Quantitative Macroeconomics and Finance with Python

1. Introduction

Machine learning has significantly improved the forecasting and analysis of macroeconomic data for generating positive portfolio returns. In this report, we select **three tickers** and regress them against Goyal-Welch macroeconomic data to generate a strategy that outperforms excess returns. To achieve this, we explore **Ordinary Least Squares (OLS)** with **ridge regression** and **Deep Neural Network models**. By comparing traditional econometric methods with more modern machine learning models, we assess their effectiveness in improving risk-adjusted returns (Sharpe ratios).

2. Ticker Selection

Since Goyal-Welch factors are based on U.S. macroeconomic data, we selected:

1. **SPY (SPDR S&P 500 ETF Trust)** – A benchmark for the U.S. equity market, ideal for predictive models & trading strategies.
2. **TLT (iShares 20+ Year Treasury Bond ETF)** – A bond ETF closely linked to monetary policy and inflation outlook.
3. **GLD (SPDR Gold Shares)** – A commodity ETF representing gold, considered a safe-haven asset during economic uncertainty.

This selection allows for meaningful observations based on the signal weights for each ETF, while the **diversified asset classes** ensure varied results.

3. Data Analysis

3.1 Data Start and End Dates

The datasets for each ticker start at different points:

- **SPY:** February 1993
- **TLT:** August 2002
- **GLD:** December 2004

3.2 Returns and Leverage Trends

- **SPY:** Significant long-term appreciation, with strong post-2008 and post-2020 growth.
- **TLT:** Stable growth from the early 2000s until 2020, followed by a decline due to rising interest rates.
- **GLD:** Strong performance in 2008 and 2020, aligning with economic crises.

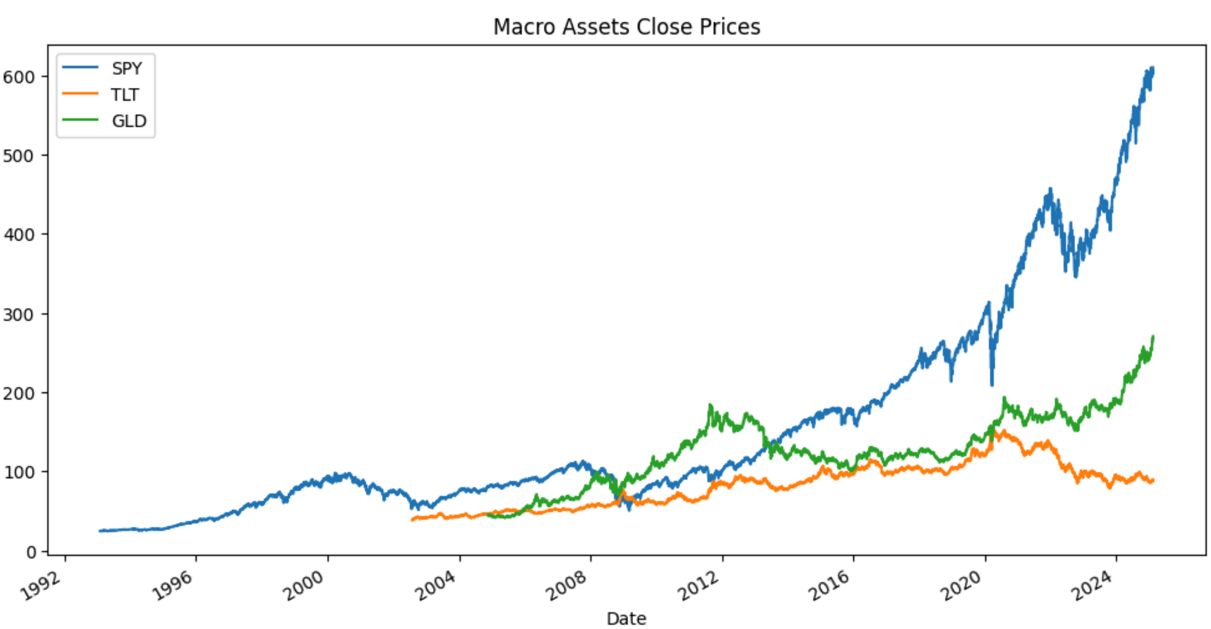


Figure 1: Close Prices of SPY, TLT, and GLD Tickers.

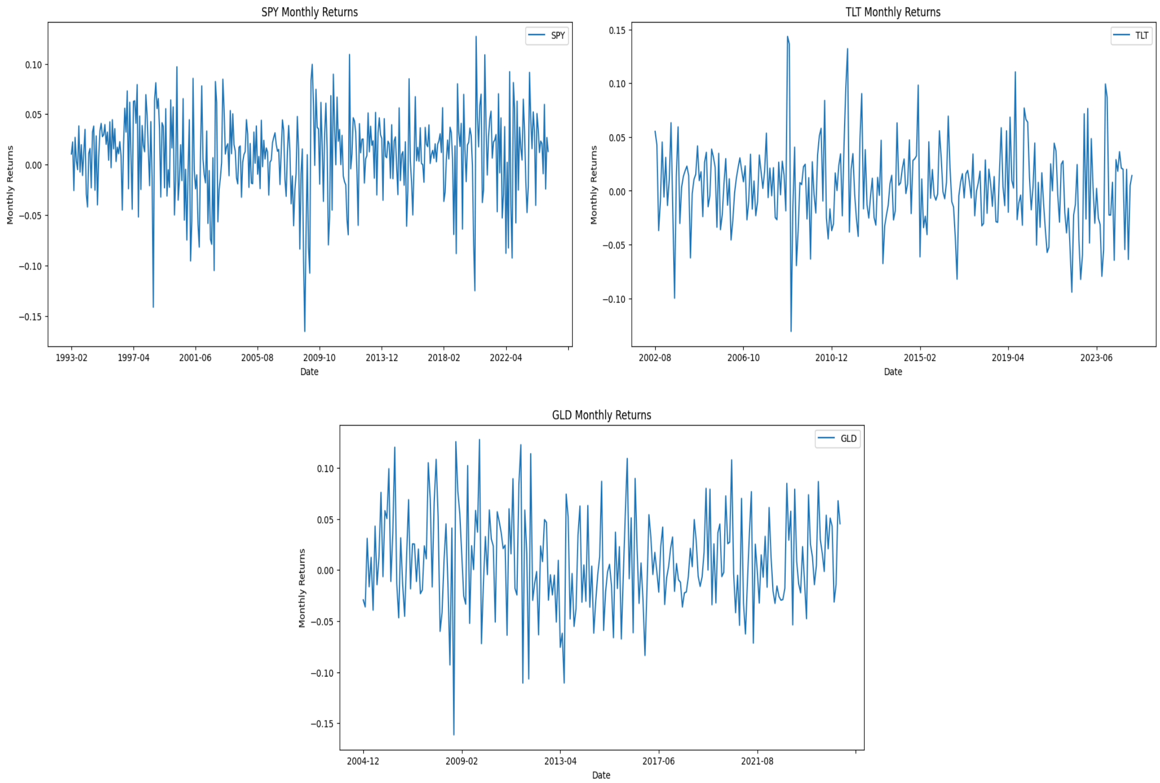


Figure 2: Returns of SPY, TLT and GLD Tickers

3.3 Goyal-Welch Data Variables

Goyal welch data in our use case is a list of 14 variables used to predict equity premiums. The list consists of the following variables:

- D12: Twelve-month moving sums of dividends paid on the S&P 500 index.
- E12: Twelve-month moving sums of earnings on the S&P 500 index.
- b/m: Ratio of book value to market value for the Dow Jones Industrial Average.
- Index: S&P500 price index
- tbl: T-bill rates
- lty: The long-term yield on US bonds
- ltr: Long-term returns on US bonds
- corpr: Corporate Bond returns

- AAA: AAA-rated bond yields
- BAA: BAA-rated bond yields
- infl: Consumer price index is used as Inflation
- Rfree: Risk-free rate is the t-bill rate
- ntis: Net equity issuance calculated using ratio of twelve-month moving sums of net issues by NYSE listed stocks divided by the total market capitalization of NYSE stocks
- svar: Stock Variance of S&P 500

Calculating correlation, we find expected patterns that some of the variables are **highly correlated**. **D12** and **E12** have a **high correlation** of **0.968**. **tbl** is **highly correlated** to **AAA**, **BAA**, **lty**, and **Rfree**. We also see **negative correlations** between **b/m** with **D12**, **E12** and **ntis** with **D12**, **E12**.

	Index	D12	E12	b/m	tbl	AAA	BAA	lty	ntis	Rfree	infl	ltr	corpr	svar
Index	1.000000	0.976952	0.971674	-0.586955	-0.233160	-0.151000	-0.187700	-0.217137	-0.388406	-0.256399	-0.010084	-0.034677	-0.019966	0.024068
D12	0.976952	1.000000	0.967866	-0.566695	-0.195115	-0.089142	-0.124923	-0.156467	-0.449804	-0.226416	-0.016082	-0.010813	0.001501	0.041149
E12	0.971674	0.967866	1.000000	-0.552464	-0.181918	-0.100869	-0.143726	-0.160729	-0.407763	-0.214650	0.000675	-0.018821	-0.017731	0.006350
b/m	-0.586955	-0.566695	-0.552464	1.000000	0.199958	0.160981	0.260409	0.197314	0.034194	0.192322	0.073090	0.009947	-0.000043	0.169768
tbl	-0.233160	-0.195115	-0.181918	0.199958	1.000000	0.888219	0.828238	0.901584	0.032109	0.987950	0.201720	0.053990	0.032672	-0.151561
AAA	-0.151000	-0.089142	-0.100869	0.160981	0.888219	1.000000	0.970474	0.988194	-0.082484	0.880914	0.159846	0.105842	0.101953	-0.047285
BAA	-0.187700	-0.124923	-0.143726	0.260409	0.828238	0.970474	1.000000	0.944715	-0.119387	0.824402	0.094729	0.116887	0.118228	0.074055
lty	-0.217137	-0.156467	-0.160729	0.197314	0.901584	0.988194	0.944715	1.000000	-0.061977	0.890538	0.184884	0.056213	0.054396	-0.099717
ntis	-0.388406	-0.449804	-0.407763	0.034194	0.032109	-0.082484	-0.119387	-0.061977	1.000000	0.032430	-0.055086	-0.045876	-0.042822	-0.039215
Rfree	-0.256399	-0.226416	-0.214650	0.192322	0.987950	0.880914	0.824402	0.890538	0.032430	1.000000	0.177283	0.091004	0.066322	-0.122932
infl	-0.010084	-0.016082	0.000675	0.073090	0.201720	0.159846	0.094729	0.184884	-0.055086	0.177283	1.000000	-0.100218	-0.095453	-0.149562
ltr	-0.034677	-0.010813	-0.018821	0.009947	0.053990	0.105842	0.116887	0.056213	-0.045876	0.091004	-0.100218	1.000000	0.827731	0.092633
corpr	-0.019966	0.001501	-0.017731	-0.000043	0.032672	0.101953	0.118228	0.054396	-0.042822	0.066322	-0.095453	0.827731	1.000000	0.019369
svar	0.024068	0.041149	0.006350	0.169768	-0.151561	-0.047285	0.074055	-0.099717	-0.039215	-0.122932	-0.149562	0.092633	0.019369	1.000000

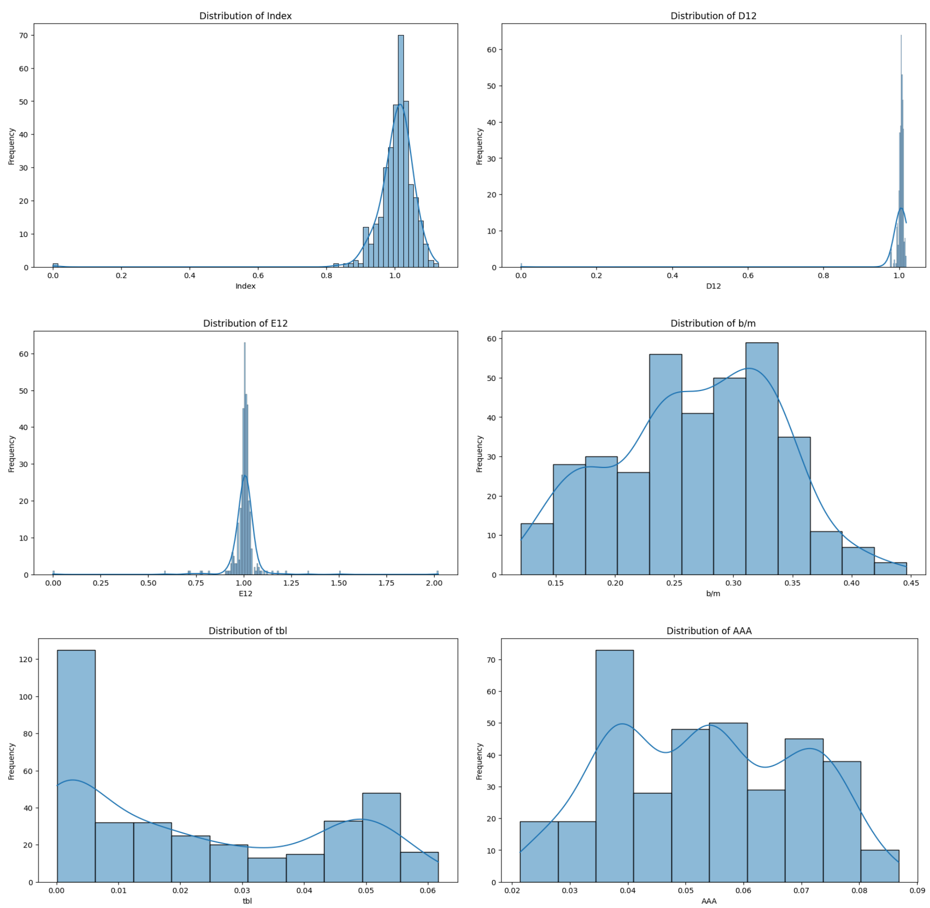
Figure 3: Correlation between Goyal Welch variables

4. OLS and Ridge Regression

4.1 Why Ridge Regression?

In this section, we will look at **deterministic features** directly from **14 Goyal Welch signals** and subsequently construct **Random Features** using **Fourier transformation** for each ticker selected.

Ridge Regression is selected over **LASSO (Least Absolute Shrinkage and Selection Operator)**. **LASSO** makes the weights zero for some signals. In financial modeling, especially when working with macroeconomic variables, every signal carries economic meaning. Setting coefficients to **0** might lead to the **loss of valuable information**.



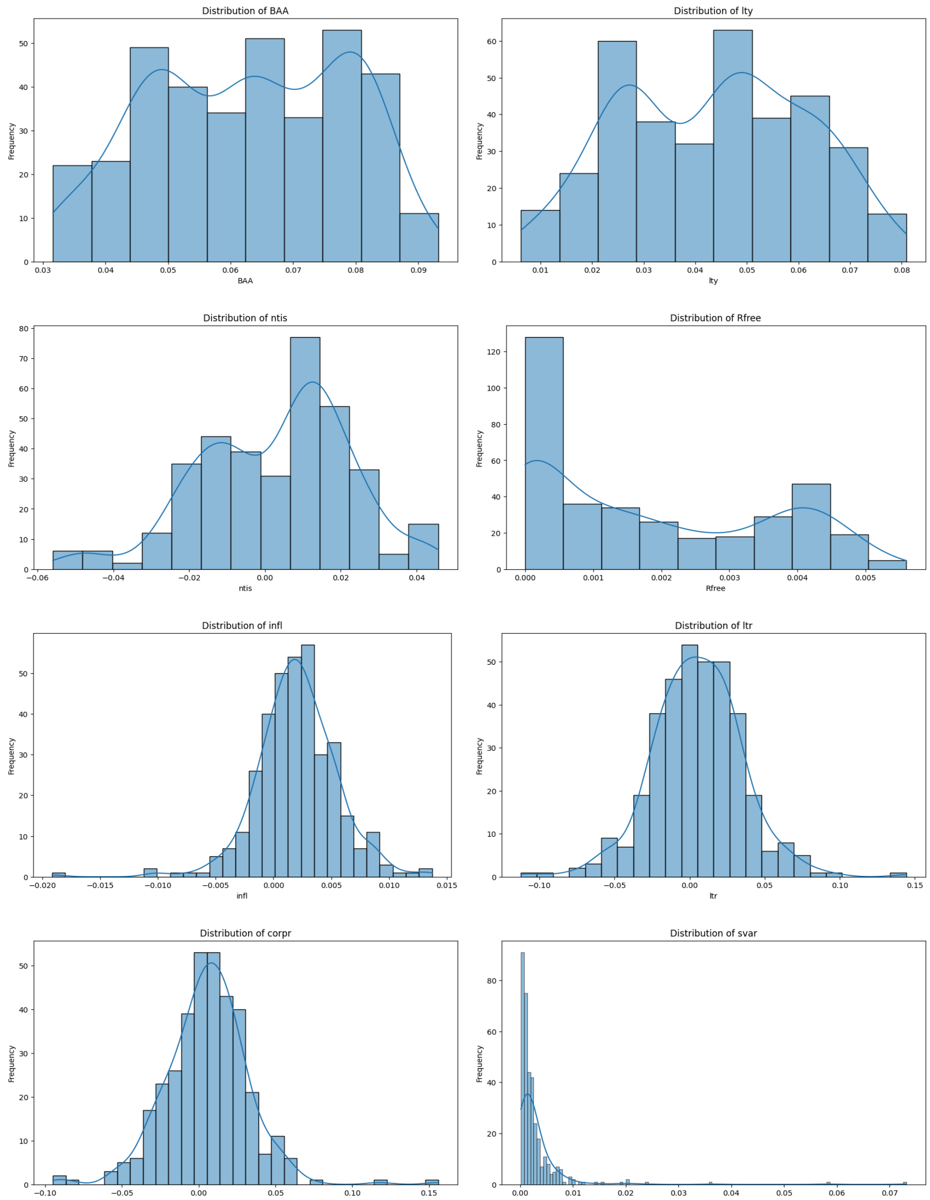


Figure 4: Distributions of Signals

4.2 Results

We observed that as the ridge penalty increases, the model assigns weights that **approach 0** for the three tickers.

4.2.1 **SPY Analysis**

1. Higher earnings per share (E/P) usually indicates **undervaluation** which is bullish for stocks. Thus, a **positive correlation** between E/P and SPY returns is as expected.
2. **Stocks** and **Bonds** markets typically have an **inverse relationship**, because investors allocate capital between riskier equities and safer fixed-income assets, depending on economic conditions. Thus, a **negative correlation** with **AAA**, **BAA**, and **long-term yield** (lty) aligns with the substitution effect - rising bond yields make equities relatively less attractive.
3. Inflation effects are likely already incorporated into earnings and bond calculations, reducing direct impacts on SPY signals.

