only index price data	Hybrid CNN-LSTM	1 day
Won, et al. (2022)	Individual stocks (XOM)	Technical - Only index price data	Other ML (non DL)	1 day
Chang, et al. (2020)	Indexes	Technical, Sentiment, Fundamental (?)	LSTM	"Short Term" (??)
Bao, et al. (2017)	Indexes - CSI 300, Nifty 50, Hang Seng index, Nikkei 225, S&P500 and DJIA index	Mix, Largely Technical	LSTM, Others	Annual
Chandra 2024	8 individual stocks	Technical	LSTM	1 day
Yan & Yu 2018	Index (S&P 500)	Technical	LSTM	1 day

Table 1 - Selected studies involving Deep Learning Models and Financial Data

Furthermore, in a study of 122 research works on the prediction of stock markets, Adekoya, et al. (2020) noted that "…the predictive timeframe of most studies was a day ahead…". They further noted that 66% of these studies focused on *technical analysis* - finding patterns in the historical price data (Adekoya, et al., 2020). Again highlighting a lack of feature modeling with macroeconomic data or other more fundamental model inputs.

Features and inputs for a deep learning model

As a starting point for the analysis, the following features were selected based on contemplating a plausible impact on U.S. stock prices in aggregate, availability of data over the long term, as well as general scholarly acceptance as meaningful economic indicators. Exotic data sets, and complex modeling were generally avoided for the sake of establishing a baseline for future study. What follows is the list of features included in this deep learning model.

CPI, Inflation

A key challenge in estimating prices of assets over time is handling the concept of inflation. There is a general consensus that inflation is something like *"the rate at which the price of goods and services increase over time"*. However, economists aren't necessarily aligned on what really causes inflation or how it is measured. For example, noted economist Milton Friedman is famous for saying:

> *"Inflation is always and everywhere a monetary phenomenon. It is a result of a greater increase in the quantity of money than in the output of goods and services which is available for spending" - Milton Friedman (Freidman, 1974)*

Yet other economists take a more *"black box approach"* with inflation and assume there is more going on in an interaction of Monetary and Fiscal policy.

Perhaps The Federal Reserve Bank of Cleveland provides the most broadly contextual explanation of inflation which highlights some of the core difficulties is quantifying it:

> *"Prices can change for different reasons and in different ways. The prices of individual goods and services can change because the supply or demand for the items has changed…so how can we tell when inflation is happening and by how much? We do so by looking at the prices of many items over time. " (Federal Reserve Bank of Cleveland)*

Jiang analyzed the impacts to inflation by changes in Federal Funds Rate, Forward Guidance provided by the Federal Reserve, as well as Large Scale Asset Purchases — Federal Reserve purchases of mortgage backed securities and other assets in the open market. Jiang's work illustrates the "main stream economic" view of inflation and points to monetary policy having measurable and quantifiable impacts on inflation as measured by the Personal Consumption Index (Jiang, 2024).

As such it is worth thinking through the implications. Is this inflation phenomenon a natural order of things, as in do prices always increase? Do all prices change in the same direction at the same time for all goods and services? Can we really measure it? Can we parse out the aggregation of concepts at play such as the supply and demand of money, and the supply and demand of goods and services? How do inflation indexes, which are utilized over long ranges of time (decades, etc.), address and factor in the emergence of new technologies such as cell phones, computers, automobiles, etc.? What about changing consumer behavior? As always the devil is in the details.

The concept of inflation is crucial to security analysis as investors seek to quantify the increase in value over time of the asset or security, and thus are attempting to parse out the true "intrinsic value” of the asset versus “real price” and/or “nominal price”. A typical technique, as illustrated by Fisher is to adjust “nominal” prices for inflation and to arrive at “real” prices. Fisher does this for assets such as housing, stocks, and bonds - in particular his famous U.S. Home Prices long range chart (Shiller 2015, pg 20).

Ceballos provides analysis that indicates investment decisions and portfolio allocations are indeed predicated in part on inflation data and inflation expectations (Ceballos, 2024). This work provides motivation to include inflation as a feature in model, even though the topic is difficult.

Federal Government Agencies publish a variety of indexes which attempt to measure inflation such as Consumer Price Index (CPI), Producer Price Index (PPI), Personal Consumption Expenditures (PCE) deflator, among many others. These measures are often times adjusted for other factors such as seasonality and can be reported on a basis that excluded components such as housing and energy. Dietrich asserts that households (general population) disproportionately attend to “headline inflation”, or measures like CPI, etc. instead of “core inflation” which energy and food as these tend to be more volatile (Dietrich 2024). This further highlights the complexity and nuance of the the concept of inflation, and areas where consensus is lacking.

