S&P 500 forecasting & Portfolio Construction

Jinghan Chen

Summary

Introduction

Methodology & Data

Analysis & Result

Conclusion & Next Steps

Introduction - Background & Motivation

Effective forecasting models are important to make informed decisions

Standard & Poor 500 (S&P 500) is a well-diversified Market Cap Weighted Index

Potential influential factors:

Macroeconomic conditions

Technical indicators

Portfolio construction strategy

Modern Portfolio Theory (MPT)

Methodology & Data

Methodology

- Monthly index data
  - Key macroeconomic indicators
  - Three technical indicators from each of the general category

- Exploratory Data Analysis
  - Lagged Return & Cross-Correlation

- Modeling & model selection
  - Autoregressive integrated moving average (ARIMA)
  - Seasonal Auto-Regressive Integrated Moving Average with eXogenous (SARIMAX)
  - Recurrent Neural Network (RNN) & Long Short-Term Memory (LSTM)

- Portfolio Construction using MPT

Data – S&P 500 Index & Macroeconomic Indicators

- **Yahoo finance API** (yfinance)
  - ^GSPC
  - Volume, High, Low, Close
- **Federal Reserve Economic Data** (FRED)
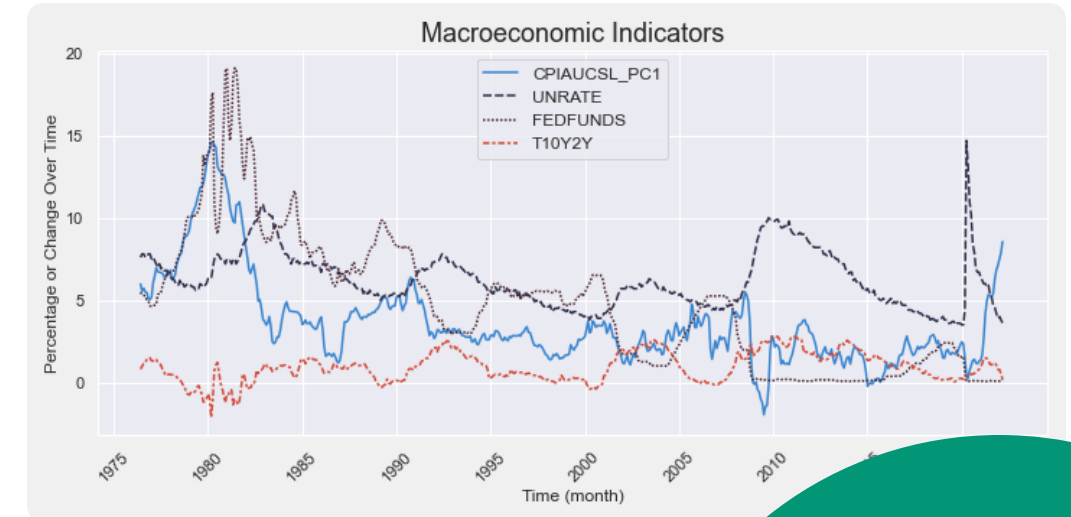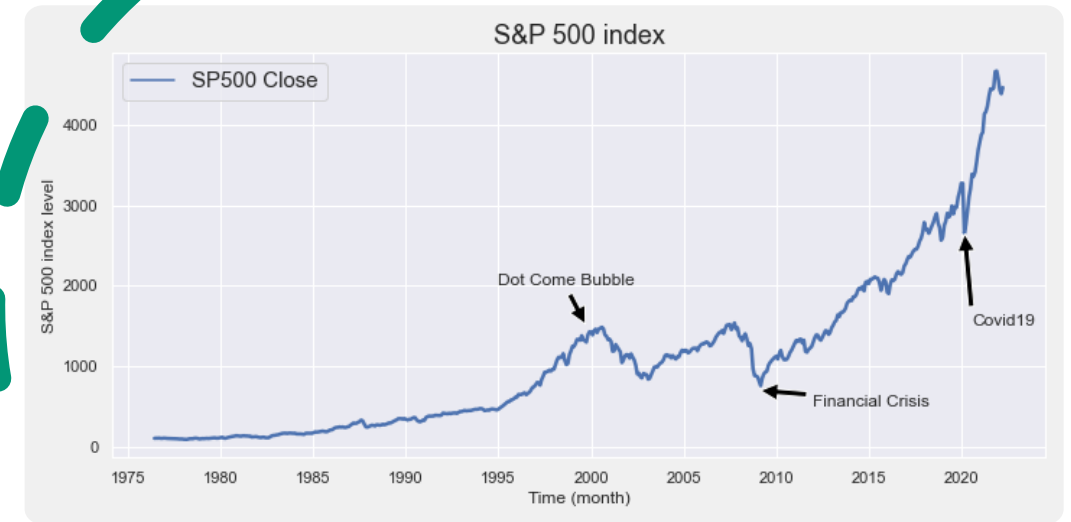
Data – Technical Indicators
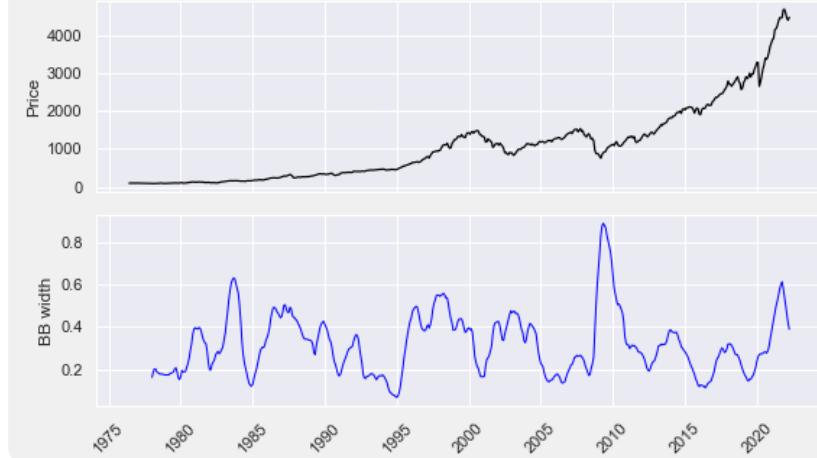
- **Technical Indicators** ([Talib](#))