	1.000000e-08	1.000000e-05	1.000000e-04	1.000000e-03	1.000000e-02	1.000000e-01	1.000000e+00	1.000000e+01	1.000000e+02	1.000000e+03
Index	-0.095420	-0.093636	-0.079863	-0.029565	-0.001275	0.001515	0.001083	0.000204	0.000022	2.256700e-06
D12	-0.000726	-0.000696	0.000878	0.004391	0.003882	0.002346	0.001319	0.000244	0.000027	2.682789e-06
E12	0.135714	0.133914	0.117590	0.054820	0.012432	0.003197	0.001274	0.000225	0.000024	2.467319e-06
b/m	0.007075	0.007460	0.009803	0.016444	0.016214	0.005098	0.000878	0.000117	0.000012	1.232416e-06
tbl	0.020991	0.020987	0.021464	0.020128	0.009216	0.002860	0.000722	0.000109	0.000012	1.167264e-06
AAA	-0.097017	-0.091308	-0.065372	-0.028150	-0.007407	0.000080	0.000648	0.000129	0.000014	1.432258e-06
BAA	-0.039822	-0.042752	-0.052150	-0.036645	-0.009549	-0.000133	0.000689	0.000140	0.000015	1.559436e-06
lty	0.115537	0.112570	0.095019	0.041401	0.005346	0.001648	0.000735	0.000129	0.000014	1.414643e-06
ntis	0.003292	0.003379	0.004385	0.009310	0.008391	0.002379	0.000432	0.000059	0.000006	6.182618e-07
Rfree	-0.007193	-0.007317	-0.008169	-0.005695	0.002781	0.001780	0.000473	0.000070	0.000007	7.536248e-07
infl	0.004531	0.004457	0.004140	0.003968	0.003814	0.001327	0.000179	0.000019	0.000002	1.966745e-07
ltr	-0.033446	-0.032798	-0.027447	-0.007334	0.001801	0.001071	0.000340	0.000056	0.000006	6.080900e-07
corpr	0.057634	0.056567	0.048577	0.019725	0.003426	0.001023	0.000392	0.000068	0.000007	7.447297e-07
svar	0.021860	0.021670	0.020334	0.014138	0.005716	-0.000204	-0.000814	-0.000164	-0.000018	-1.820748e-06

Figure 5: SPY Signal Weights across Ridge Penalties

For SPY, the ridge penalty of **0.1** outperforms the market with a Sharpe Ratio of **0.73**; for random features, none of the models beat the market. As we **increase** the ridge penalty, which mitigates overfitting, we see an increasing trend in the Sharpe ratio, reinforcing the effectiveness of regularisation. Importantly, ensemble averaging of the models is able to achieve a Sharpe Ratio of **0.61**.

```
Sharpe Ratio =excess_returns    0.56
0.0                 0.23
0.00001             0.23
0.0001              0.27
0.001               0.40
0.01                0.57
0.1                 0.73
1.0                 0.55
10.0                0.54
100.0               0.53
1000.0              0.53
1e-08_complex       0.26
1e-05_complex       0.25
0.0001_complex      0.22
0.001_complex       0.24
0.01_complex        0.19
0.1_complex         0.12
1.0_complex         0.19
10.0_complex        0.45
100.0_complex       0.48
1000.0_complex      0.54
mean                0.61
dtype: float64
```

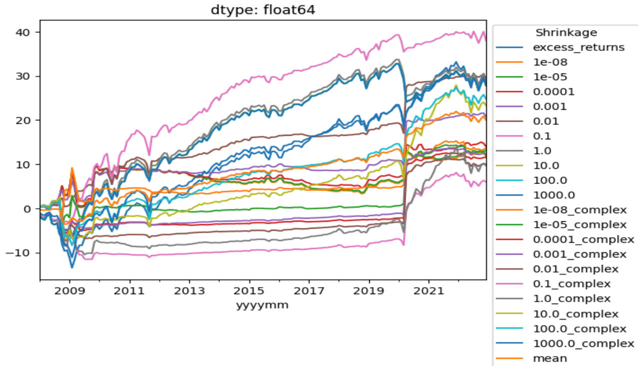


Figure 6: SPY Regression Plot and Sharpe Ratios

4.2.2 TLT

1. **Negative E/P** is the expected relationship as higher **E12** indicates stronger equity returns typically indicating a **higher** opportunity **cost** for **bond investments**. As equities become more attractive, capital would flow out of bonds which **lowers bond prices** and **increases yields**.
2. A **positive beta** between **AAA yields**, **Long-term yield**, and **TLT** is expected. This is due to the positive, direct relationship between treasury yields and bond returns. Higher yields on high-quality bonds increase demand for long-duration, fixed-income securities like TLT.
3. A positive relationship with **Stock Variance** underscores the **flight-to-safety** effect. This refers to the fact that increased

market volatility might drive investors into bonds as a risk-averse asset class.

	1.000000e-08	1.000000e-05	1.000000e-04	1.000000e-03	1.000000e-02	1.000000e-01	1.000000e+00	1.000000e+01	1.000000e+02	1.000000e+03
Index	-0.076228	-0.077208	-0.080116	-0.053933	-0.017481	-0.002634	0.000505	0.000138	1.561462e-05	1.580889e-06
D12	0.088479	0.089745	0.094068	0.064580	0.012173	0.001459	0.001093	0.000215	2.357722e-05	2.380962e-06
E12	-0.113128	-0.112347	-0.105979	-0.073634	-0.025829	-0.004274	-0.000463	-0.000046	-4.529121e-06	-4.526540e-07
b/m	0.063898	0.062666	0.053382	0.018457	-0.001083	0.000079	0.000495	0.000098	1.069674e-05	1.080012e-06
tbl	0.065754	0.063789	0.051965	0.023068	-0.003486	-0.004611	-0.001416	-0.000202	-2.122431e-05	-2.133978e-06
AAA	0.385202	0.366156	0.252321	0.064213	0.019379	0.005926	0.001338	0.000205	2.197721e-05	2.213475e-06
BAA	-0.502665	-0.486817	-0.385358	-0.143095	-0.014266	0.001549	0.000662	0.000111	1.201205e-05	1.210976e-06
lty	0.039933	0.045788	0.076110	0.075032	0.025340	0.006461	0.001294	0.000195	2.077014e-05	2.091512e-06
ntis	-0.046810	-0.045496	-0.036520	-0.009795	0.002275	0.000967	0.000434	0.000076	8.278239e-06	8.349070e-07
Rfree	-0.111231	-0.109010	-0.094981	-0.054607	-0.015767	-0.005475	-0.001510	-0.000214	-2.254747e-05	-2.267055e-06
infl	0.019915	0.020808	0.026472	0.038004	0.028282	0.006631	0.001084	0.000147	1.551237e-05	1.559778e-06
ltr	-0.044883	-0.041914	-0.022932	0.018207	0.015559	0.002167	0.000149	0.000007	5.332326e-07	5.153371e-08
corpr	0.038970	0.037233	0.024796	-0.014205	-0.017374	-0.003566	-0.000570	-0.000077	-8.103809e-06	-8.147886e-07
svar	0.144274	0.141962	0.126682	0.082342	0.033631	0.005582	-0.000210	-0.000114	-1.328332e-05	-1.349000e-06

Figure 7: TLT Signal Weights across Ridge Penalties

In contrast to theoretical expectations, **TLT** shows a poor Sharpe Ratio of **0.03**, suggesting little predictive power. **Lower ridge penalties** are shown to **outperform** the larger ridge penalty cases both for deterministic and random features, suggesting that some flexibility in coefficients is necessary to capture meaningful signals. The highest **TLT model SR** is **0.11** which outperforms the market. However, random Features mostly fail to provide a positive Sharpe Ratio. Most models struggle around the 2020 Covid crash.

```
Sharpe Ratio=excess_returns    0.03
0.0                  -0.04
0.00001              -0.03
0.0001                0.01
0.001                 0.08
0.01                  0.08
0.1                   0.04
1.0                   0.02
10.0                  0.03
100.0                 0.03
1000.0                0.03
1e-08_complex         0.11
1e-05_complex         0.08
0.0001_complex        0.03
0.001_complex        -0.05
0.01_complex         -0.15
0.1_complex          -0.22
1.0_complex          -0.31
10.0_complex         -0.26
100.0_complex        -0.08
1000.0_complex        0.01
mean                 -0.04
dtype: float64
```

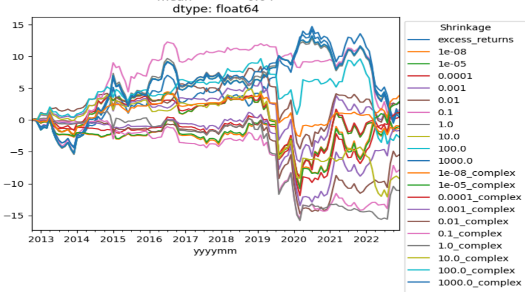


Figure 8: TLT Regression Plot and Sharpe Ratios

4.2.3 GLD Analysis

1. As **gold** is considered a **safe asset** and as a hedge during financial stress a positive beta with **AAA, BAA, corporate returns**, and stock variance is expected. These factors serve as indicators of financial stress, prompting investors to shift towards gold as a protective asset.
2. As **long-term bond yield** and **returns increase**, Bonds will be more attractive assets than Gold, as gold is non-yielding. A **negative beta** is an interesting result.

	1.000000e-08	1.000000e-05	1.000000e-04	1.000000e-03	1.000000e-02	1.000000e-01	1.000000e+00	1.000000e+01	1.000000e+02	1.000000e+03
Index	-0.356993	-0.349376	-0.291068	-0.092496	0.002662	0.005456	0.002345	0.000395	0.000043	4.294519e-06
D12	0.312412	0.308526	0.271956	0.113327	0.018848	0.007054	0.002837	0.000475	0.000051	5.157615e-06
E12	0.089363	0.086033	0.065820	0.026458	0.015183	0.003797	0.000483	0.000052	0.000005	5.239638e-07
b/m	-0.113800	-0.110085	-0.086779	-0.034730	-0.000129	0.004507	0.001475	0.000240	0.000026	2.596315e-06
tbl	-0.053751	-0.055335	-0.059491	-0.032065	-0.001857	0.002282	0.000331	-0.000015	-0.000003	-2.838709e-07
AAA	0.450288	0.420687	0.268119	0.072503	0.023244	0.009161	0.002526	0.000377	0.000040	4.021954e-06
BAA	0.109389	0.111483	0.111409	0.061618	0.027786	0.008945	0.001907	0.000264	0.000028	2.786951e-06
lty	-0.402932	-0.378232	-0.243397	-0.036769	0.014007	0.008180	0.002363	0.000348	0.000037	3.708263e-06
ntis	0.103752	0.100747	0.082255	0.042541	0.021349	0.006836	0.001180	0.000156	0.000016	1.645044e-06
Rfree	0.099474	0.098755	0.089934	0.040614	0.006957	0.002595	0.000262	-0.000029	-0.000004	-4.345714e-07
infl	-0.022885	-0.024063	-0.031360	-0.046635	-0.040766	-0.010365	-0.000524	0.000015	0.000003	2.958373e-07
ltr	-0.275414	-0.269346	-0.232275	-0.133741	-0.034270	-0.001893	-0.000223	-0.000038	-0.000004	-4.129437e-07
corpr	0.231719	0.230304	0.218601	0.158136	0.061466	0.011246	0.000905	0.000054	0.000005	4.544011e-07
svar	0.006698	0.006336	0.006165	0.012431	0.010727	0.001801	-0.001475	-0.000316	-0.000035	-3.520809e-06

Figure 9: GLD Signal Weights across Ridge Penalties

3. The **ridge penalty** of **0.1** works best, possibly showing **overfitting** for **lower** and **underfitting** for **higher values.** Interestingly, the highest Sharpe Ratio of **0.52** occurs at a much **larger penalty** (**1000.0**), suggesting that **extreme regularisation** may help filter out the noise and emphasize more persistent economic relationships.

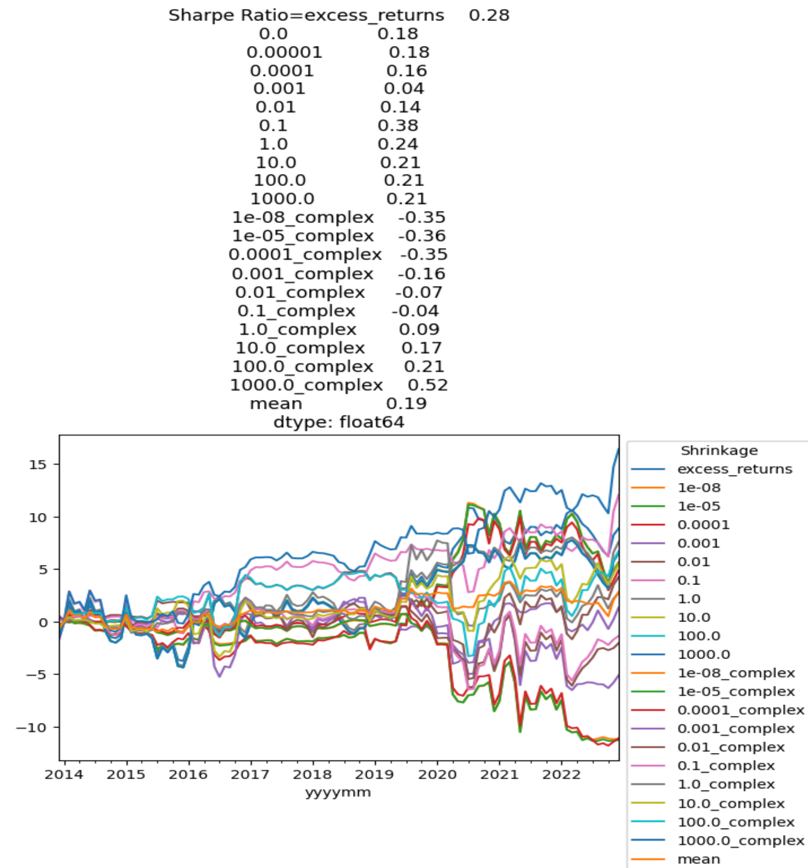


Figure 10: GLD Regression Plot and Sharpe Ratios

4.3 Alpha Analysis

Alpha represents the **excess return** that cannot be explained by regressed market factors. A higher alpha indicates the predicted returns have **unexplained excess** over regressed excess returns, while **negative** means the opposite. A **negative beta** indicates a possible defensive behavior related to regressed data.

For the best-performing random features model, we see an alpha of **0.1** for **SPY**, **0.5** for **TLT**, and **3.2** for **GLD**.

```
Model: 1000
apha: 0.7  Excess Returns beta: 5.6

Model: 100
apha: 0.7  Excess Returns beta: 5.6

Model: 10
apha: 0.7  Excess Returns beta: 5.6

Model: 1
apha: 0.7  Excess Returns beta: 5.5

Model: 0.1
apha: 1.2  Excess Returns beta: 5.1

Model: 0.001
apha: 1.7  Excess Returns beta: 0.8

Model: 0.0001
apha: 1.8  Excess Returns beta: -0.4

Model: 1e-05
apha: 1.7  Excess Returns beta: -0.6

Model: 1e-08
apha: 1.7  Excess Returns beta: -0.7
```

```
1000.0_complex
alpha: 0.1 Excess Returns beta: 34.3 1000 beta: -1.8

100.0_complex
alpha: 0.8 Excess Returns beta: 24.4 100 beta: -6.4

10.0_complex
alpha: 0.9 Excess Returns beta: 11.9 10 beta: -5.2

1.0_complex
alpha: -0.1 Excess Returns beta: 6.8 1 beta: -4.8

0.1_complex
alpha: 0.4 Excess Returns beta: 3.7 0.1 beta: -2.9

0.001_complex
alpha: -1.2 Excess Returns beta: 1.8 0.001 beta: 3.0

0.0001_complex
alpha: -1.0 Excess Returns beta: 2.3 0.0001 beta: 1.9

1e-05_complex
alpha: 0.3 Excess Returns beta: 1.6 1e-05 beta: 2.4

1e-08_complex
alpha: 1.2 Excess Returns beta: -0.7 1e-08 beta: 0.1
```

Figure 11: Alpha and Beta for SPY regressions

```
Model: 1000
apha: 0.0  Excess Returns beta: 13.8

Model: 100
apha: 0.0  Excess Returns beta: 13.8

Model: 10
apha: 0.0  Excess Returns beta: 13.4

Model: 1
apha: -0.0  Excess Returns beta: 10.5

Model: 0.1
apha: 0.1  Excess Returns beta: 2.5

Model: 0.001
apha: 0.3  Excess Returns beta: -2.7

Model: 0.0001
apha: 0.1  Excess Returns beta: -1.8

Model: 1e-05
apha: -0.1  Excess Returns beta: -1.1

Model: 1e-08
apha: -0.1  Excess Returns beta: -1.0
```

```
Model: 1000.0_complex
alpha: -0.1 Excess Returns beta: -0.9 1000 beta: 5.9

Model: 100.0_complex
alpha: -0.2 Excess Returns beta: -3.2 100 beta: 3.1

Model: 10.0_complex
alpha: -1.1 Excess Returns beta: -4.6 10 beta: 4.3

Model: 1.0_complex
alpha: -2.3 Excess Returns beta: -6.4 1 beta: 5.7

Model: 0.1_complex
alpha: -1.4 Excess Returns beta: -7.2 0.1 beta: 6.9

Model: 0.001_complex
alpha: -0.6 Excess Returns beta: -1.8 0.001 beta: 8.7

Model: 0.0001_complex
alpha: 0.2 Excess Returns beta: -2.7 0.0001 beta: 8.8

Model: 1e-05_complex
alpha: 0.5 Excess Returns beta: -2.7 1e-05 beta: 7.3

Model: 1e-08_complex
alpha: 0.5 Excess Returns beta: -2.7 1e-08 beta: 7.1
```

Figure 12: Alpha and Beta for TLT regressions

```
Model: 1000
apha: -0.9  Excess Returns beta: 12.3

Model: 100
apha: -0.9  Excess Returns beta: 12.3

Model: 10
apha: -0.8  Excess Returns beta: 11.9

Model: 1
apha: -0.2  Excess Returns beta: 8.6

Model: 0.1
apha: 1.2  Excess Returns beta: 0.9

Model: 0.001
apha: 0.8  Excess Returns beta: -2.2

Model: 0.0001
apha: 0.0  Excess Returns beta: 2.5

Model: 1e-05
apha: -0.1  Excess Returns beta: 3.4

Model: 1e-08
apha: -0.1  Excess Returns beta: 3.5
```

```
Model: 1000.0_complex
alpha: 3.2 Excess Returns beta: -2.6 1000 beta: 3.0

Model: 100.0_complex
alpha: 2.7 Excess Returns beta: -4.2 100 beta: 3.3

Model: 10.0_complex
alpha: 1.9 Excess Returns beta: -5.1 10 beta: 4.8

Model: 1.0_complex
alpha: 1.0 Excess Returns beta: -4.4 1 beta: 4.1

Model: 0.1_complex
alpha: -0.9 Excess Returns beta: -6.0 0.1 beta: 7.2

Model: 0.001_complex
alpha: -0.3 Excess Returns beta: -2.1 0.001 beta: 2.7

Model: 0.0001_complex
alpha: -1.1 Excess Returns beta: -2.6 0.0001 beta: -1.3

Model: 1e-05_complex
alpha: -1.2 Excess Returns beta: -1.0 1e-05 beta: -2.2

Model: 1e-08_complex
alpha: -1.2 Excess Returns beta: -0.7 1e-08 beta: -2.3
```

Figure 13: Alpha and Beta for GLD regressions

4.4 Changing Normalization Method

The choice of **normalization** technique **significantly impacts** model stability and predictive accuracy. The **Min-Max normalization** method has a drawback when **handling extreme values**, such as those observed during the **2008 financial crisis**, as it **compresses** signals into a **narrow range**. Alternative normalization techniques include **Z-score**, **robust**, and **log normalization**. However, since Goyal-Welch signals are not **highly skewed** and **ridge regression** requires centering around **0**, log normalization can be omitted.