In future iterations of this study, a review of different “features” for inflation may prove fruitful. However, for the sake of simplicity, and because of a long range of available data, CPI is used as a feature in the model. (U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers)

Figure 3 shows a long range plot of the Consumer Price Index (CPI) dating back to 1913. It is interesting to note that this plot suggests a sort of “natural order” or strong signal in inflation that machine learning models may be able to pick up on.

One further note on inflation: the preceding comments are not intended as a critique of the techniques and measures of inflation. These are difficult endeavors, and much great work has gone into wrestling with difficult to measure, and perhaps somewhat abstract economic subjects. However, these subjects remain clouded and solid bedrock, and scholarly consensus remains elusive. Hopefully applied modern mathematics, abundance of data, and availability of computing power can begin to steer us in new scholarly directions on the subject.

For simplicity, CPI is used as a model input and was sourced from the U.S. Bureau of Labor Statistics through the St. Louis Federal Reserve (U.S. Bureau of Labor Statistics).

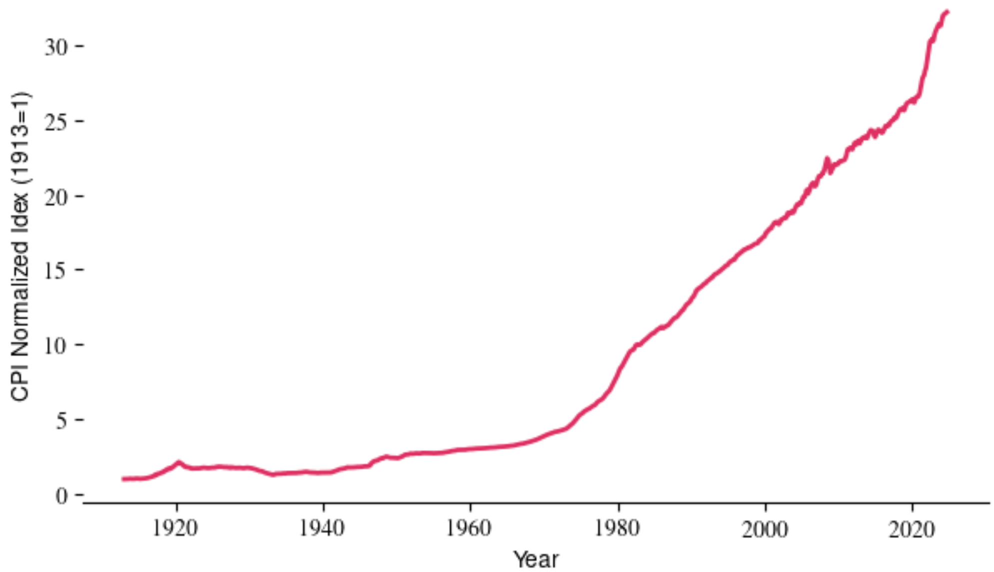

Figure 3 - Historic CPI dating back to 1913

Monetary Supply, M2

Given the ambiguity around inflation an additional macroeconomic variable is included. The monetary supply is a measure of how much currency is in circulation and can be measured many ways and is typically looked at for analysis on inflation. For the sake of simplicity and data availability M2 is used and the data was sourced from the St. Louis Federal Reserve (Board of Governors of the Federal Reserve System)

Federal Funds Rate

The Federal Funds Rate, generally speaking is an underlying interest rate that has a variety of impacts to all other interest rates in the U.S. such as car loans, bonds, mortgages, etc. As such it is a tool used by the Federal Reserve to impact inflation which garners widespread attention in the financial news and the investment community. Jiang quantified impacts to inflation from changes in the Federal Funds Rate, as well as other variables (Jiang 2024) which serves as evidence for the impact on economic activity and investing. Data for the Fed Funds Rate is included in this model and was accessed through the St. Louis Federal Reserve (Board of Governors of the Federal Reserve System (US)).