- **Average Directional Movement Index** (ADX)
    - Strength of a trend. The higher the magnitude of ADX, the stronger the trend

- **Relative Strength Index** (RSI)
    - Momentum of a trend by determining if a stock/ index is overbought or oversold

- **Bollinger Band width**
   - Implied volatility of the stock/ index movement
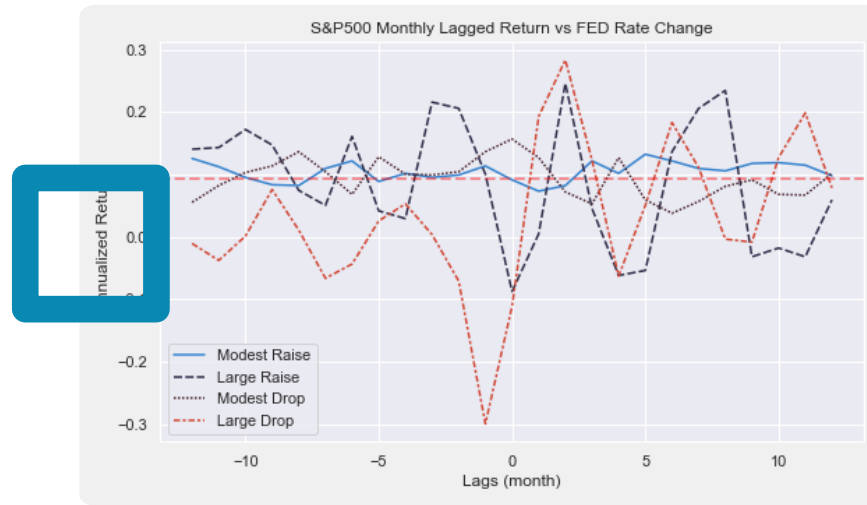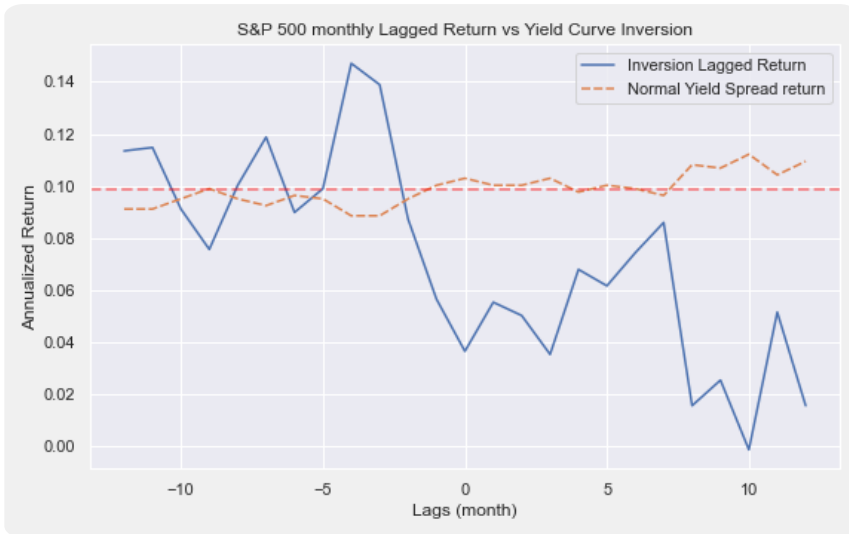
Analysis – Univariate
Lagged Return