Given that the **signal distributions** are **nearly symmetric**, Z-score normalization is a strong candidate for ensuring a stable and well-scaled dataset.

Results:

- **SPY** improved Sharpe Ratios with **z-score** and **robust normalization**.
- **TLT** performed worse with the **z-score**.
- **GLD** showed **Sharpe Ratio** improvement with **robust normalization**.

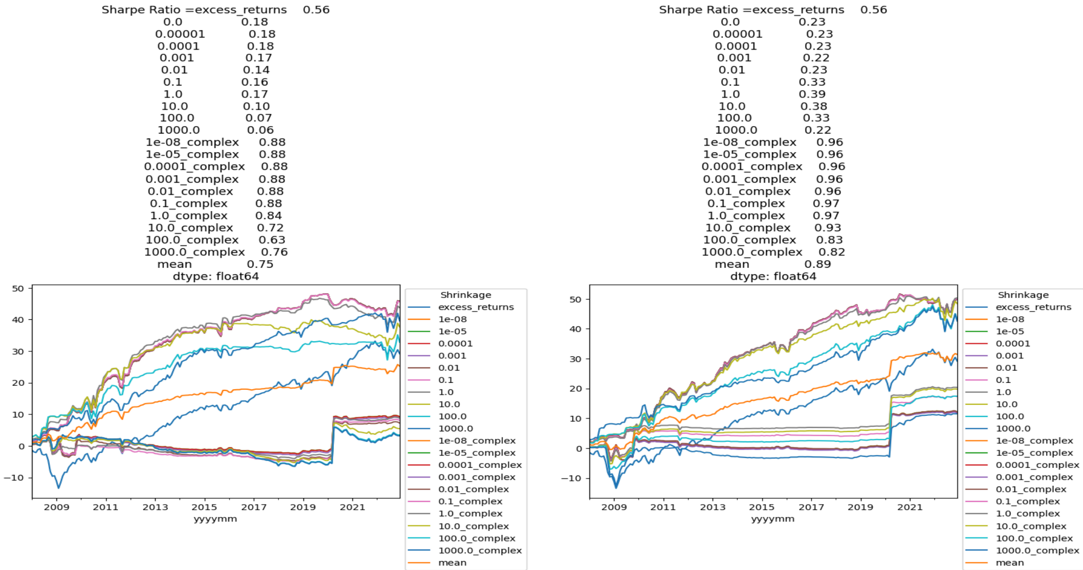


Figure 14: Results with z-score and robust normalization for SPY

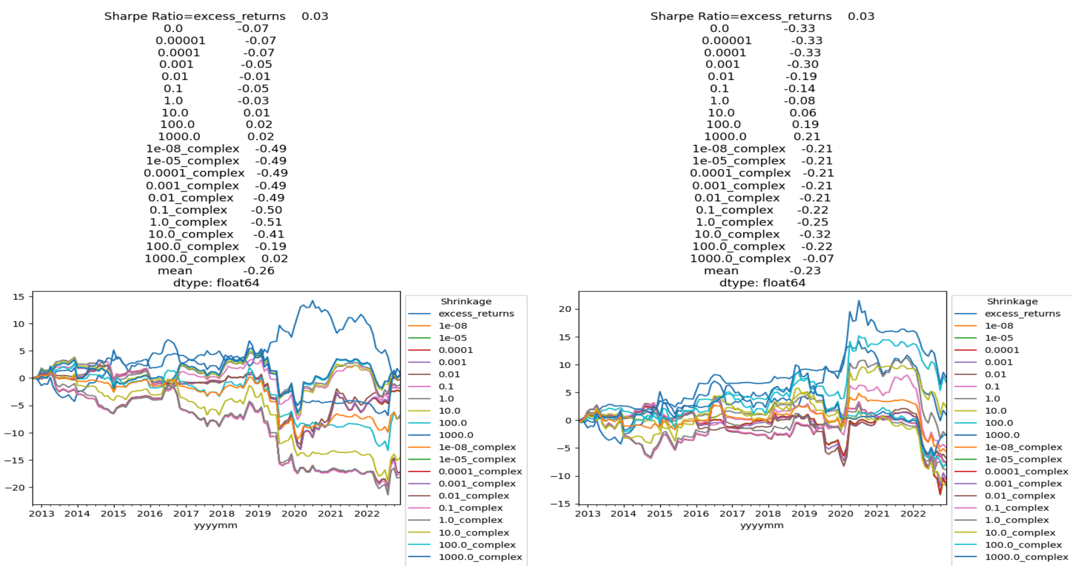


Figure 15: Results with z-score and robust normalization for TLT

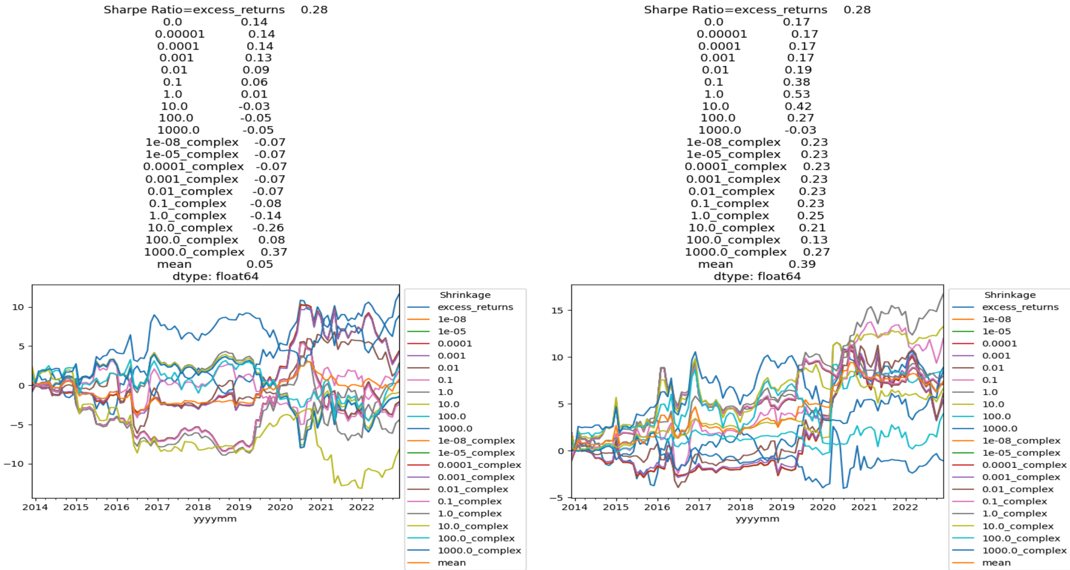


Figure 16: Results with z-score and robust normalization for GLD

5.1 Deep Learning

In this section, we explore **deep learning methods** and **adjust parameters** to achieve **optimal results**. **Deep Learning excels** over **OLS regression** when data exhibits **complex nonlinear relationships**. It achieves this by constructing interconnected layers with **nonlinear activation functions**.

<u>5.1.1 SPY Result Analysis</u>

In an analysis of the **SPY** ticker, Best Sharpe ratios are selected when regressed against different parameters. **Final models** are constructed using the **averaging method** of models generated by **different sample seeds**, The **second** is **ridge regressing** over the **last hidden layer**. The results of the regression are shared in the table below.

Category	Model 1	Model 2	Model 3
Number of seed	25	25	45
Width	40	40	230
Ridge Penalty	0.007	0.005	0.001
Learning Rate	0.085	0.08	0.017
Number of Epochs	185	200	300
Actual Sharpe ratio	0.56		
Averaging Predictions	0.57	0.47	0.71
Ridge Regression	0.89	1.03	0.51

The table demonstrates that **Model 1** outperforms the market in both **ridge regression** and **averaging predictions**. However, **Model 2** achieves the **highest** performance with **ridge regression**, attaining an SR of **1.03** compared to the market SR of **0.56**, by utilizing a **lower ridge penalty** and a **higher number of epochs**. Meanwhile, **Model 3** delivers **strong results** using the **averaging predictions** with **0.71 SR** method rather than **ridge regression**, benefiting from a slightly **higher number of seeds**. This can be attributed to the **averaging approach**, which produces smoother results. The graph comparison is shown below.

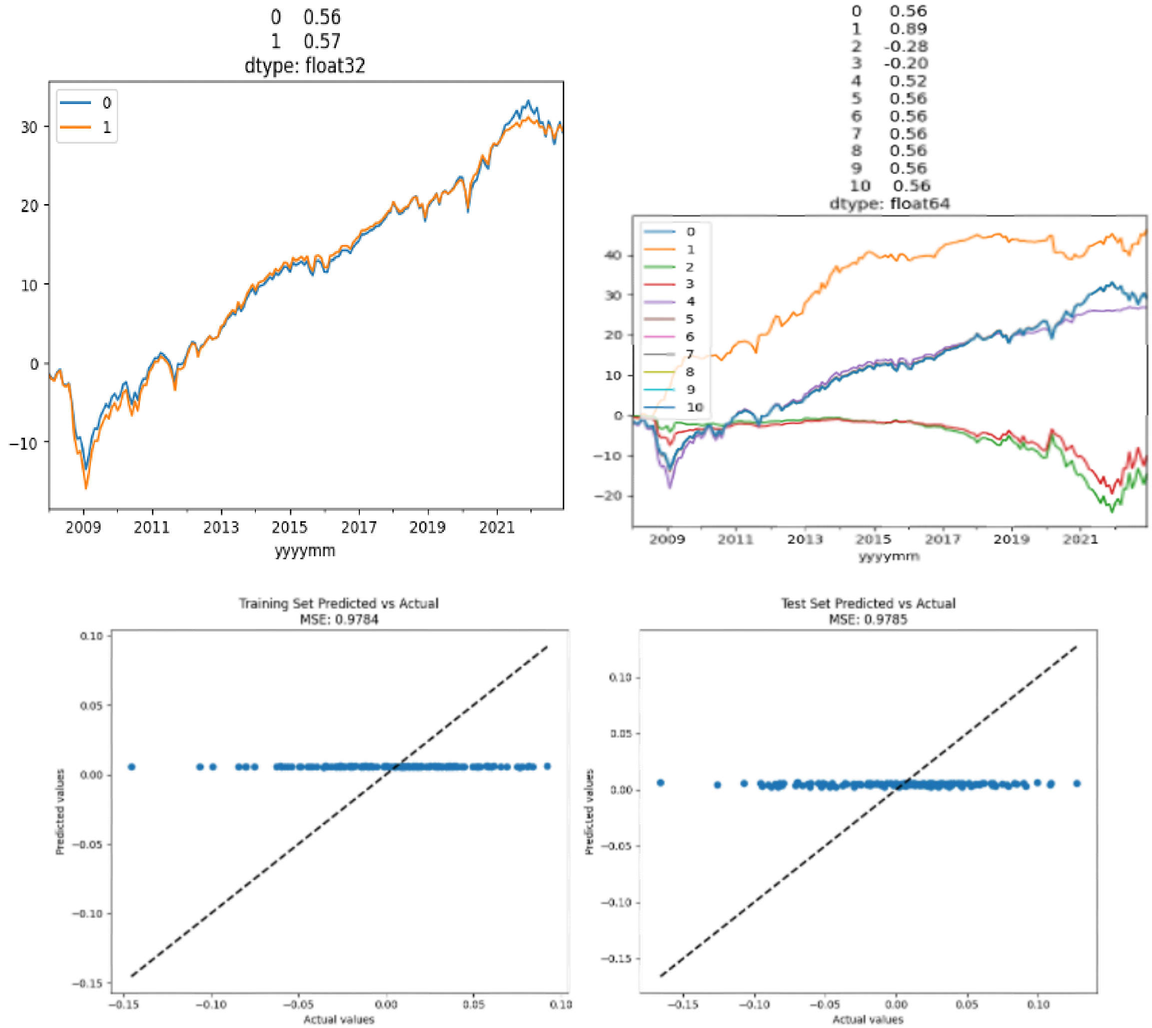


Figure 17: Model 1 results using averaging predictions and ridge regression

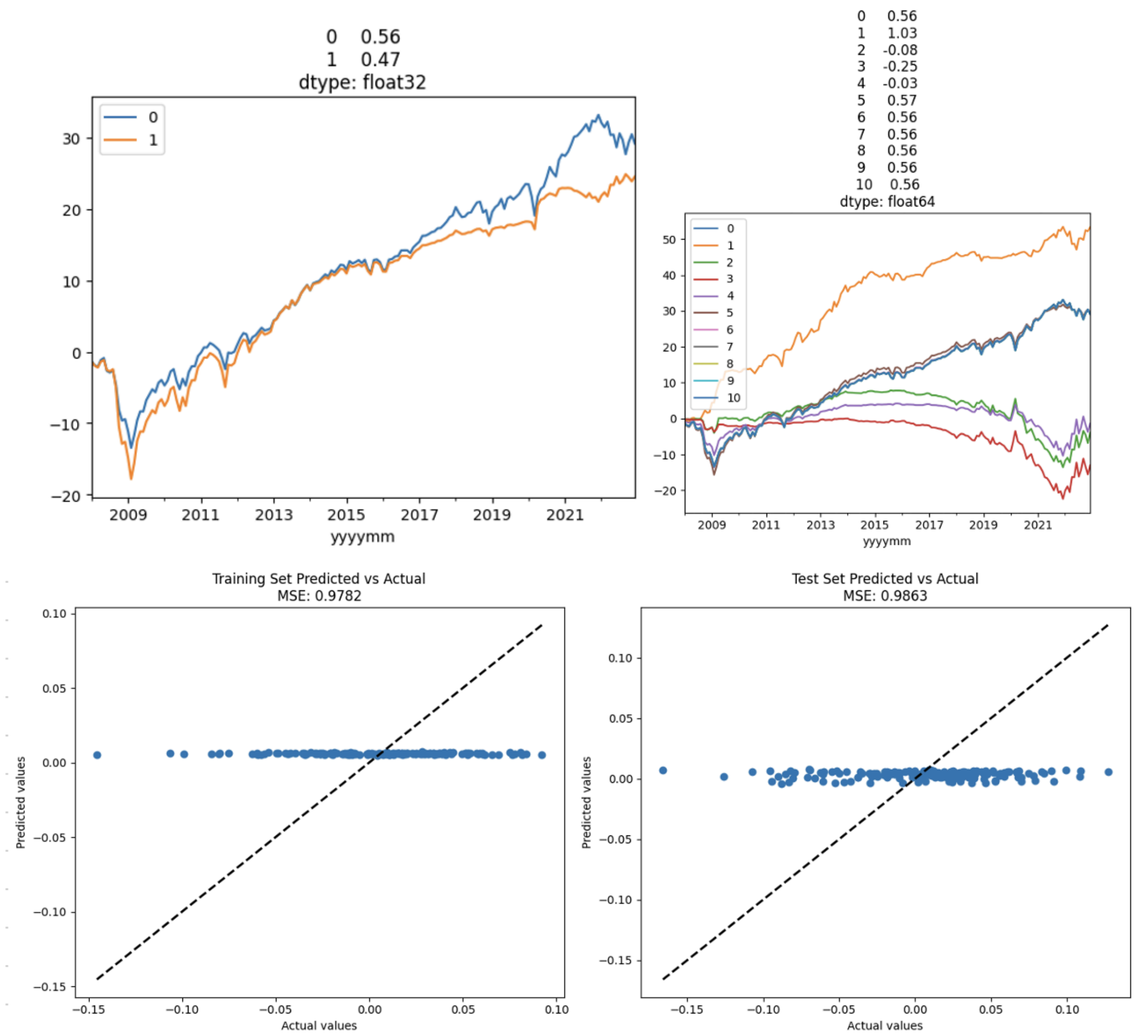


Figure 18: Model 2 results using averaging predictions and ridge regression

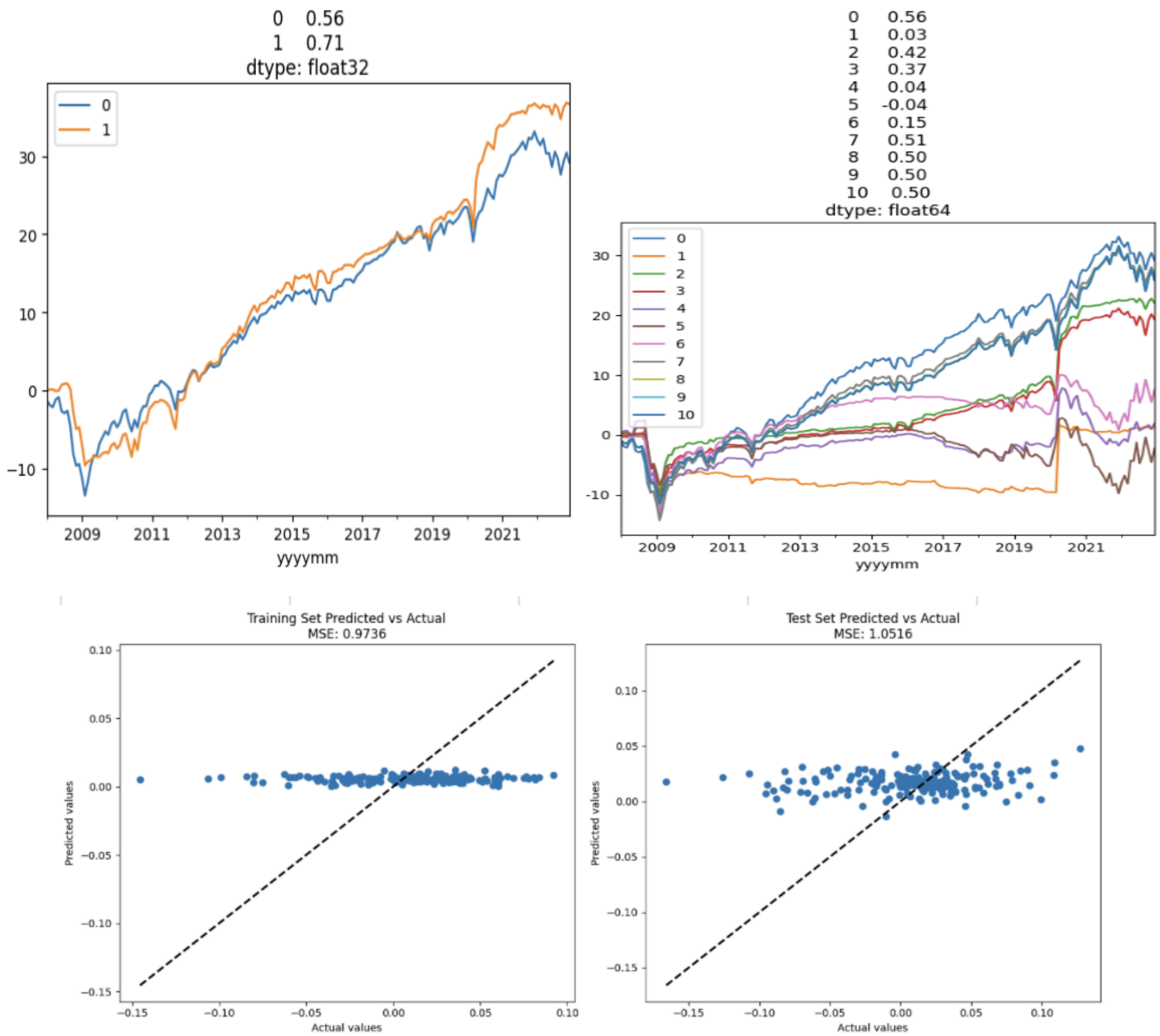


Figure 19: Model 3 results using averaging predictions and ridge regression

The top-left graph represents the averaging method with data normalization, where label 0 corresponds to the actual Sharpe ratio and label 1 to the model's Sharpe ratio.