Triple AAA Corporate Bond Yields

All kinds of Interest rates tend to impact one another to some degree, similar to the way the Federal Funds Rates underpin most interest rates. Additionally Corporate bonds offer investors yet another investment alternative to stocks, municipal bonds, precious metals, real-estate, etc. Data for AAA rated corporate bonds from Moody's was sourced through the St. Louis Federal Reserve and are included as a part of the model (Moody's, Moody's Seasoned Aaa Corporate Bond Yield).

Housing Prices

For at least the last two generations (Baby Boomers, etc.) housing is generally considered the largest purchase for most families and individuals in the course of a lifetime. Housing typically involves financing, as well as personal tradeoffs in purchases of other financial assets including retirement saving, personal savings, etc. As such one can begin to conceptualize many ties between housing prices, interest rates, US stock market and many other asset prices.

Nobel Prize winning Economist Robert J. Shiller studied housing price data and has published this data in his book *Irrational Exuberance.* Housing data for the model was sourced from Shiller's website (Shiller, R. J. *US National Case Shiller Home Price Data)* which goes as far back the late 1800's.

Gold Prices and Gold to Silver Ratio

Financial analysis is generally the study and analysis of relative value. Gold and silver are assets that have long been used to store value and as a medium of exchange throughout human history. Also, many people consider "The Gold Standard" to a be true measure of value. Again, like inflation, these concepts are very much up for scholarly debate, however, investors do have a choice to purchase among a wide variety of assets including gold and silver. The price of gold and silver are included as features in the model. Officer and Williams (Officer, L. H., Williamson, S. H., *The Price of Gold, 1257-Present)* provide an excellent historical dataset on gold prices which is used in this model.

Macro Economic Growth - Population, GDP, Energy, US Unemployment Rate

Both global and U.S. population data are included in the model. Population growth is generally accepted as a key component in "economic growth" and Lucas provides a clear example of an economic growth model, which builds upon work by Robert Solow and Edward Dennison, where population is the central variable in the mathematical integral of his model for economic growth (Lucas 1988, pg 7). The key idea is that knowledgeable and skilled workers are the primary sources of economic output, and it naturally follows that workers are measured by population. For this model world population was sourced from Our World In Data (Ritchie, Rodés-Guirao, Mathieu, Gerber, Ortiz-Ospina, Hasell, Roser)

A similar assertion can be made in regards to energy. Energy, in all forms, is a key component in economic output, or in aiding workers to produce economic output. For this model a simple estimate of total historical energy consumption was sourced from Our World In Data (Ritchie, Rosado, Roser).

This leaves us with a question. How do we quantify "economic output"? There are various measures including Gross Domestic Product (GDP), Gross National Product (GNP), as well as more tangential measures such as per-capital-income as used by Lucas (Lucas 1988, pg 7). For this model GDP was selected, for simplicity, ubiquitous historical use, and availability of data (U.S. Bureau of Economic Analysis, Gross Domestic Product).

A further element added to population, given the workers and economic output rationale, is the U.S. unemployment rate. Data from the U.S. Bureau of Labor Statistics on the Unemployment Rate is included in the model and was sourced from the Federal Reserve Bank of St. Louis (U.S. Bureau of Labor Statistics, Unemployment Rate).

U.S. Federal Fiscal Metrics - Federal Debt, Federal Tax Receipts

Federal Debt and Federal Tax receipts are the primary means of funding and financing the Federal government. Federal taxes, it can be argued, impact individual decisions on purchases of goods, services, assets, etc. Additionally, Federal stimulus and fiscal policy can be implemented through borrowings. Both are included in the model.

Federal Tax Receipts data from the U.S. Office of Management and Budget data was sourced through the Federal Reserve Bank of St. Louis (U.S. Office of Management and Budget, Federal Receipts). Federal Debt data from the Council of Economic Advisers (US) is also sourced through Federal Reserve Bank of St. Louis (Council of Economic Advisers (US), Gross Federal Debt).

Baseline

Before building a prediction model a baseline must be established with which to compare results. As described earlier, the S&P 500 index appears to exhibits exponential growth similar to the hypothetical compound growth curves plotted for 5% and 7% annual rate of return. Additionally, investors utilize this "annual rate of return" concept and terminology when discussing investment performance. A reasonable staring point for the Baseline model is to model an exponential curve of the form:

$$Y = a(1+g)^x$$

Where $\boldsymbol{a}$ is the value of $\boldsymbol{Y}$ at $\boldsymbol{x = 0}$, and $\boldsymbol{g}$ is the rate of growth and $\boldsymbol{x}$ is time. Utilizing the *curve_fit* method within *Scikit-Learn* the parameters for $\boldsymbol{a}$ and $\boldsymbol{g}$ can be estimated for given dataset. This was the method used for the Baseline.