Consumer Price Index - Inflation

- Deflationary environment
  - Worst performance, most volatile

- Healthy & low inflationary environment
  - Tend to outperform

- High inflationary environment
  - Tend to outperform before dipping below the average after 5 months

S&P 500 monthly Lagged Return vs Yield Curve Inversion

S&P500 Monthly Lagged Return vs FED Rate Change
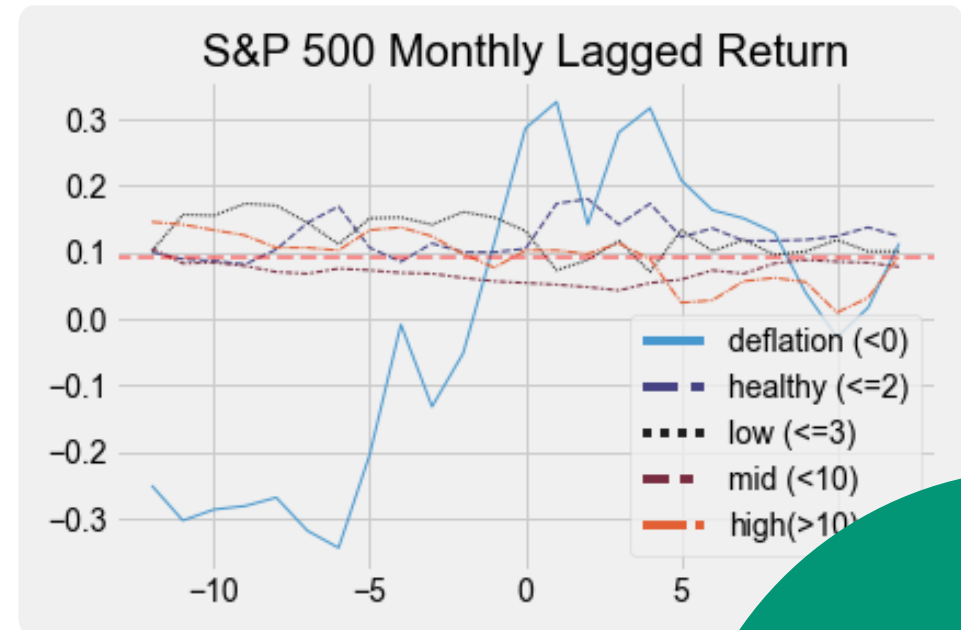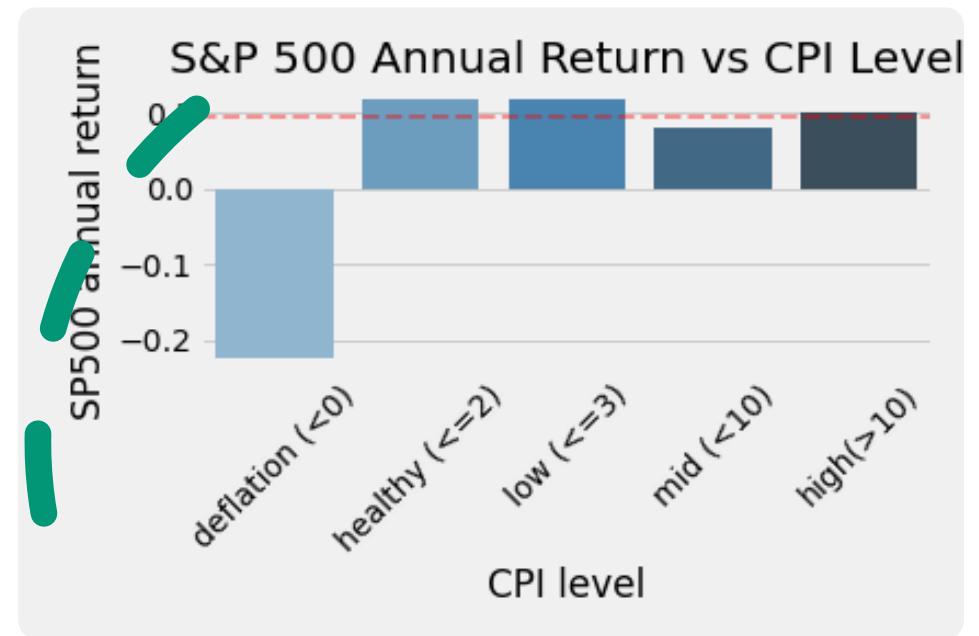
# Impact of the Federal Reserve (FED)