The top-right graph illustrates SR from ridge regression results, with a shrinkage list from 10^{-8} to 1000.

The bottom dotted graph represents the predictions using the averaging method.

In **Model 2**, the dotted graph visualizes the **averaging prediction** results, while the best **ridge regression** result of 1.03 SR is obtained using a shrinkage value of $\mathbf{10^{-8}}$, outperforming the actual market performance.

Since **Model 2** achieved the highest Sharpe ratio using **ridge regression**, further parameter variations were conducted on:

- Epochs: (100, 200, 250, 300, 400)
- Learning Rates: (0.01, 0.03, 0.08, 0.1, 0.15)
- Ridge Penalties: (0.0005, 0.0025, 0.005, 0.006, 0.0075)

A detailed parameter variation trial is provided in the attachment section at the end of this document.

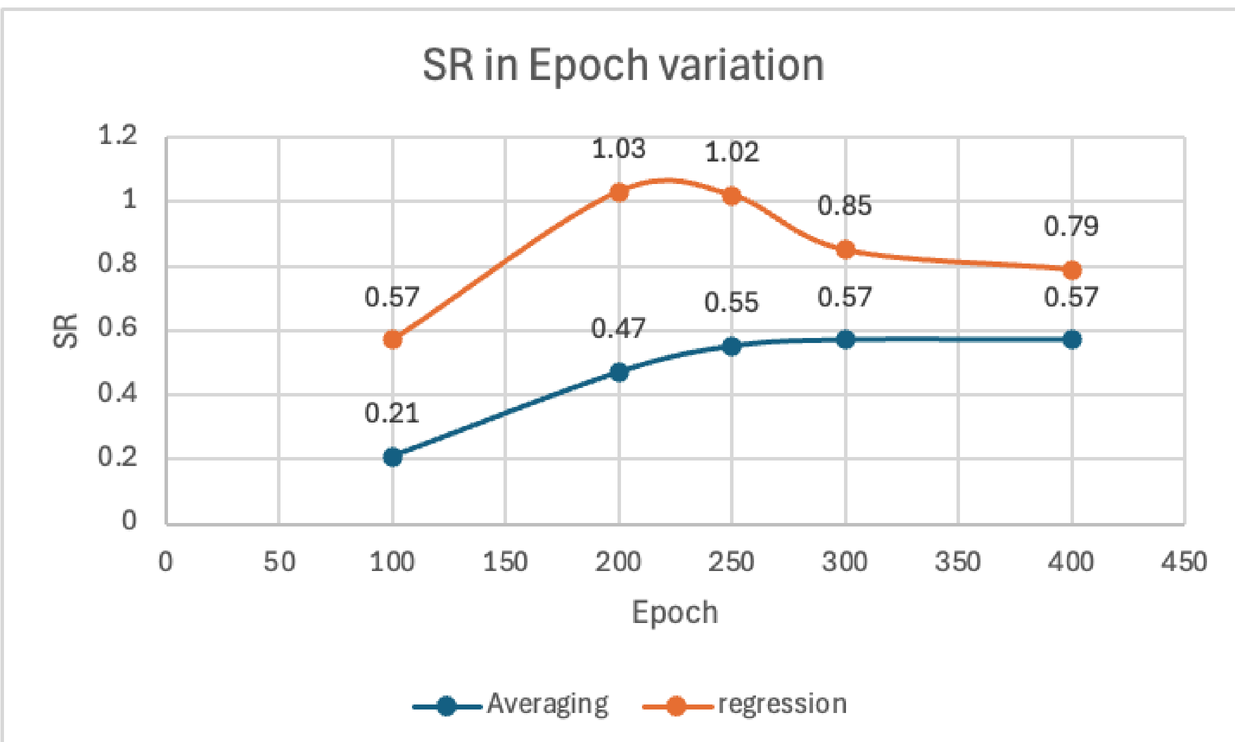


Figure 20: Epoch Variation

Optimal Epoch Range: The **best performance** for **ridge regression** is observed at around **200 epochs**, where the Sharpe Ratio reaches its peak **(1.03)**. After that, performance deteriorates. For the averaging method, the performance increases until **300** epochs to **0.57** Sharpe Ratio then it remains stable.

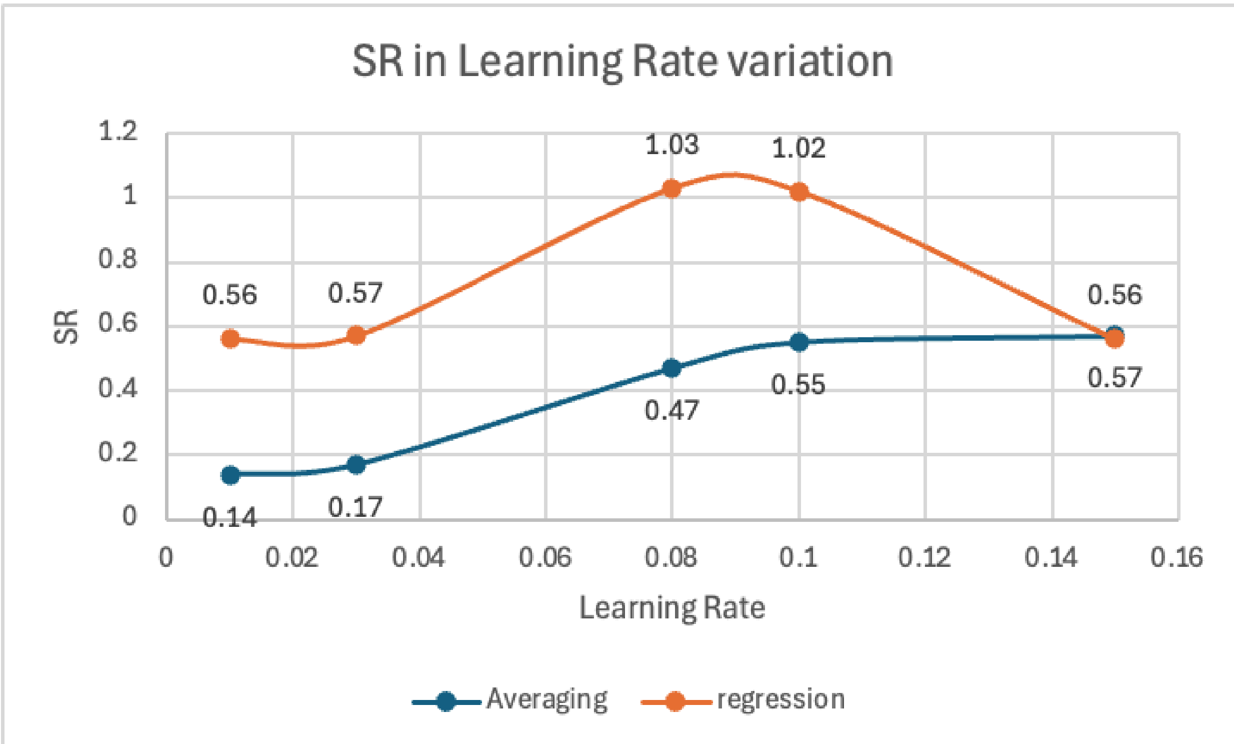


Figure 21: Learning Rate Variation

Optimal learning rate range: The best performance for **ridge regression** is observed at around **0.08 learning rate**, where the Sharpe Ratio reaches its peak (**1.03**). After that, performance **deteriorates**. For the averaging method, the performance increases until a learning rate of **0.1** for a **0.55** Sharpe Ratio and remains constant around **0.57**.

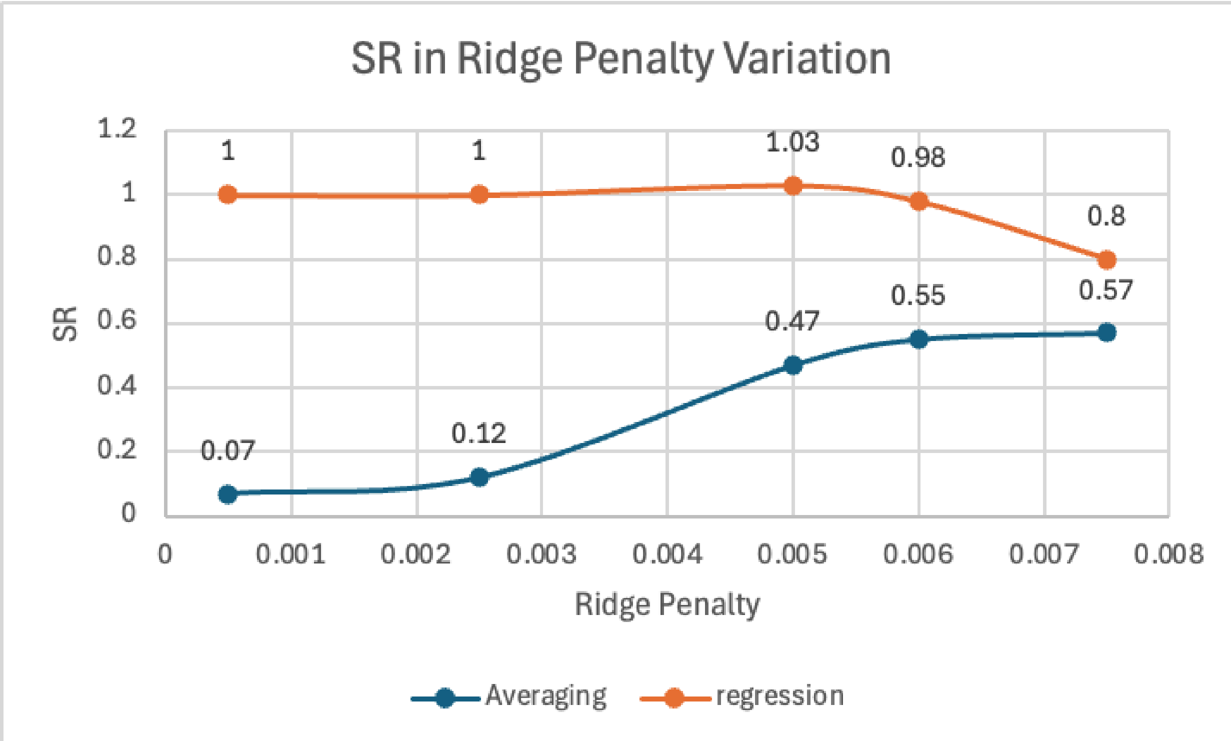


Figure 22: Ridge Penalty Variation

Optimal Ridge Penalty: in ridge regression method, the **Sharpe ratio** (SR) remains relatively constant between **1** and **1.03** for **ridge penalty** values ranging from **0.0005** to **0.005**. Beyond this range, **increasing** the **ridge penalty** leads to a **decline** in the **Sharpe Ratio**.

In contrast, for the **averaging method**, performance improves as the **ridge penalty increases**, reaching an **SR** of **0.55** at a **penalty** of **0.006**. The **SR** then rises slightly to **0.57** and **stabilizes** at that level.

5.1.2 GLD Result Analysis

Optimal results were obtained under different parameter settings—one leveraging **averaging predictions** and the other utilizing the last hidden layer for **ridge regression** (with a shrinkage range from 10^{-8} to 1000) as shown in the below table.

Category	Model 1	Model 2
Number of seed	10	10
Width	230	200
Ridge Penalty	0.001	0.035
Learning Rate	0.017	0.01
Epoch	200	250
Actual Sharpe ratio	0.28	
Averaging Predictions	0.36	0.50
Ridge Regression	0.35	0.29

The first model exhibited **balanced performance**, achieving a **Sharpe Ratio** (SR) of **0.36** for averaging predictions and 0.35 for ridge regression. This suggests that **regularisation** helped **stabilize** the predictions while **preserving meaningful signal relationships**. In contrast, the **second model** performed better with **averaging,** reaching an **SR** of **0.50**, while **ridge regression** lagged at **0.29** at 0.1 shrinkage list. These results indicate that **averaging techniques** can **mitigate overfitting** and enhance **predictive performance**. The **higher Sharpe Ratios** suggest that combining different predictions provides a more robust signal extraction. Furthermore, the **underperformance** of **ridge regression** in **Model 2** might imply that the **selected shrinkage level** constrained certain variations in the

gold market's risk-return tradeoff. Notably, both models outperformed the market's risk-adjusted return, which had an **SR** of **0.28**.

Figure 23: GLD Model 1 results for averaging prediction and ridge regression

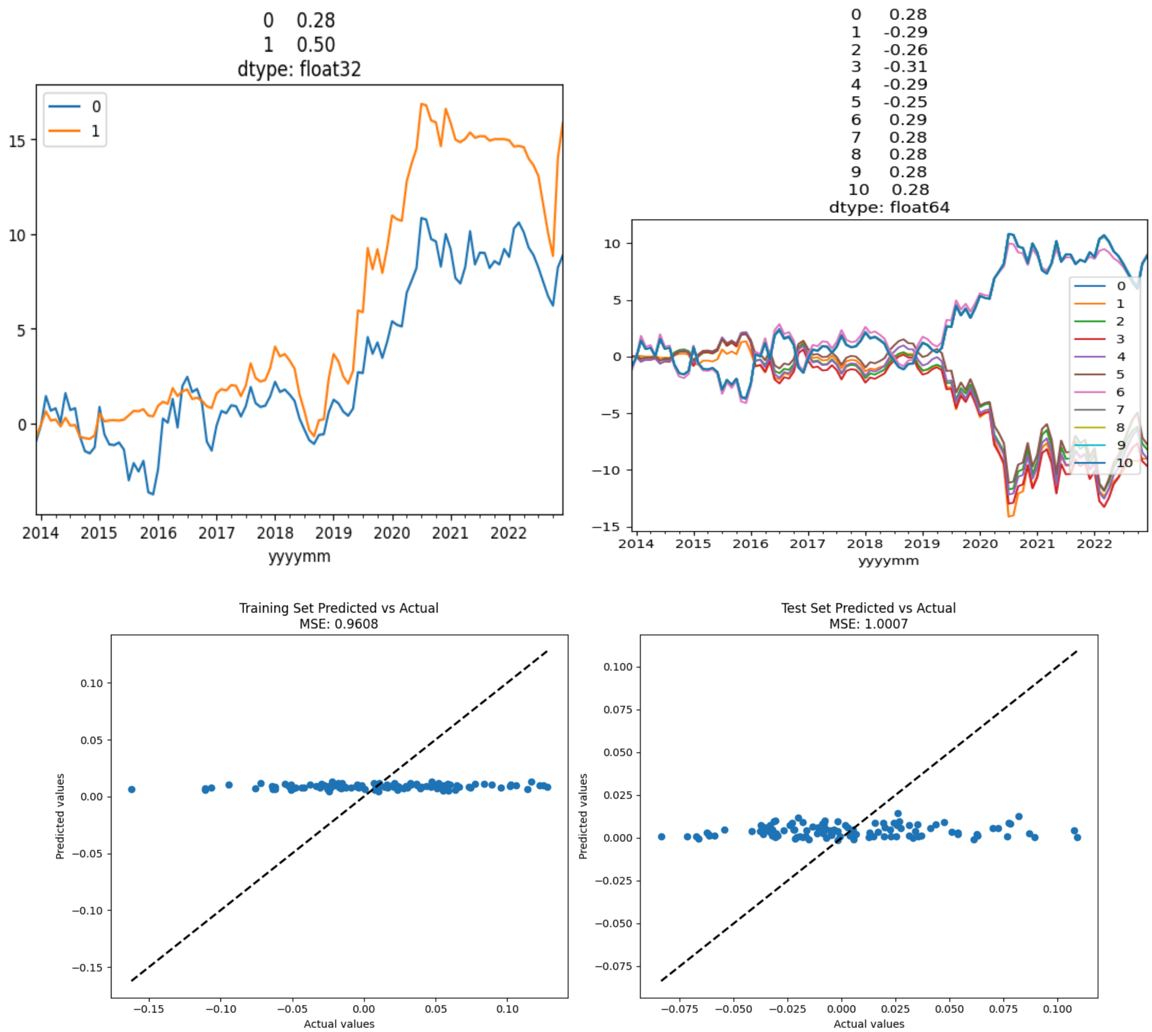


Figure 24: GLD Model 2 results for averaging prediction and ridge regression

The top-left graph represents the averaging method, where 0 corresponds to the actual Sharpe ratio and 1 represents the model's Sharpe ratio.

The top-right graph displays SR from ridge regression results, with shrinkage values ranging from 10^{-8} to 1000.

The bottom dotted graph represents the predictions using the averaging method.

Since **Model 2** achieved the **highest Sharpe ratio** using **averaging predictions**, additional variations were explored across different parameters, including:

- Epochs: (100, 150, 250, 300, 350)
- Learning Rates: (0.001, 0.005, 0.01, 0.02, 0.025)
- Ridge Penalties: (0.01, 0.02, 0.035, 0.04, 0.05)

A graphical comparison of these parameter variations is provided below, with detailed individual data points included in the report attachment section.