Data Partitioning

Data for the S&P 500 index is accessed through yFinance python module which sources data from Yahoo! finance (Aroussi). Looking ahead to the machine learning model the S&P 500 data was partitioned into a *training set, validation set,* and a *test set.* For the Baseline model the exponential best fit curve was calculated from the *training* and *validation* data. The growth rate calculated was used to estimate values of the S&P 500 over the range of test data points.

Baseline Results

Figure 4 illustrates the best fit curve estimation base on the *training* and *validation* data. The extrapolation is plotted in yellow and the target predictions are plotted in green over the *test* data set. Note baseline underestimates the compound growth over the *test* data.

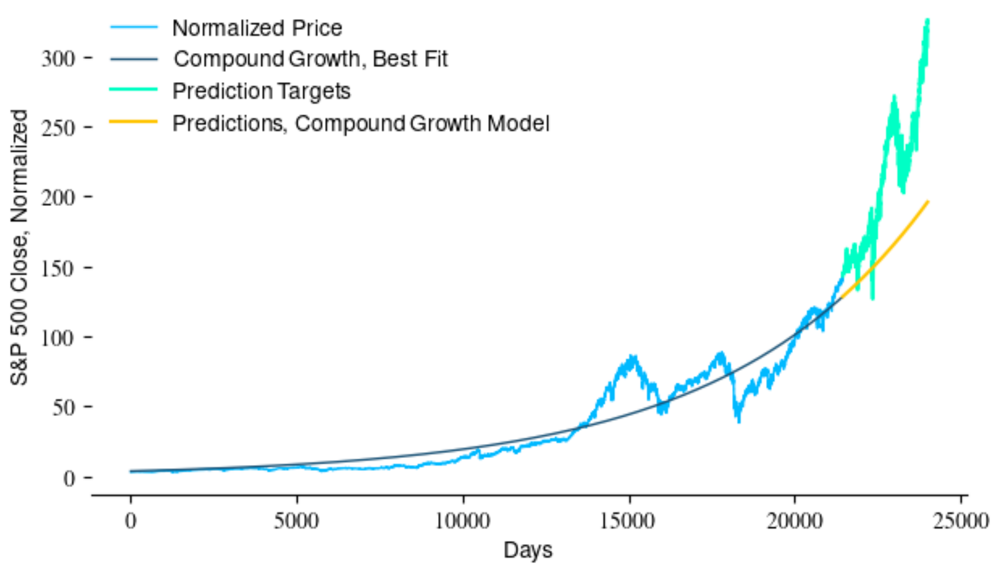

Figure 4 - Baseline prediction results

Deep learning model

Data Configuration

The feature data was aggregated and combined with the S&P 500 index data to create the data set, or *features*. The *labels* or *targets* for the data are the same S&P 500 index data, just shifted by 5 years into the future. It is important to note that most of the feature data is not available on the same frequency. Some data is annual, some data is quarterly or monthly, while the price data is available daily. A decision was made early on to interpolate the missing data, however I've come to learn that deep learning models can handle "sparse" or missing data. While I did not go back and re-configure the dataset, I look forward to exploring the topic of "sparse data" in detail for future iterations of this model.

Data Partitioning

The date range for all data used in the model is from January 1, 1959 to September 30, 2024. This date range was chosen largely as a factor of overlap of available data. Specifically more more modern macroeconomic variables such as the Fed Funds Rate and M2 Money Supply are limited to after 1959. As mentioned earlier the data was partitioned into *trialing*, *validation*, and *test* data sets. In total there are 14 features, 1 label, and 24,014 rows of data. The *training* data is the first 15,000 rows (~63%), the *validation* data is the 15,000 to 18,539 rows (~15%), while the *test* data is the last 5,475 rows (~22%).

Deep Learning Model

The "simple" Deep Learning model used has 3 sequential layers consisting of 2 LSTM layers and a dense layer to produce a single output, specifically the predicted value of the S&P 500 five years out. The model was run with 20 epochs and mean squared error is used for the loss. These configurations and hyper-parameters were chosen primarily for simplicity as starting point.