- Normal yield spread is around the market return

- Outperform around 4 months going into the yield curve inversion as investors anticipate FED policies outlook to significantly impact yield curve
  - Rally after the yield curve inverted

- React differently based on the FED rate changes
  - Outperform under a modest raise follow by a small dip after the rate hike
  - Extremely volatile under large drop

Analysis — FED vs Inflation

FED vs Inflation

FED Funds Change level	mean				count			
	drop	large drop	large raise	raise	drop	large drop	large raise	raise
CPI level								
deflation (<0)	0.1522	NaN	NaN	0.4877	8.0	NaN	NaN	6.0
healthy (<=2)	0.2281	-0.6289	0.1862	0.0776	48.0	3.0	1.0	79.0
high(>10)	0.1809	0.0521	0.0571	0.1697	3.0	8.0	10.0	10.0
low (<=3)	0.2488	-0.3987	-0.0701	0.1028	41.0	2.0	1.0	98.0
mid (<10)	0.0791	0.0011	-0.2699	0.0683	87.0	16.0	8.0	121.0
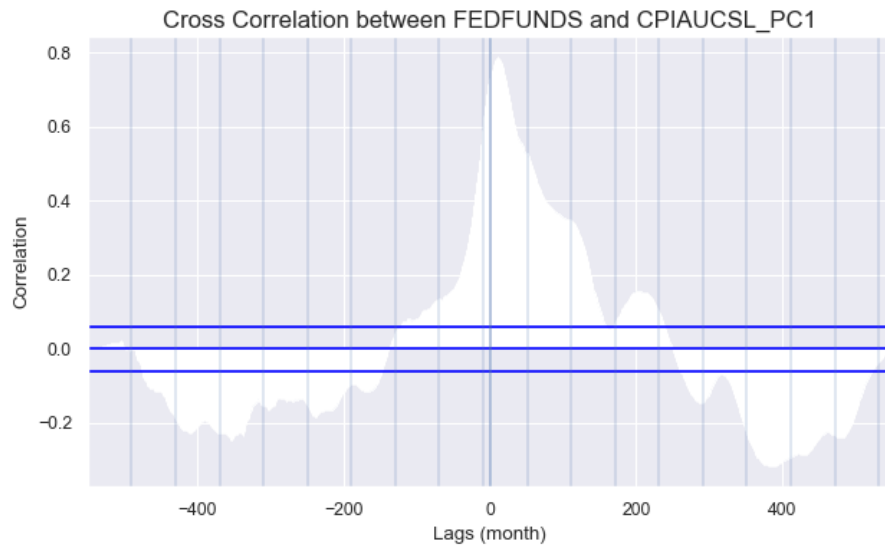
- Tend to raise the FED funds rate when the CPI is raising
- Tend to have negative returns when the FED has large policy changes
- More policy changes in the mid-level inflation environment
- Tends to **raise** the FED funds rate in an inflationary environment to tighten the economy and **drop** the FED funds rates to stimulate the economy