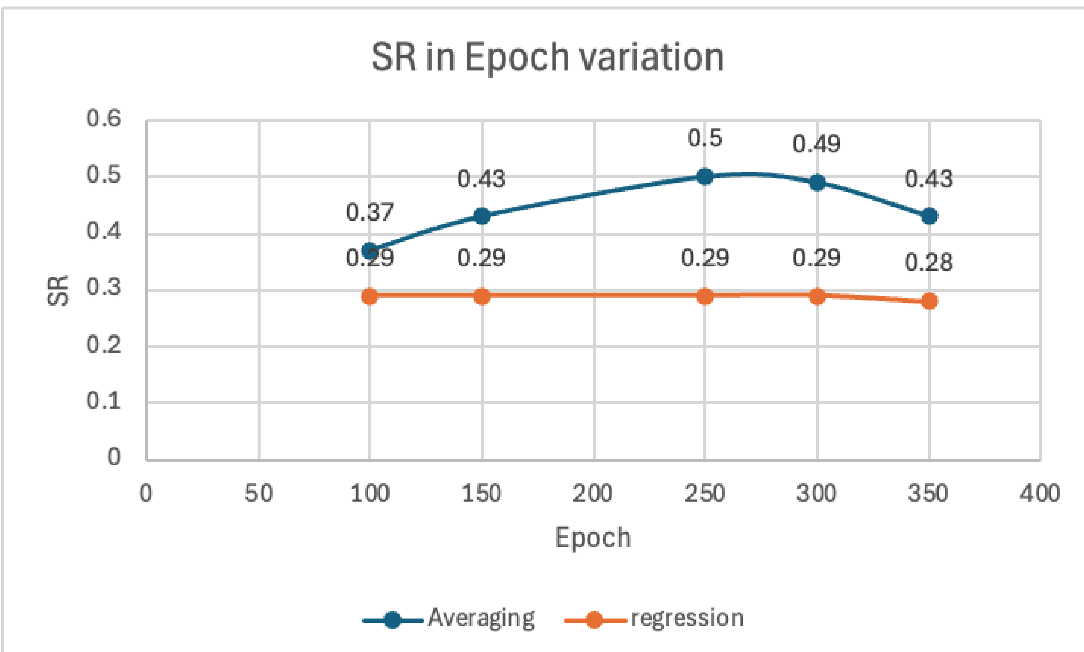


Figure 25: Epoch Variation

Optimal Epoch Range: The **best performance** for **averaging** is observed at around **250 epochs**, where SR reaches its peak **(0.5)**. After that, **performance deteriorates**. While for the **regression** method, the performance quite **stable** across different parameter

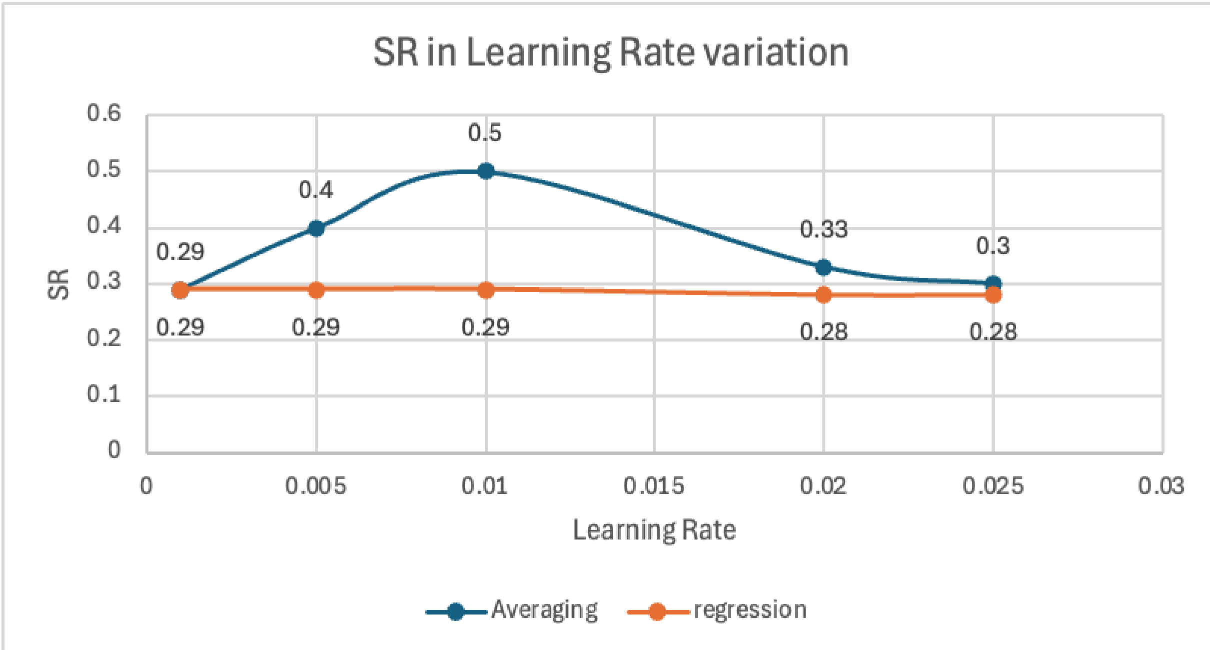


Figure 26: Learning Rate Variation

Optimal Learning Rate Range: The best performance for **averaging predictions** is observed at around **0.01**, where the **Sharpe Ratio** reaches its **peak (0.5)**. After that, **performance deteriorates**. While for the regression method, the performance is quite stable across different parameters.

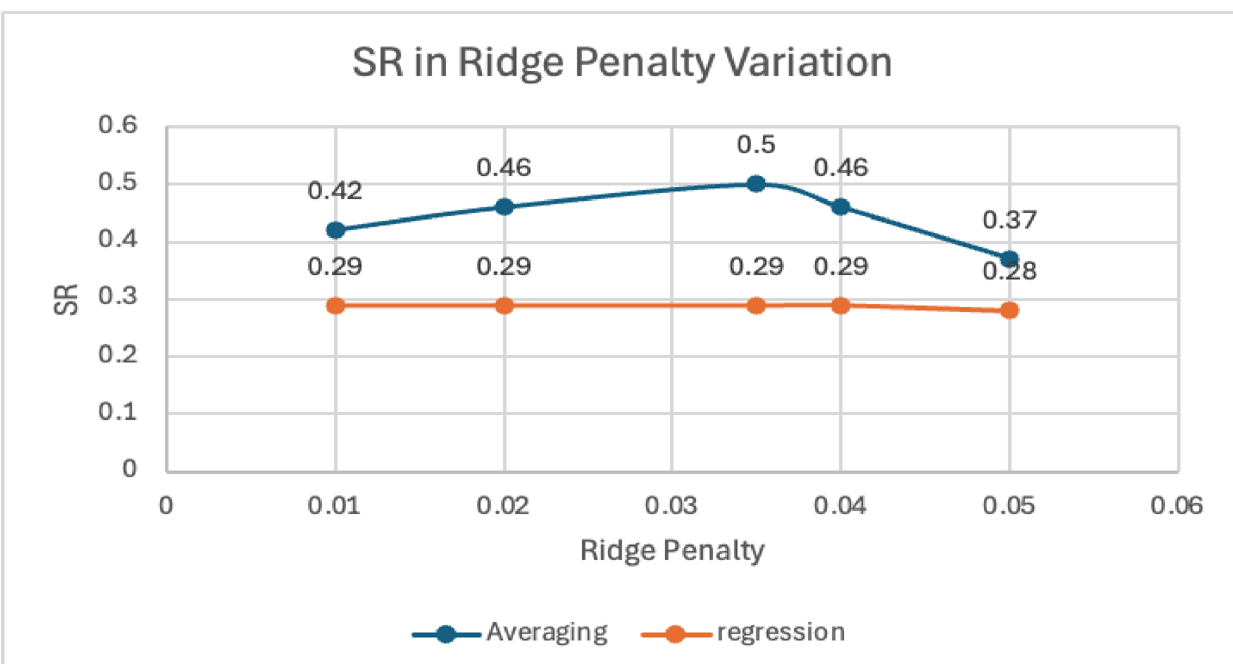


Figure 27: Ridge Penalty Variation

Optimal ridge penalty variation: The best performance for averaging predictions is observed at around **0.035 ridge penalty**, where the **Sharpe Ratio** reaches its **peak (0.5).** After that, performance deteriorates. While for the **regression method**, the performance is **relatively stable** across different parameters.

5.1.3 TLT Result Analysis

In the analysis of the **TLT ticker**, the optimal results were achieved with ridge regression. The performance of the **averaging method** remained **relatively constant** across different parameter ranges, as illustrated in the variation plots below.

Category	Model
Number of seed	45
Width	200
Ridge Penalty	0.00001
Learning Rate	0.12
Epoch	400
Actual Sharpe ratio	0.03
Averaging Predictions	-0.15
Ridge Regression	0.54

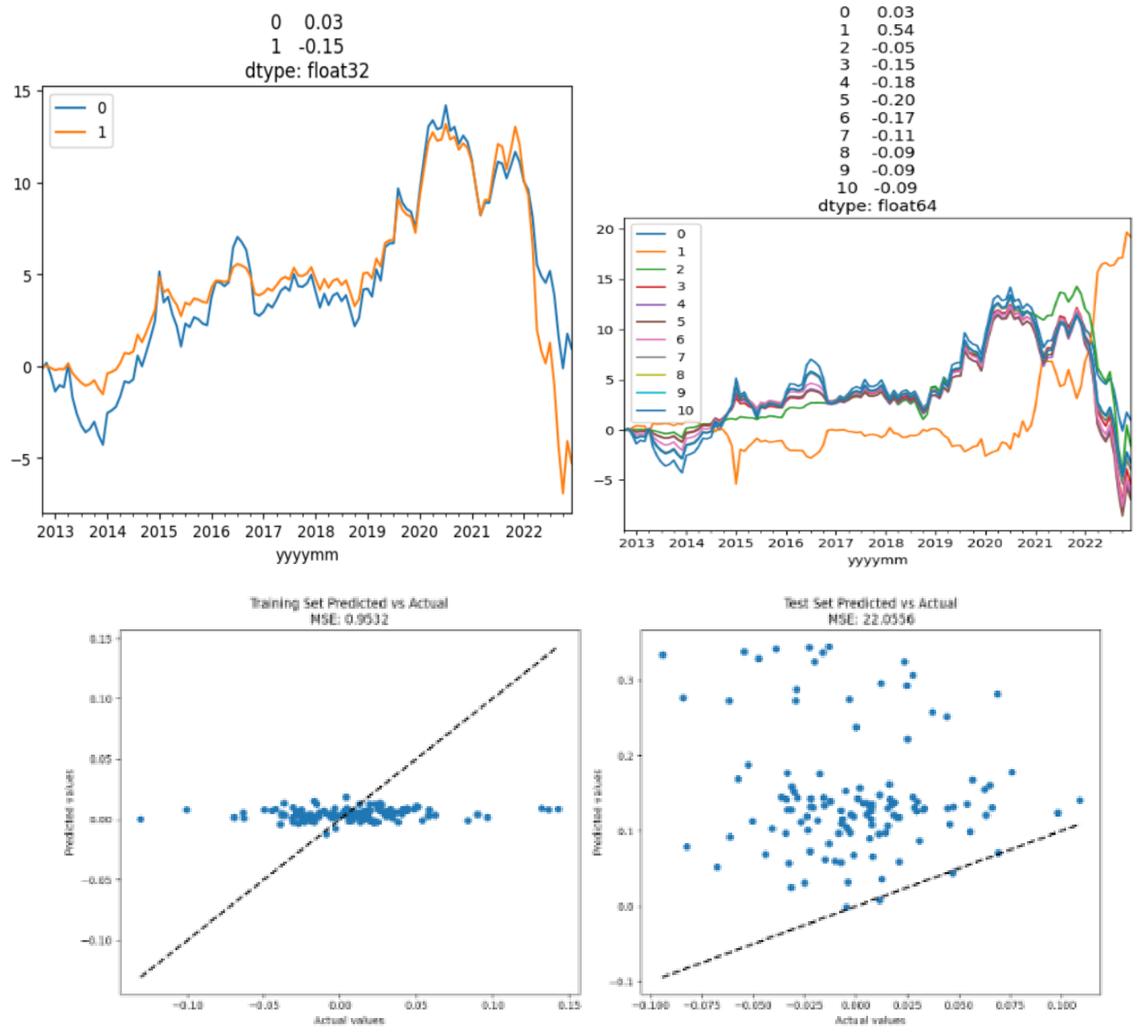


Figure 28: TLT Model results for averaging predictions and Ridge Regression

The top-left graph represents the averaging method with data normalization, where 0 corresponds to the actual Sharpe ratio and 1 to the model's Sharpe ratio.

The top-right graph illustrates ridge regression results ranging from 10^{-8} to 1000 (shrinkage list) The highest Sharpe ratio (0.54) for ridge regression is observed at the yellow line which shrinkage list performance at 10^{-8}

The bottom dotted graph visualizes the predictions using the averaging method.

To explore parameter sensitivity, variations were conducted in:

- Epochs: (200, 300, 400, 500, 800)
- Learning Rates: (0.03, 0.06, 0.12, 0.24, 0.48)
- Ridge Penalties: (0.0000025, 0.000005, 0.00001, 0.00005, 0.0001)

The detailed model result plots for all variations can be found in the attachment at the end of this report.

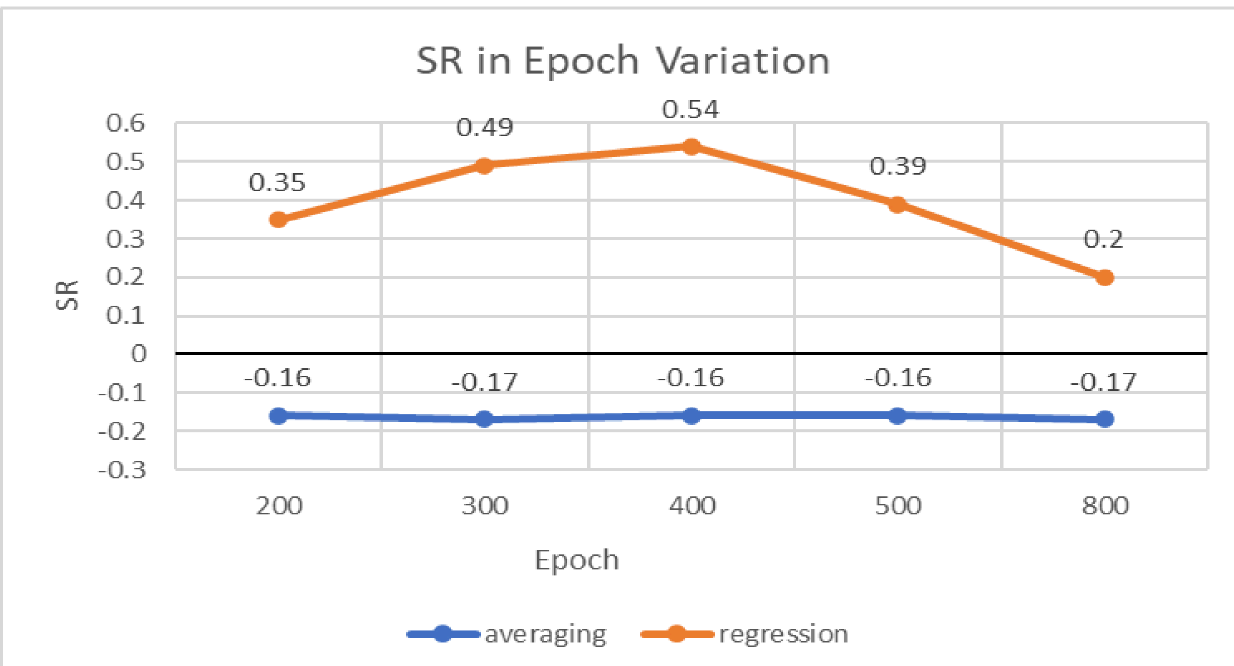


Figure 29: Epoch Variation

This graph depicts the variation in the **Sharpe Ratio** (SR) across different epochs for both Averaging and Ridge Regression methods.

- **Ridge Regression Performance**: **Ridge regression** consistently **outperforms simple averaging**, yielding a significantly **higher SR** and demonstrating superior risk-adjusted returns.

- **Optimal Epoch Range**: The best performance for **ridge regression** occurs around **400 epochs**, where the SR peaks at **0.54**. Beyond this point, performance declines.
- **Overfitting Trend**: The SR deterioration after **400 epochs** suggests potential **overfitting**, where the model begins capturing noise rather than meaningful patterns.
- **Averaging Ineffectiveness**: The persistently **negative SR** for **averaging** indicates that it is **not a viable** method for generating profitable predictions in this context.

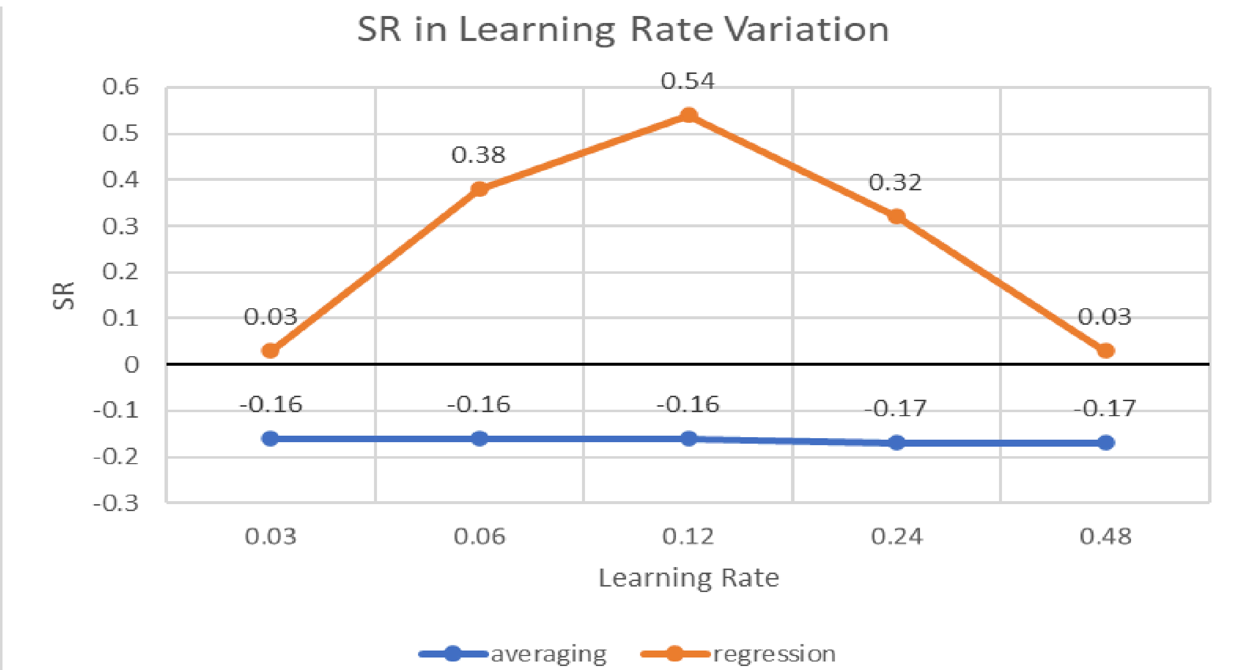


Figure 30: Learning Rate Variation

This graph illustrates the variation in the **Sharpe Ratio** (SR) across different learning rates for **Averaging** and **Ridge Regression** methods. The **highest SR** of **0.54** is achieved at a learning rate of **0.12**, suggesting an **optimal balance** between **learning speed** and **model generalization** for **ridge regression.** In contrast, the **Averaging method** consistently produces **negative** SR values across all **learning rates**, indicating that **simple averaging** is not an effective prediction method in this scenario.

Effect of Learning Rate on Performance:

- **Too Low (0.03)**: The model learns **too slowly**, leading to suboptimal SR.
- **Optimal (0.12)**: **Best performance**, suggesting this is the ideal learning rate.
- **Too High (0.24, 0.48)**: SR declines, possibly due to the model failing to **converge** or **learning inefficiently**.