Results

The results of the target values and predicted values are plotted in figure 5. At first glance this result seems uninspiring. The predicted values, at times are off by a wide margin, say roughly 30% or more. Also, the model does seem to pick up "market shocks" such as the Covid pandemic at around the 900 day mark in figure 5. The model tends to not track the shaper ups and downs of the stock market. Further more the loss as plated in Figure 6 does not indicate a model that is improving, or learning, over the epochs. Rather, the validation loss begins to drift upwards over the epochs which is a sign of over fitting as explained by Chollet and others (Chollet, 2017).

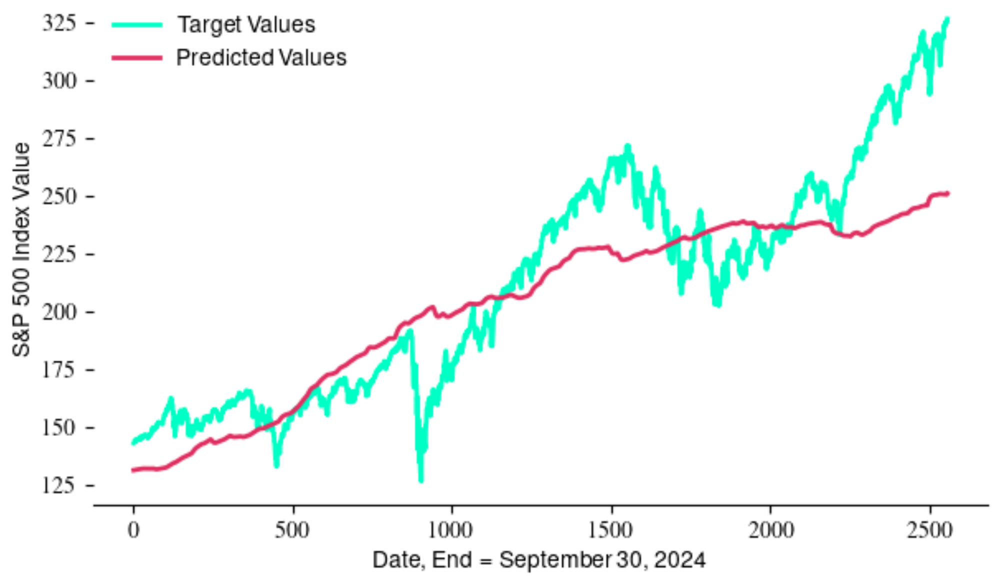

Figure 5 - LSTM Model Results, predictions vs. targets

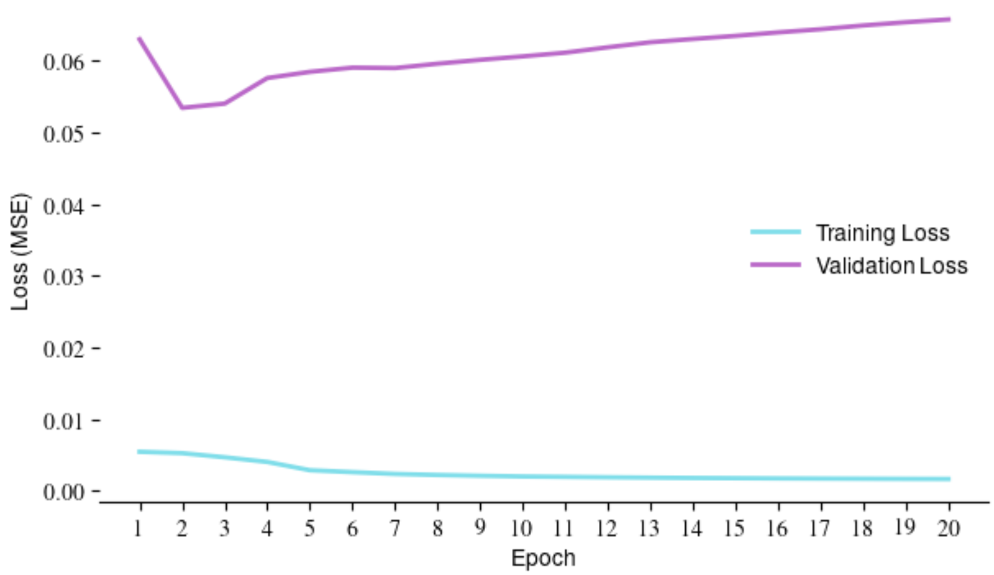

Figure 6 - Loss of LSTM Model

However, there are a couple of promising indications in the results. First, the model does seem to pick up the general upward trend exhibited in compound growth. Second, while this simple model exhibits initial signs over fitting, the validation loss and the training loss are on the same order of magnitude and are not wildly off from one another. Additionally, the upward drift of the validation loss is not pronounced or dramatic. Perhaps the features chosen for the model are not contributing any meaningful information. Feature engineering is certainly an area that warrants more attention for this model. This gives me hope that over fitting may be addressable in fine tuning the hyper-parameters of the model and/or selecting new features.