Cross Correlation between FEDFUNDS and CPIAUCSL_PC1

Analysis – Technical Indicators

Technical Indicators

- ADX key levels: **25 & 50**
  - Sideways movement, continue in the direction of the breakout
  - Cross over vs cross under
- RSI key levels: **30 & 70**
  - Overbought vs oversold
  - Cross over 70 > cross over 30
  - Trade when RSI crosses over 70 as it has the highest expected return and lowest volatility and realize profit/ cut loss when RSI crosses under key levels
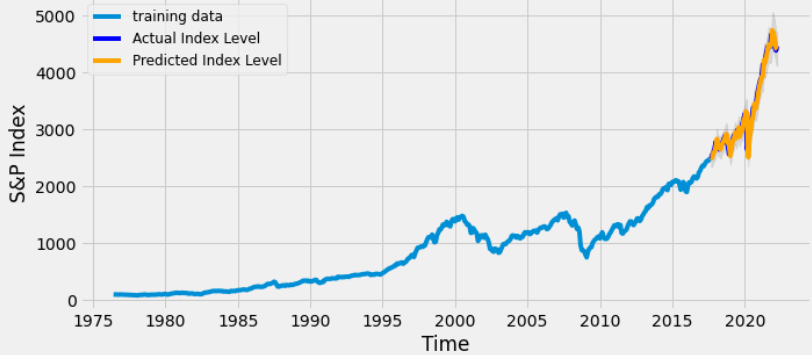
ADX	mean	std
under 25 cross	0.163515	0.185007
over 25 cross	0.038100	0.181219
under 50 cross	0.224131	0.161028
over 50 cross	-0.018400	0.266430

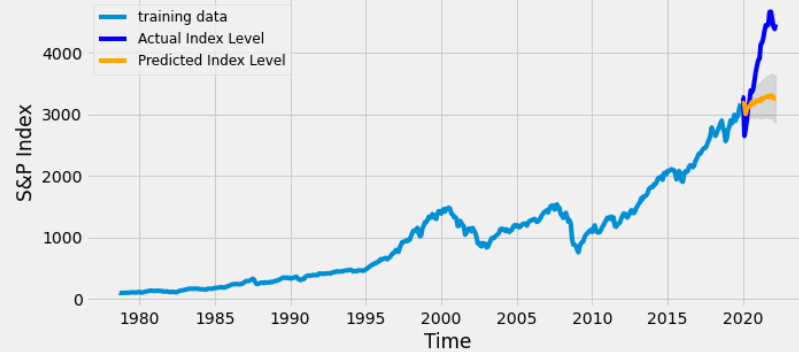| RSI	mean	std
under 30 cross	0.085254	0.343876
over 30 cross	-0.064515	0.301103
under 70 cross	0.150838	0.139509
over 70 cross	0.078462	0.161460

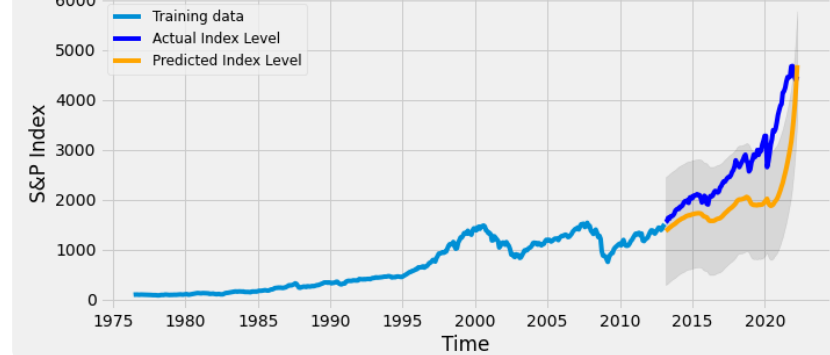# Analysis – Modeling & Portfolio Construction