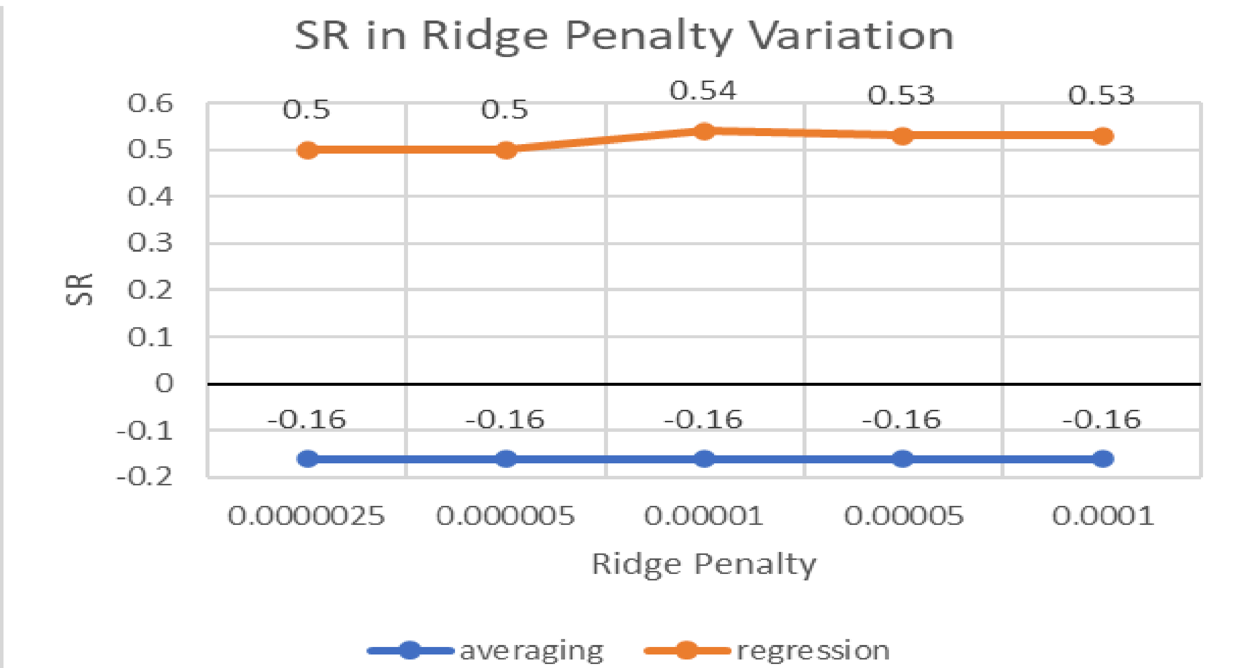


Figure 31: Ridge Penalty Variation

This graph illustrates the variation of the Shape Ratio (SR) across different ridge penalties for Averaging and Ridge Regression.

Ridge regression results maintain a **high SR (~0.5)** across all Ridge Penalty values. The slight peak at **0.00001 (SR = 0.54)** suggests an optimal penalty value for performance improvement. The SR remains consistently at **-0.16**, showing no impact from **Ridge Penalty** changes for average results.

Regularization Impact:

- **Too Low (0.00000025 - 0.000005):** The SR is stable but slightly lower.
- **Optimal (0.00001)**: The **highest SR** of **0.54** is achieved.
- **Too High (0.00005 - 0.0001)**: The SR remains **stable** but slightly declines.

6. Conclusion & Limitations

In this report, we regressed the selected **ticker** returns against **Goyal-Welch signals** using various **machine-learning techniques**. We experimented with different **ridge penalties**, **normalization methods**, **seed numbers**, **epochs**, and **learning rates** to optimize the **Sharpe Ratio**. These are key result comparisons between random features regression with gradient descent Sharpe ratio (SR)

	SPY	GLD	TLT
Actual Market SR	0.56	0.28	0.03
OLS with Random Features Model SR	0.73	0.52	0.11
Deep Learning Model SR	1.03	0.5	0.54

From the results, We can infer that the **deep learning model** mostly **outperforms** the **OLS model** with a **higher Sharpe ratio** except **GLD** which has quite a **similar** result. This indicates that **deep learning** models effectively capture **nonlinear relationships** in financial data, enhancing **risk-adjusted returns**. However, for **gold**, the absence of such **substantial improvement** may imply that **traditional methods** are already **sufficient** for modeling its returns.

While the analysis provides valuable insights, several **limitations** must be acknowledged. These constraints affect the scope, accuracy, and interpretability of the results. Addressing these issues in future research could enhance the **robustness** and **applicability** of the findings.

- **Computational Constraints:** The analysis was conducted within the **limits** of **available computational** resources, restricting the ability to **test larger** and **more complex** models.
- **Data Limitations:** The study is constrained by the availability of relevant **economic** and **financial** data. While the selected variables provide meaningful insights, additional macroeconomic indicators could improve predictive performance.
- **Collinearity Among Factors**: Many macroeconomic variables exhibit **high collinearity**, which can affect the **interpretability** and **stability** of the models. **Strong correlations** between predictive factors may lead to **redundancy** and **reduced** model effectiveness. **Advanced feature selection** techniques or **dimensionality reduction** methods could mitigate this problem,
- **Short Timeframe**: The dataset for **GLD** and **TLT** covers a relatively **short period**, limiting the ability to assess **long-term trends** and model stability across different **economic cycles.**
- **Economic Sensitivity**: The models remain **vulnerable** to **unpredictable macroeconomic** shocks such as financial crises

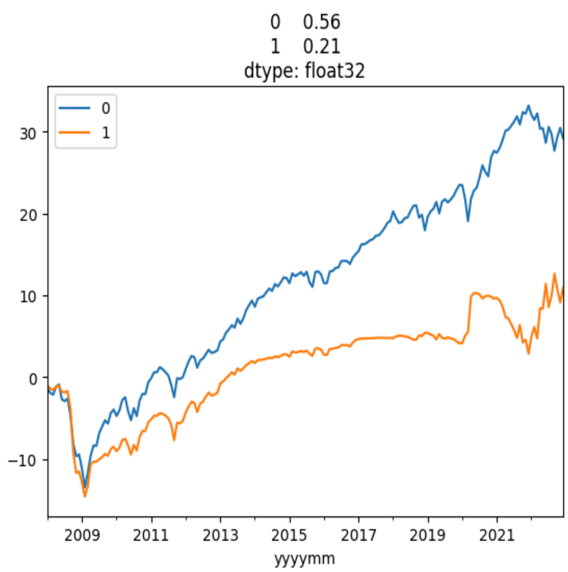
and **geopolitical events**, which could impact robustness in **real-world applications**.

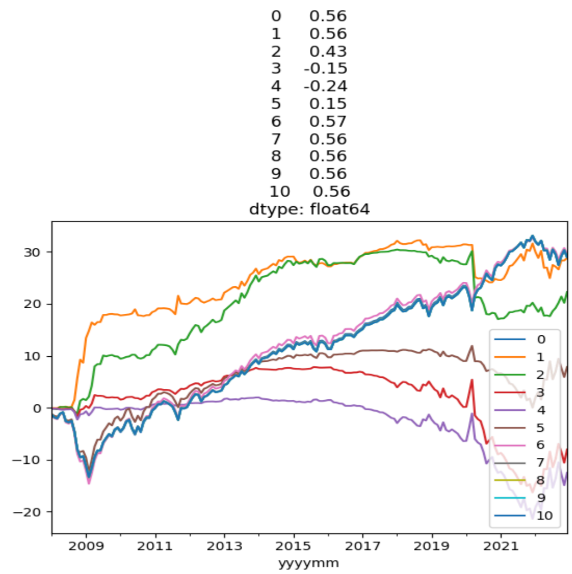
Overall, this report demonstrates the **effectiveness** of **deep learning techniques** in **financial modeling**, especially for equity-based instruments like **SPY**. The improvements in **Sharpe Ratios** show the potential of **leveraging machine learning** for **asset return predictions**. Nevertheless, challenges such as **data constraints**, **computational limitations**, and **economic uncertainties** must be addressed in the future to further **improve predictive accuracy** and allow **good applicability** of **models** in ever-changing financial marke