If we zoom out and see the longer range results in context we start to see some promise as plotted in figure 7. In context of the baseline approach, the Simple Deep Learning model actually appears to perform much better. The Simple Deep Learning model does not exhibit nearly as much under estimation as the Baseline and in fact the estimation error looks more balanced between positive and negative *"misses"*.

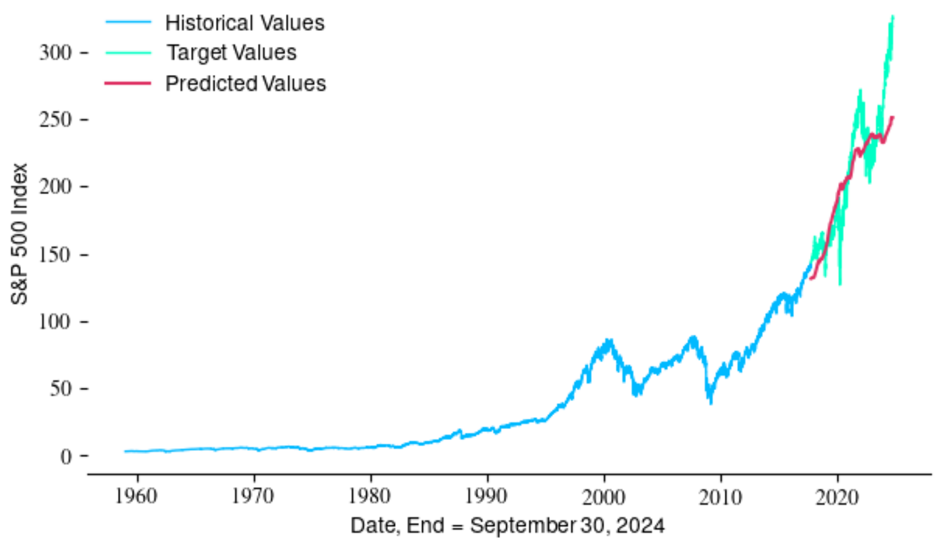

Figure 7 - Simple Deep Learning Model Results in context

Table 2 shows error results side-by-side for the Baseline and the Simple Deep Learning Model. Notably, the Simple Deep Learning model does better on the metrics of *MSE, MAE* and importantly, *Mean Absolute Percentage Error (MAPE)* by a factor of more than 2.

Metric	Baseline	Simple DL
MSE:	3642.45827	662.83498
MAE:	51.66609	19.75187
MAPE:	22.4%	9.0%

Table 2 - Error comparison

Further examination of of *Percentage Error* illustrates some interesting artifacts. First, both models show long tails on the over-estimation side (right). Second the LSTM, or Simple model has a characteristic bi-modal distribution indicating it does not model sharp spikes and dips well. Third the Exponential model shows a chronic under-estimation noted earlier.