S&P 500 Index ARIMA Prediction

S&P 500 Index SARIMAX Prediction

S&P 500 Index Single Layer LSTM Prediction
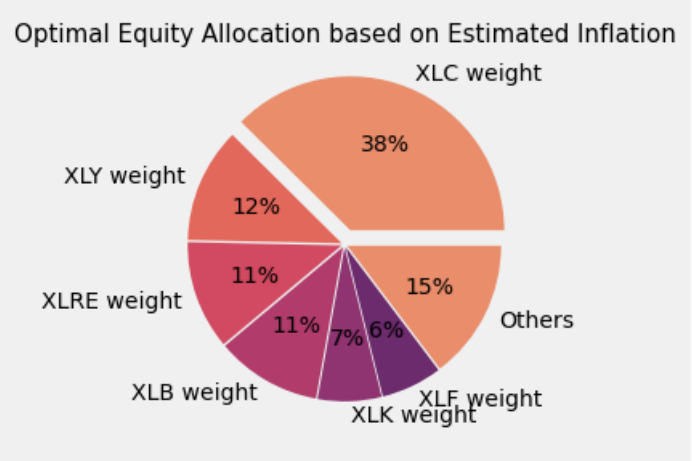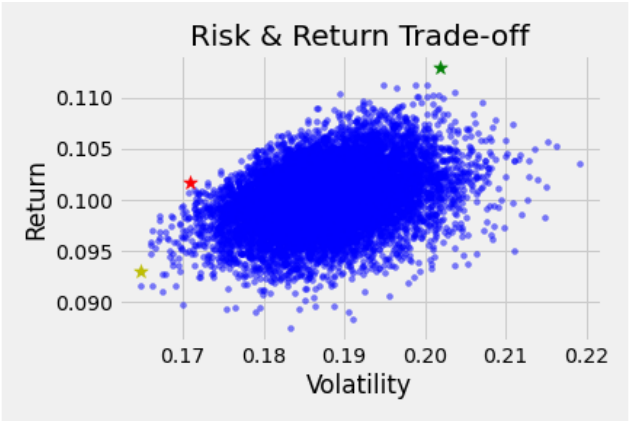
Analysis – Modeling & Model Selection

- **auto_arima** – AIC

- Deep learning model architecture

- Cross-validation

- Root Mean Square Error (RMSE) as the metric to measure model performance across different model classes

Analysis – Forecasting Result

- Best Model: **ARIMA (0, 1, 1)**

- Point estimate of the index level: **4484.35**

- 95% confidence interval: **[4384.23, 4584.47]**

- Annualized monthly return: **4.24%**

- Assuming the average inflation in 2022 is 7.5%, the annual return of the S&P 500 is **11.74%** since equity is inflation hedged

Analysis – Portfolio Construction

Optimal Equity Allocation based on Estimated Inflation

Sector	S&P 500 MWI	S&P 500 EWI
Consumer Discretionary	12.5%	11.8%
Industrials	7.8%	14.2%
Information Technology	29.2%	15.1%
Financials	10.7%	13.2%
Health Care	13.3%	13.2%
Energy	2.7%	4.0%
Consumer Staples	5.9%	6.5%
Utilities	2.5%	5.7%
Real Estate	2.8%	6.0%
Materials	2.6%	5.6%
Communication Services	10.2%	4.6%

- XLC is overweighted
- XLY is about the same weight as the normal S&P index
- XLRE, XLK, and XLF are underweighted
- XLB is overweighted

Conclusion & Next Steps

Conclusion & Next Steps

- Predicting the S&P 500 index and allocating portfolios based on the forecasted return
- Univariate lagged return, the impact of the FED, forecasting model, and portfolio construction
- Validate different trading strategies under certain market conditions
  - Back testing and simulation
- Incorporate return patterns into the model & account for lags
- Develop deep learning methods
  - Multivariate deep learning models & extract the market sentiment from news and social media
- Forecasting inflation to have a better return target to build the optimal portfolio