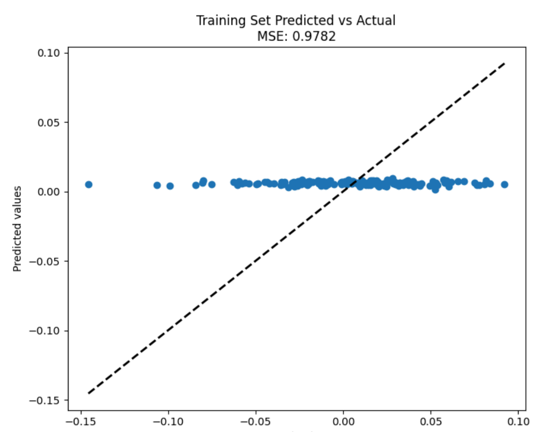
7. ANNEX

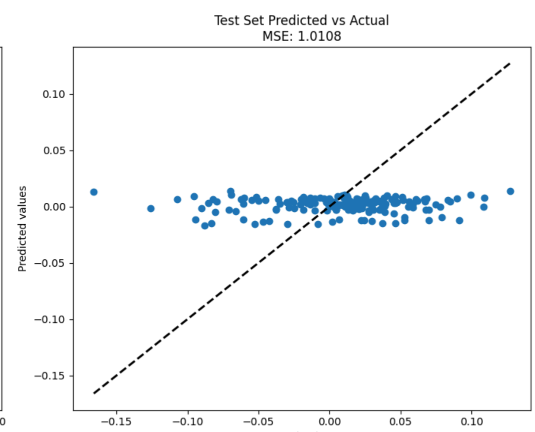
SPY Gradient Descent Epoch Variation

seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.08	100

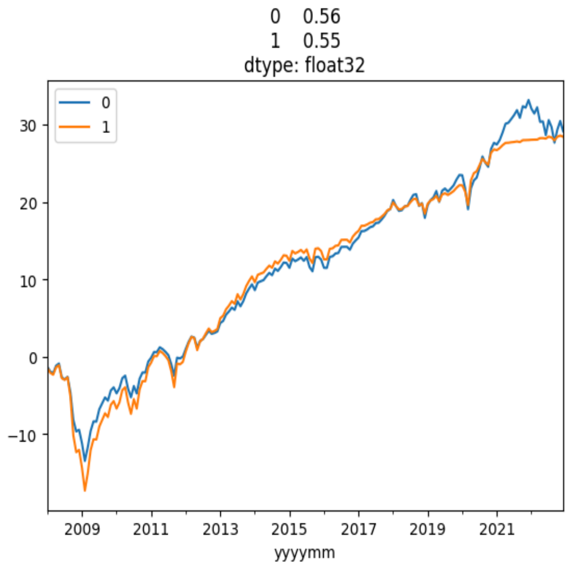


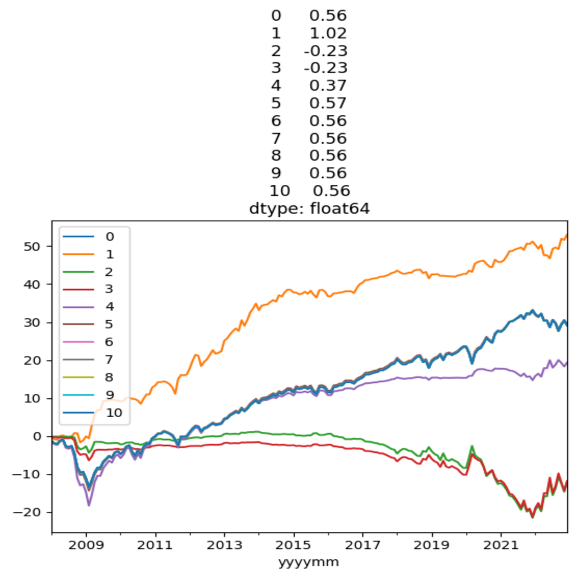


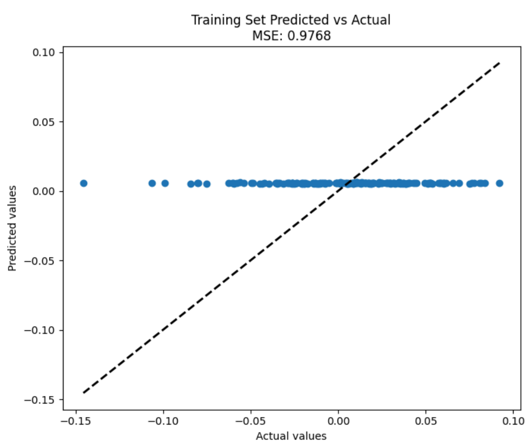


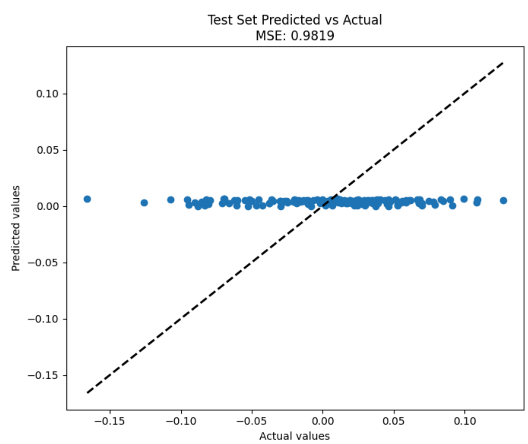


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.08	250

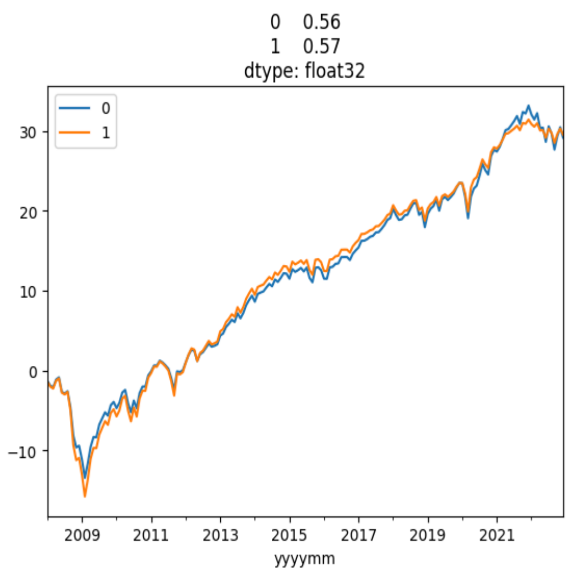


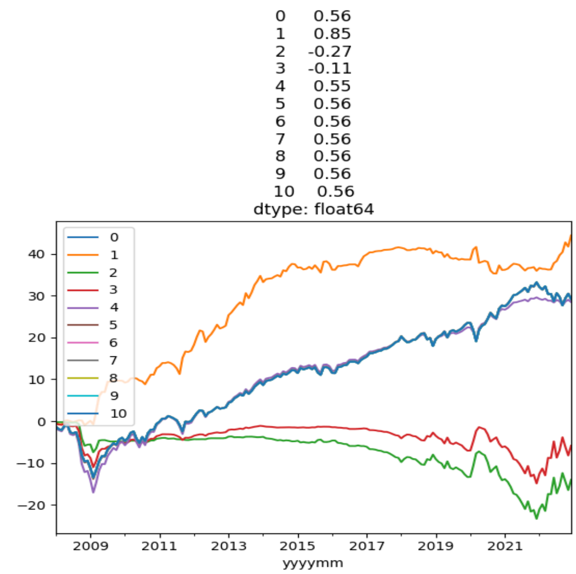


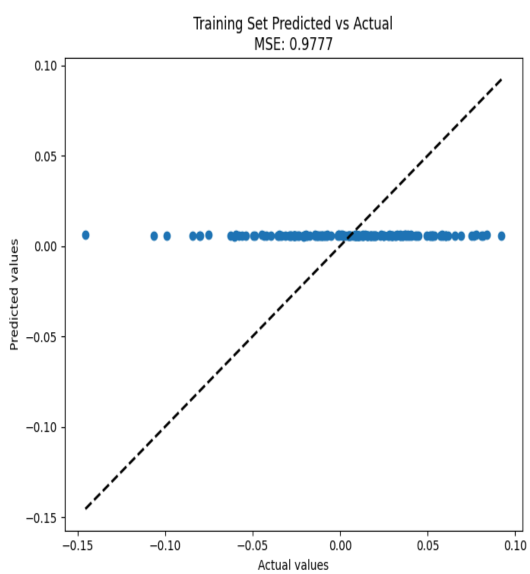


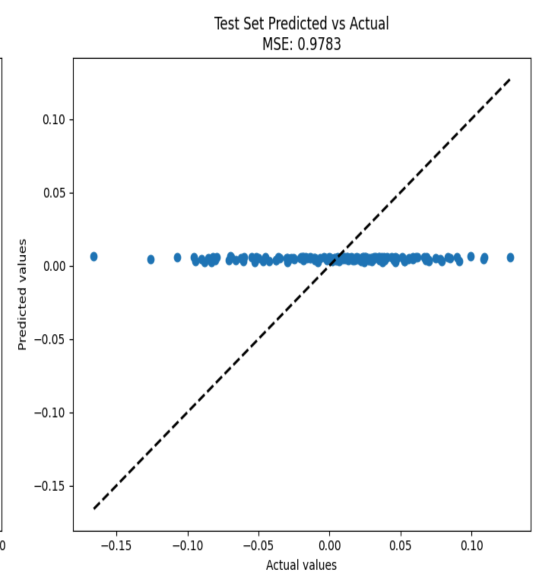


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.08	300

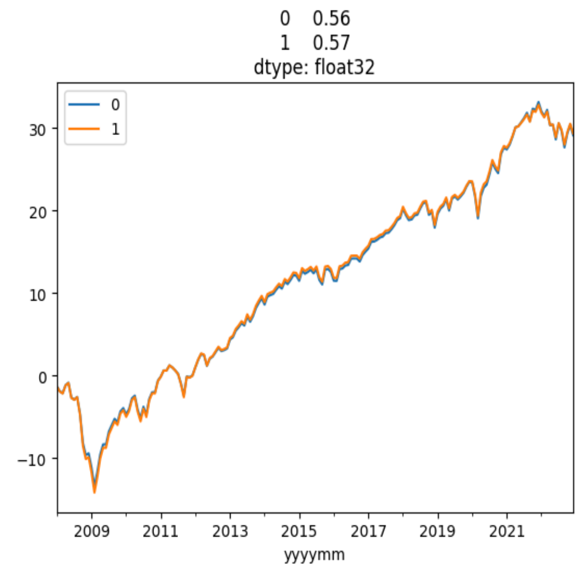


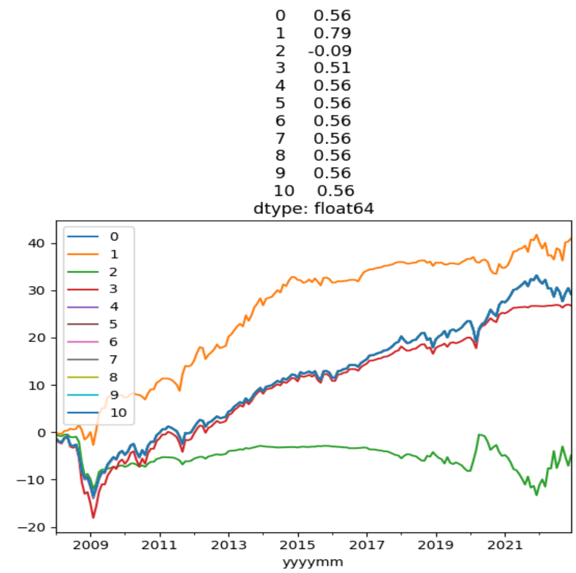


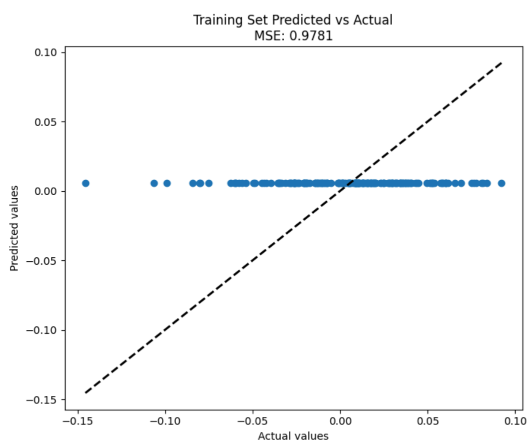


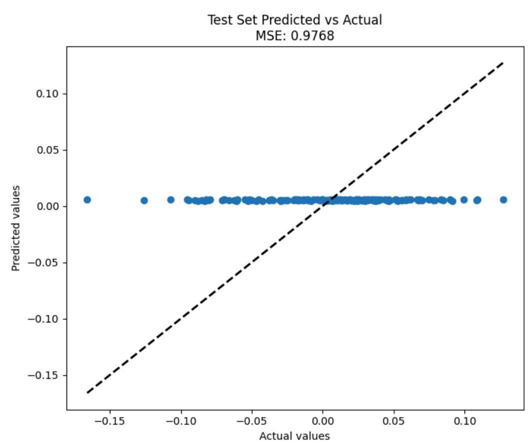


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.08	400



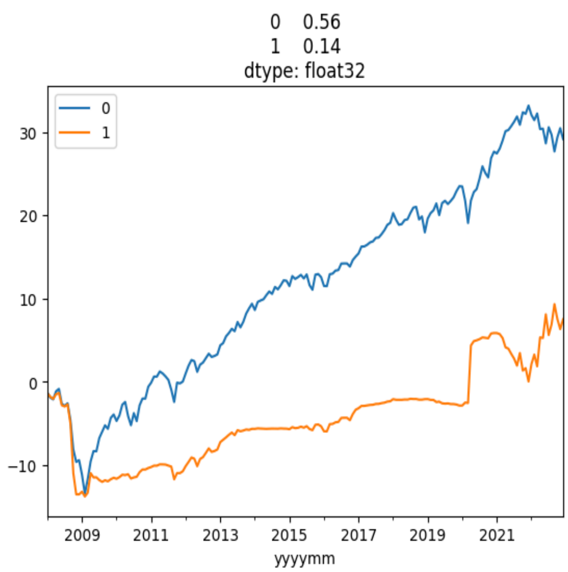


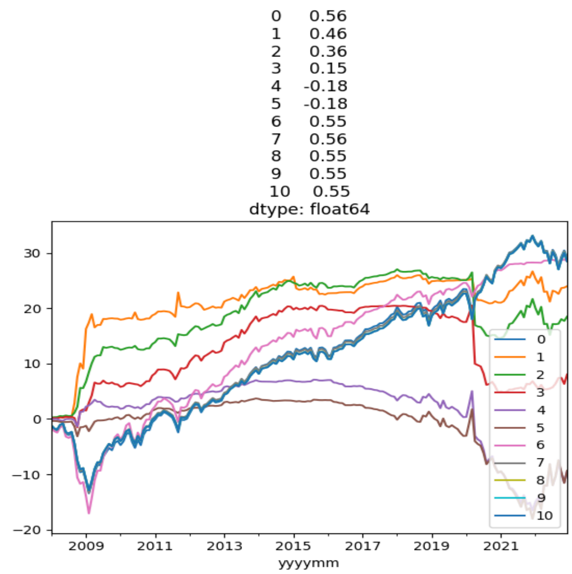


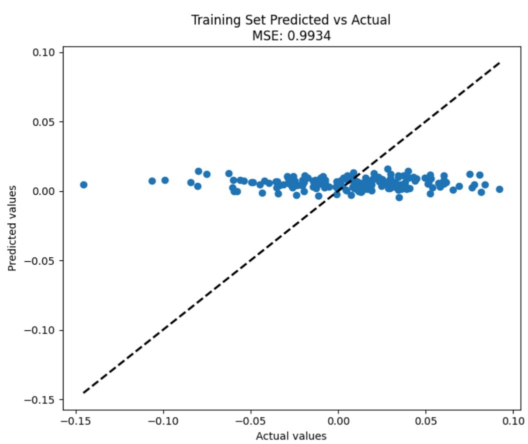


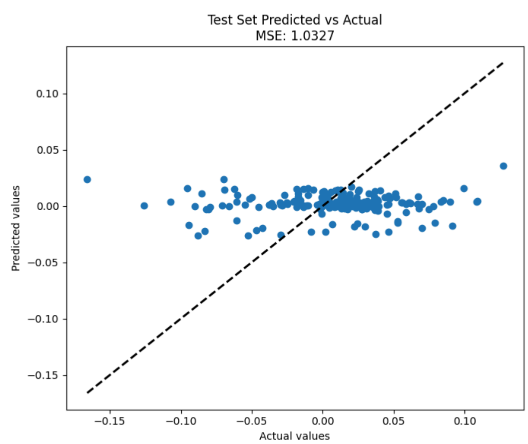
SPY Variation Learning Rate

seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.01	200

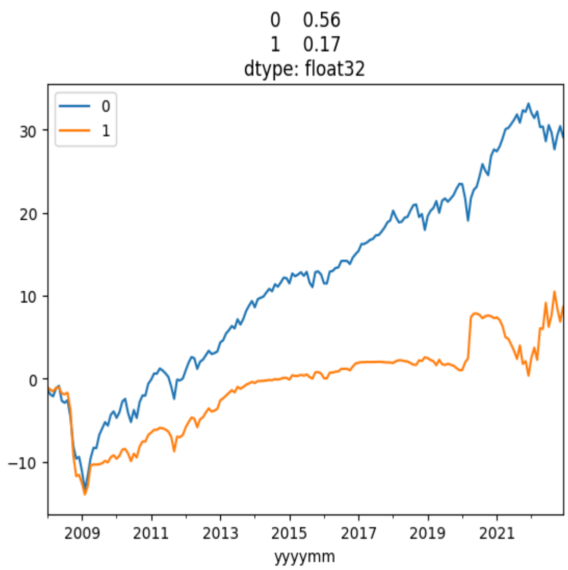


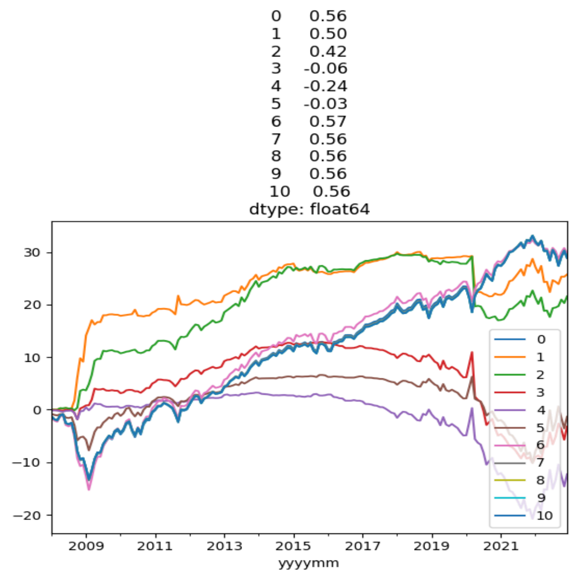


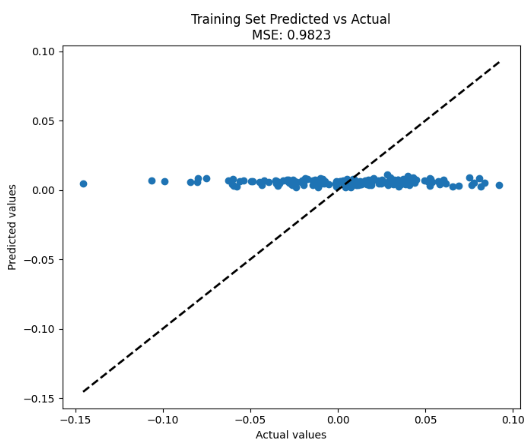


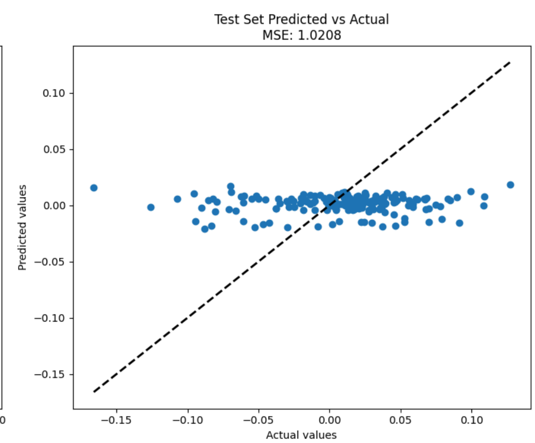


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.03	200

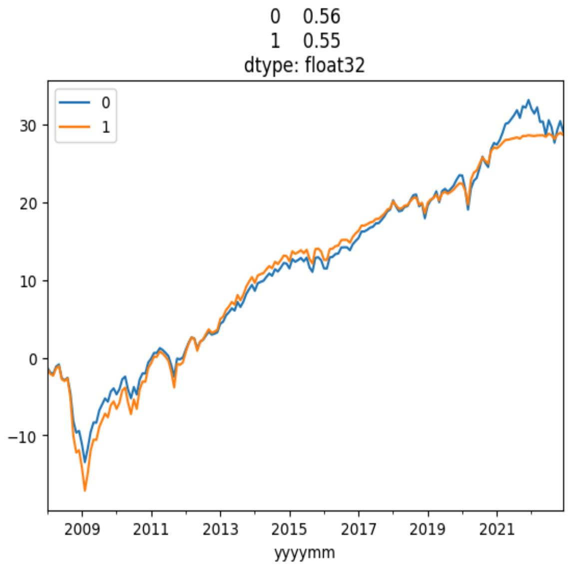


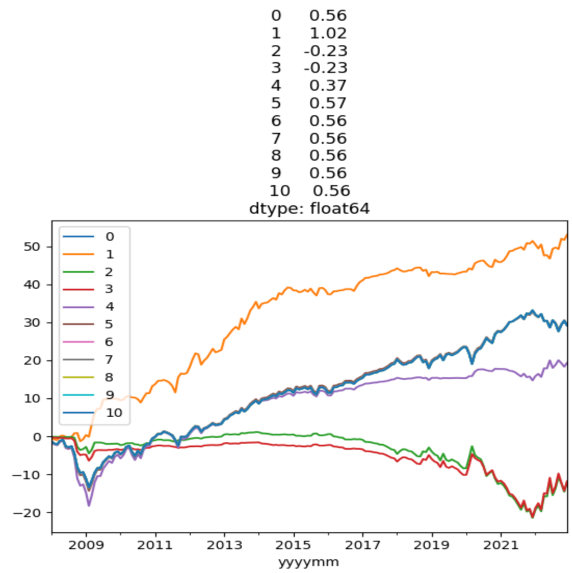


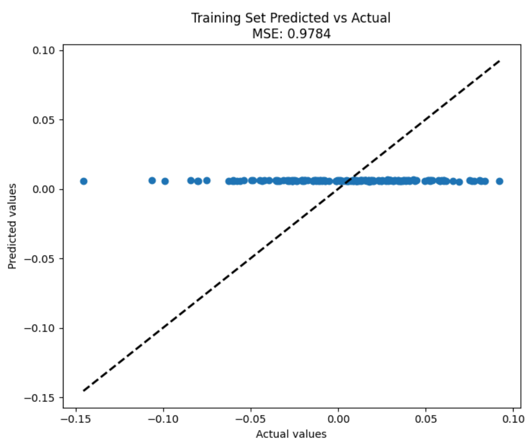


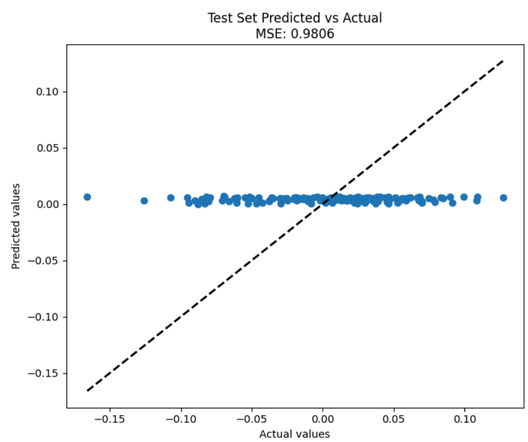


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.1	200

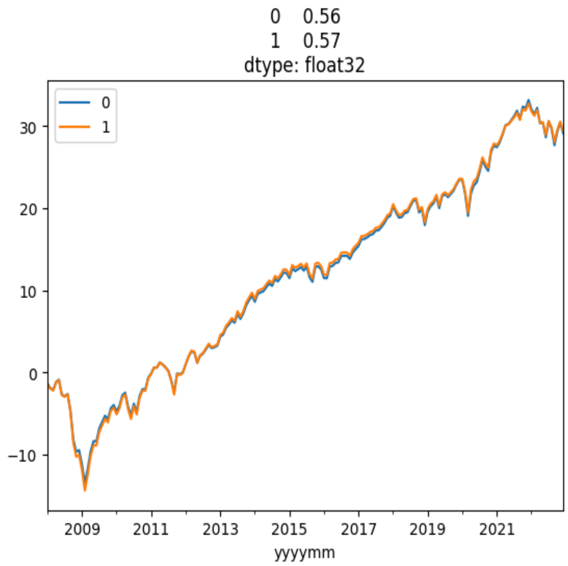


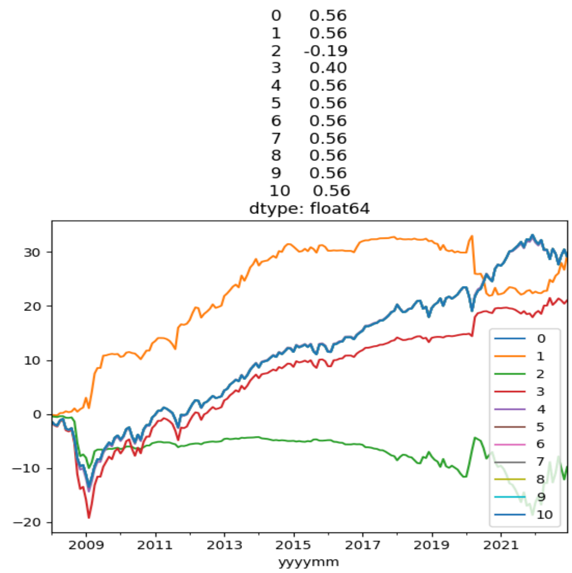


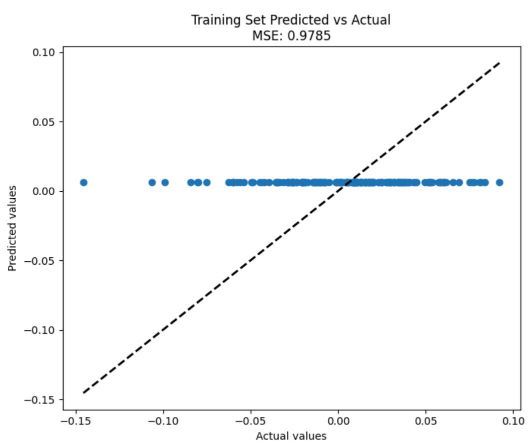


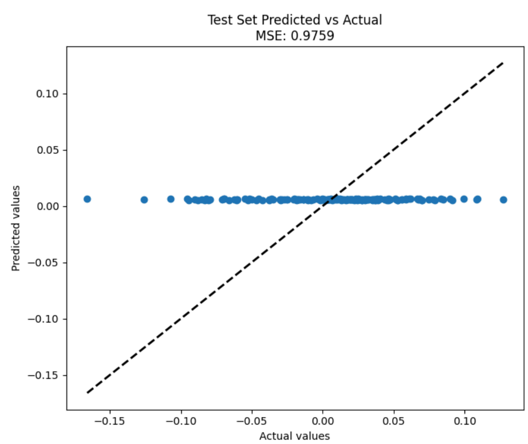


seed	width	ridge penalty	learning rate	epoch
25	40	0.005	0.15	200





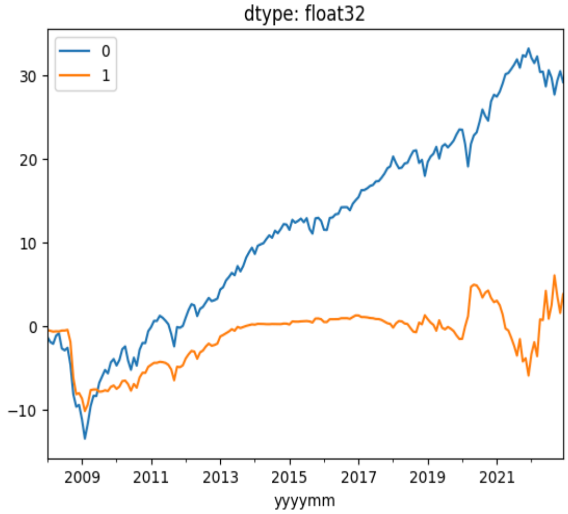




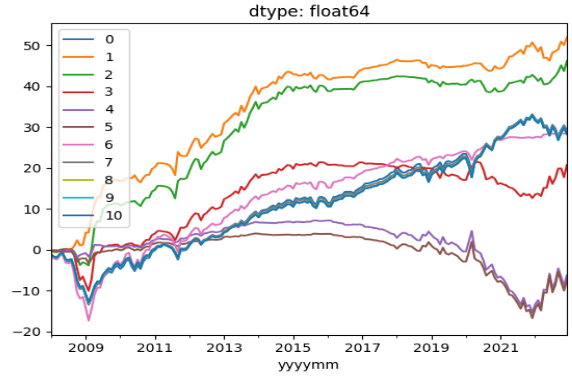
SPY Variation Ridge Penalty

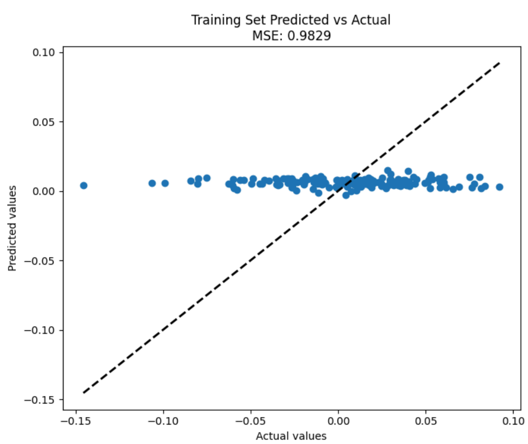
seed	width	ridge penalty	learning rate	epoch
25	40	0.0005	0.08	200