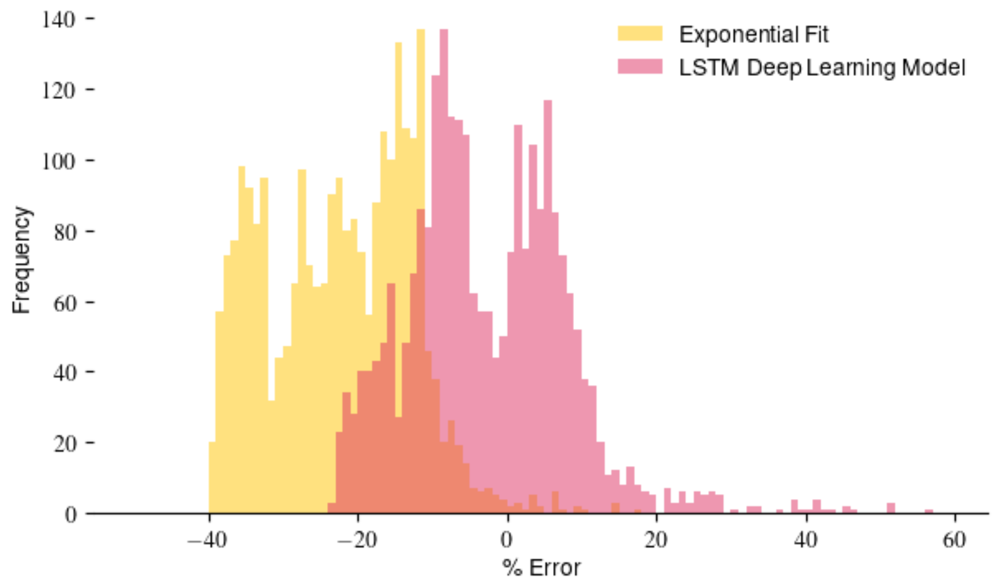

Figure 8 - Histogram of percent error

A challenge was encountered when *Mean Absolution Percentage Error (MAPE)* was attempted as a measure of error. Mean Absolute Percentage Error seems to be the intuitively appropriate measure of error to use given the relation to time and increasing value over time. However, when MAPE was selected as a measure of error, the model would "crash". The source of the error was driven by the normalization of features and data which resulted in a value of 0 for some of the data points. This resulted in impossible values for the MAPE error calculation.

In a brief review literature it appears that this is not a unique situation especially in time series forecasting. In fact Kim and Kim (2016) propose a different approach form similar forecasting applications in which error is measured by *mean arctangent absolute percentage error* (Kim and Kim, 2016). In this approach the arctan function is applied to the ratio created in the absolute percentage error calculation. The advantage is that *arctan* is defined over all values from -∞ to +∞. This provides obvious calculation advantages over the *absolute percentage error* calculation, however, the intuitive advantages of *absolute percentage error* are not necessarily inherent in the the *mean arctangent absolute percentage error (MAAPE).* Generally speaking, the in the MAAPE approach is a comparison of 'angles' which is not nearly as intuitive as percentages. I look forward to exploring further applications of other error approaches in future iterations of the model as this may strengthen the results and intuition of the model.

Conclusions

Did I find the philosophers stone or the Midas touch? No, this is hardly a result ready for Wall Street. However, it is important to note that this was a fairly “simple” deep learning model and the fine tuning of hyper-parameters was very limited. Yet the preliminary results contain some notes of promise and clues for further exploration. Creating a simple deep learning model that significantly outperforms a baseline approach of exponential curve fit and estimation is no small result.

I look forward to further research on this topic to better understand which models might be better suited to the problem at hand. Can I incorporate more complex models and layers including CNNs? Does fine tuning of hyper parameters improve results? Can I expand the feature window from 3 years to some long time period? Can I expand the prediction interval our further than 5 years?

Additionally, I look forwards to investigating the feature variables. Are the variables that can be removed to simplify the model? Do other measures, say Core CPI, or GNP improve or add information to the model? Instead of modeling the S&P 500 returns, can we refactor the data and model to investigate topics such as inflation?

There are myriad of directions this can go. However, the results are better than I could have hoped for, and I feel this is an excellent entry point into applying Deep Learning models to financial markets. I look forward to further study and testing.

References:

Jiang, S. (2024) The effect of monetary policies on inflation: A fiscal perspective. E*conomic Letters, Volume 241,* https://doi.org/10.1016/j.econlet.2024.111805

S&P Global. *The Guage of the U.S. Large Cap Market* https://www.spglobal.com/spdji/en/documents/additional-material/sp-500-brochure.pdf

Hochreiter, S., Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation 9(8):1735-1780, 1997* https://doi.org/10.1162/neco.1997.9.8.1735

Friedman, M. A. (1974) Inflation is Created by Government and No One Else. The Collected Works of Milton Friedman, *The Hoover Institute* https://miltonfriedman.hoover.org/internal/media/dispatcher/271092/full

Federal Reserve Bank of Cleveland, What is inflation? https://www.clevelandfed.org/center-for-inflation-research/inflation-101/what-is-inflation-start#

Ceballos, L., Ng, O. (October 2024) Do investors care about inflation risk? Evidence from global bond portfolio allocation. *Economic Letters Volume 243 11195* https://doi.org/10.1016/j.econlet.2024.111955

Dietrich, A. M. (2024) Consumption categories, household attention, and inflation expectations: Implications for optimal monetary policy. *Journal of Monetary Economics, Volume 147* https://doi.org/10.1016/j.jmoneco.2024.103594

Lucas Jr, R. E. (1988). On the mechanics of economic development. Journal of monetary economics, 22(1), 3-42.

Kim, H., Kim, S. (2016) A new metric of absolute percentage error for intermittent demand forecasts, *International Journal of Forecasting*, Volume 32, Issue 3 https://doi.org/10.1016/j.ijforecast.2015.12.003

Graham, B., Dodd, D. L., Security Analysis *McGraw-Hill Book Company, Inc.*, 1951

Adekoya, A. F., Nti, I. K., Weyori, B. A. (2020, April) A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review* https://doi.org/10.1007/s10462-019-09754-z

Faaljou, H., Mansourfar, G., Rezaei, H. (May 1, 2021) Stock price prediction using deep learning and frequency decomposition. Experience Systems with Applications Volume 169 https://doi.org/10.1016/j.eswa.2020.114332

Won, D., Yoon, S. W., Yun, K. K., Interpretable stock price forecasting model using genetic algorithm-machine learning regressions and best feature subset selection, *Expert System with Applications Volume 213, Part A* https://doi.org/10.1016/j.eswa.2022.118803

Chang, V., Lin, W., Pang, X., Zhou, Y. (January 12, 2018) An innovative neural network approach for stock market prediction. Journal of Supercomputing 76, 2098–2118 (2020). https://doi.org/10.1007/s11227-017-2228-y

Bao W, Yue J, Rao Y (July 14, 2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLOS ONE 12(7): e0180944* https:/doi.org/10.1371/journal.pone.0180944

Chandra, K. S. (2024) Deep learning framework for stock price prediction using long short-term memory. *Soft Computing (2024) 28:10557–10567* https://doi.org/10.1007/s00500-024-09836-3

Yan, X., Yu, P., (December 17, 2018) Stock price prediction based on deep neural networks. *Neural Computing and Applications (2020) 32:1609–1628* https://doi.org/10.1007/s00521-019-04212-x

Smith, A. (1776) An inquiry into the nature and causes of the wealth of nations. *Simon and Brown 2010*

Chollet, F. *Deep Learning with Python.* Manning Publications Co., 2017

Reamer, N., Downing, J., *Investment A History.* Colombia Business School Publishing, 2016 https://doi.org/10.7312/ream16952

Buffett, W. E. (2018) Letter to the Shareholders of Berkshire Hathaway Inc. 2017 https://www.berkshirehathaway.com/letters/2017ltr.pdf

Shiller, R. J. *Irrational Exuberance.* Princeton University Press, 2015

Aroussi, R. *yFinance 0.2.49 Download market data from Yahoo! Finance's API* https://pypi.org/project/yfinance/

Shiller, R. J. *US National Case Shiller Home Price Data.* Shiller Data https://shillerdata.com/

Ritchie, H., Rosado, P., Roser, M. (2020) *Energy Production and Consumption* Our World In Data https://ourworldindata.org/energy-production-consumption

Officer, L. H., Williamson, S. H., *The Price of Gold, 1257-Present.* MeasuringWorth, 2024 http://www.measuringworth.com/gold/

Ritchie H., Rodés-Guirao, L., Mathieu, E., Gerber, M., Ortiz-Ospina, E., Hasell, J., Roser M.

(2023) *Population Growth* Our World in Data https://ourworldindata.org/population-growth

U.S. Bureau of Economic Analysis, Gross Domestic Product [GDP], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/GDP

U.S. Bureau of Labor Statistics, Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/UNRATE

Board of Governors of the Federal Reserve System (US), Federal Funds Effective Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/FEDFUNDS

Moody's, Moody's Seasoned Aaa Corporate Bond Yield [AAA], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/AAA

U.S. Bureau of Economic Analysis, Population [B230RC0Q173SBEA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/B230RC0Q173SBEA https://fred.stlouisfed.org/series/B230RC0Q173SBEA

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCNS], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/CPIAUCNS

Board of Governors of the Federal Reserve System (US), M2 [M2SL], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/M2SL, November 16, 2024.

U.S. Office of Management and Budget, Federal Receipts [FYFR], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/FYFR

Council of Economic Advisers (US), Gross Federal Debt [FYGFD], retrieved from FRED, Federal Reserve Bank of St. Louis https://fred.stlouisfed.org/series/FYGFD