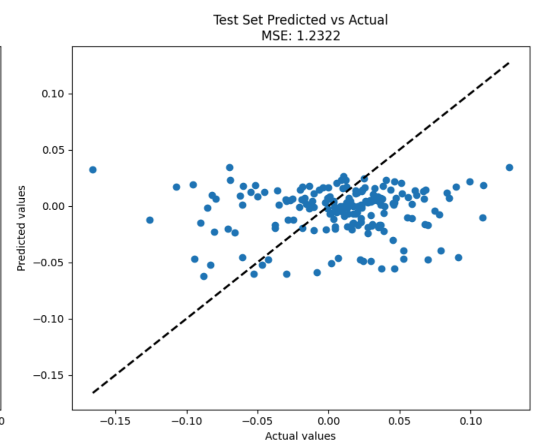
0   0.56
1   0.07
dtype: float32



0   0.56
1   1.00
2   0.89
3   0.40
4   -0.12
5   -0.15
6   0.55
7   0.55
8   0.55
9   0.55
10   0.55
dtype: float64

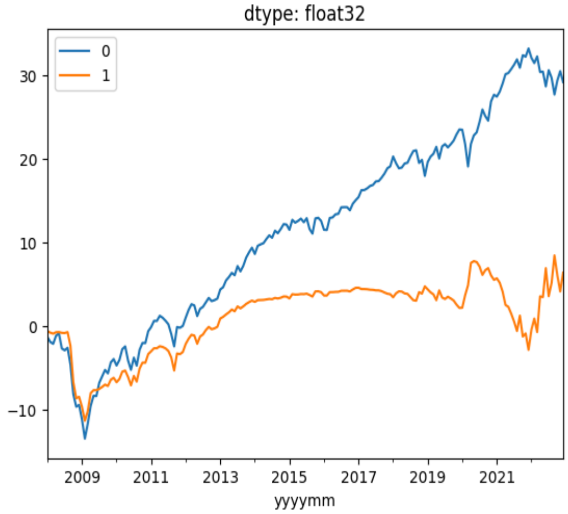




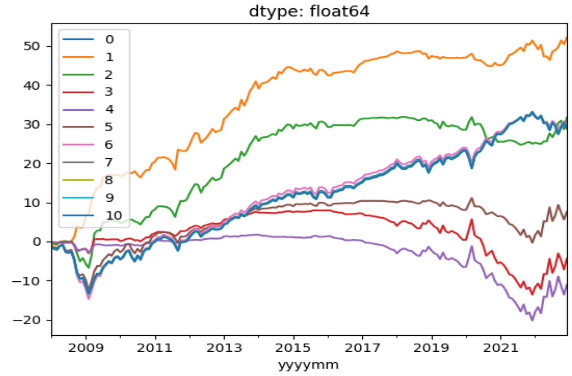


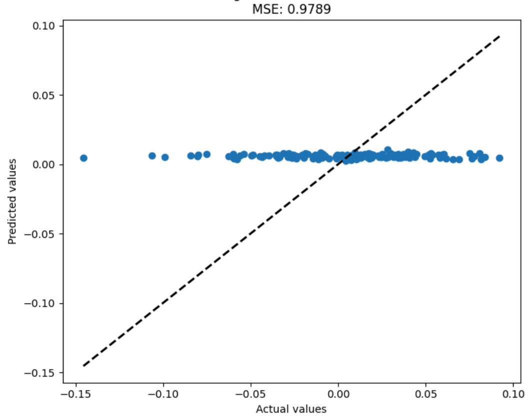
seed	width	ridge penalty	learning rate	epoch
25	40	0.0025	0.08	200

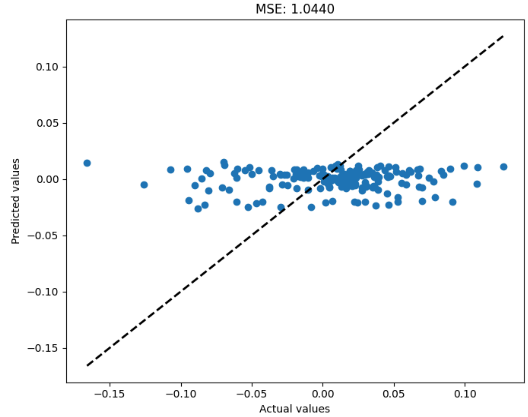
0 0.56
1 0.12
dtype: float32



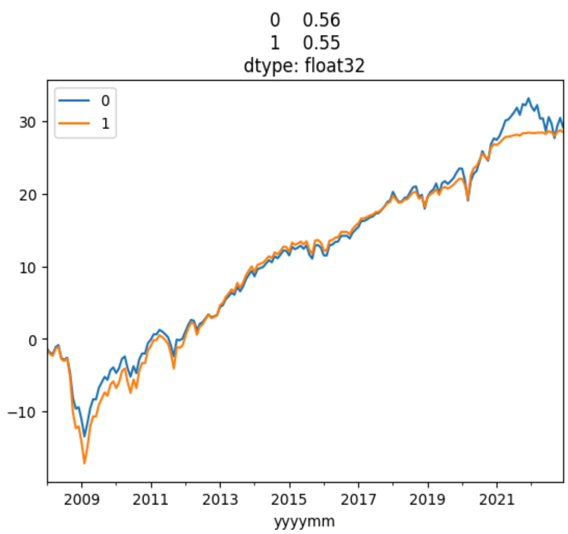
0 0.56
1 1.00
2 0.61
3 -0.08
4 -0.21
5 0.15
6 0.57
7 0.56
8 0.56
9 0.56
10 0.56
dtype: float64

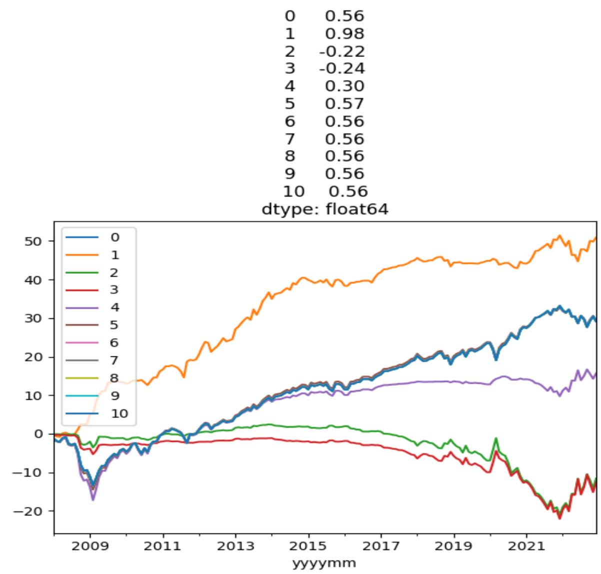


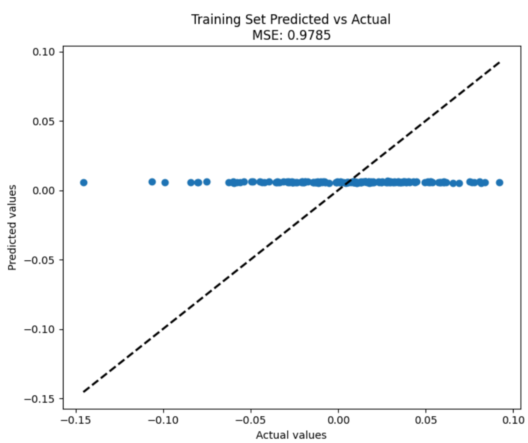


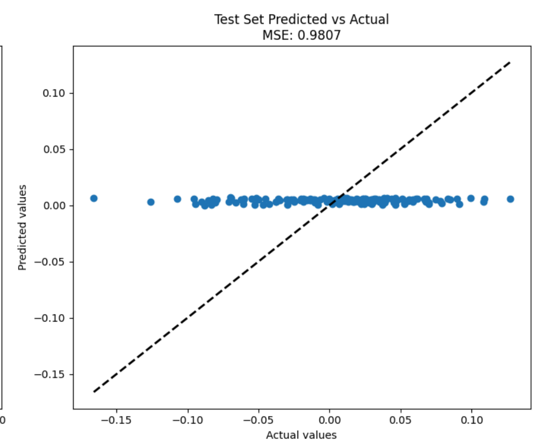


seed	width	ridge penalty	learning rate	epoch
25	40	0.006	0.08	200



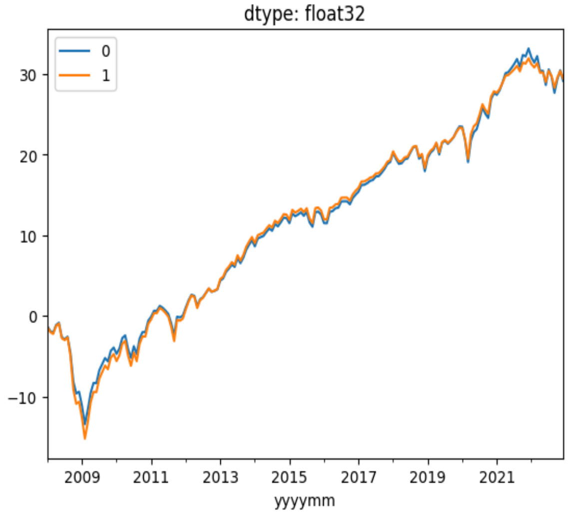




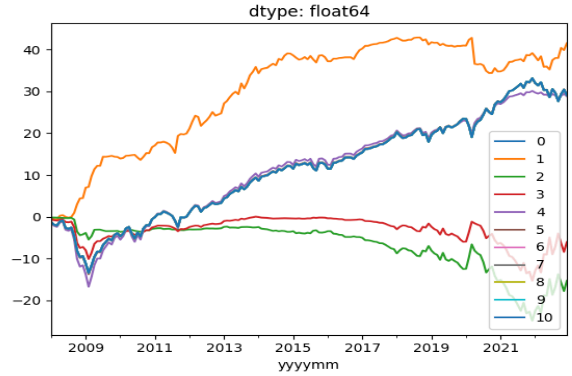


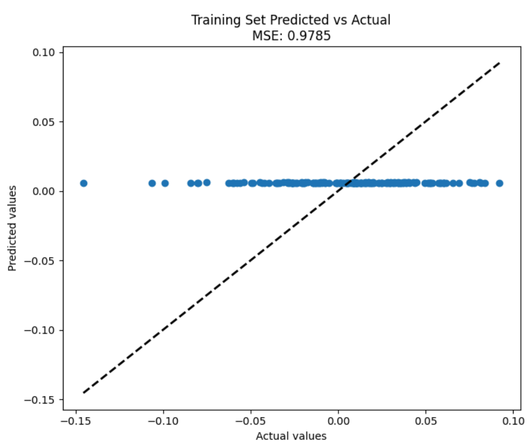
seed	width	ridge penalty	learning rate	epoch
25	40	0.0075	0.08	200

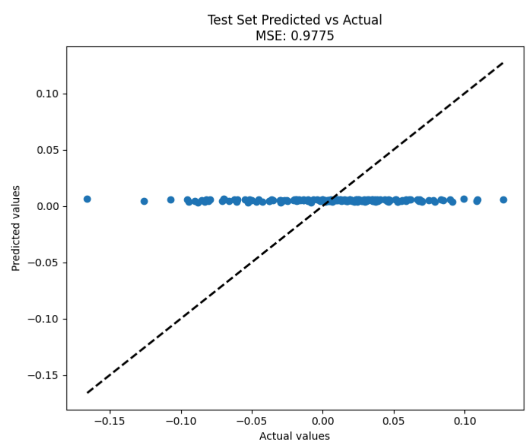
0 0.56
1 0.57
dtype: float32



0 0.56
1 0.80
2 -0.30
3 -0.12
4 0.56
5 0.56
6 0.56
7 0.56
8 0.56
9 0.56
10 0.56
dtype: float64

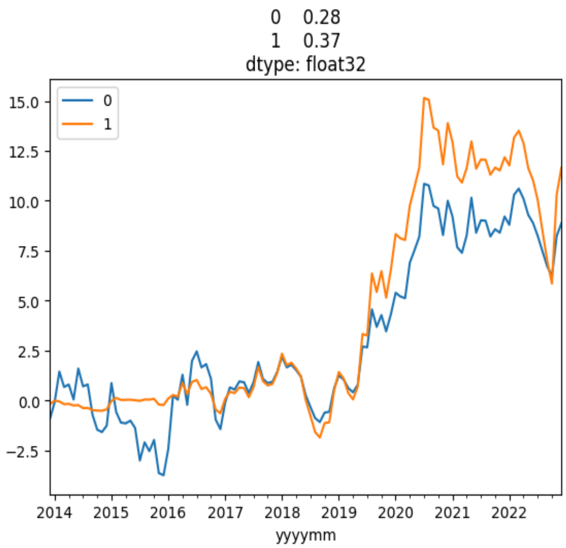


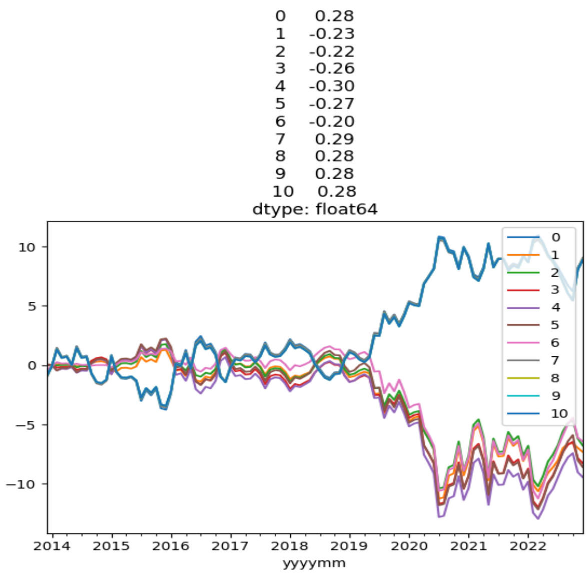


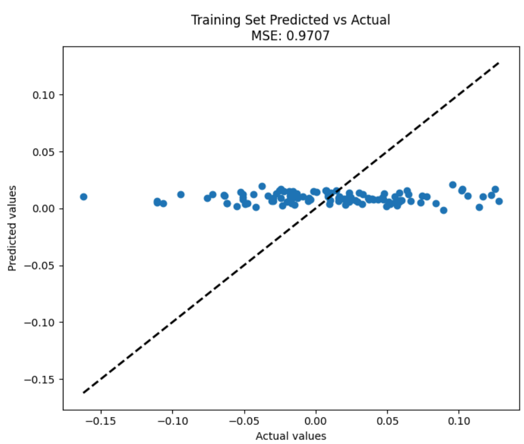


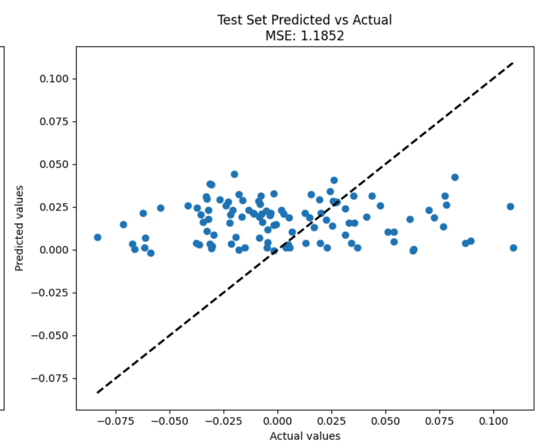
GLD Gradient Descent Epoch Variation

seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.01	100



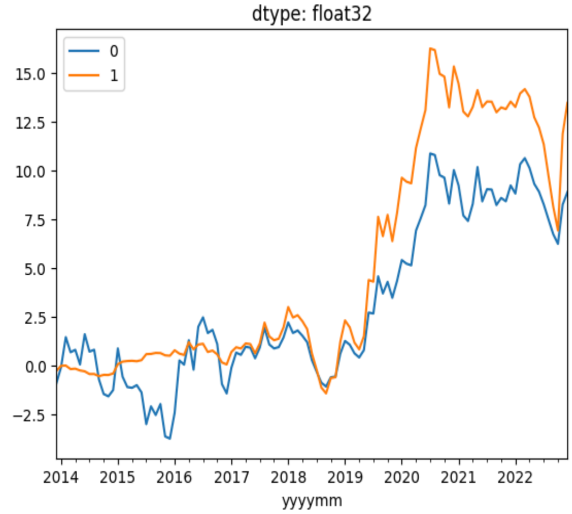




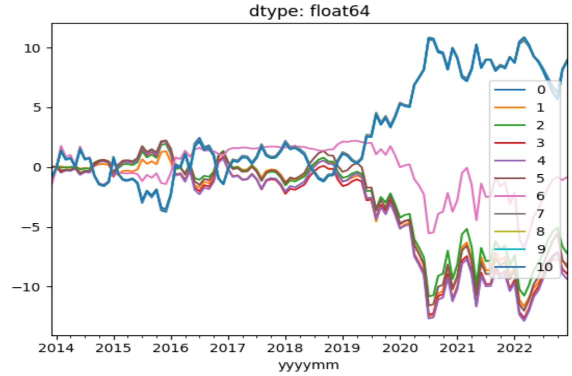


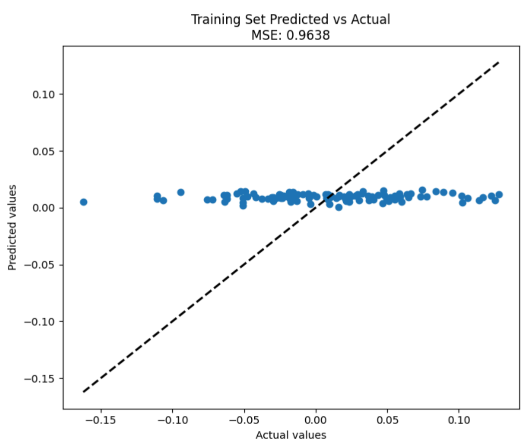
seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.01	150

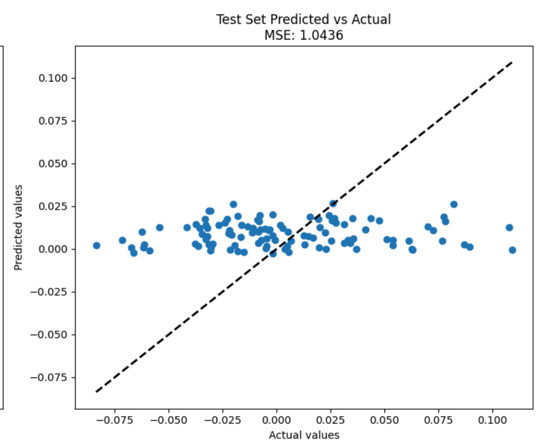
0   0.28
1   0.43
dtype: float32



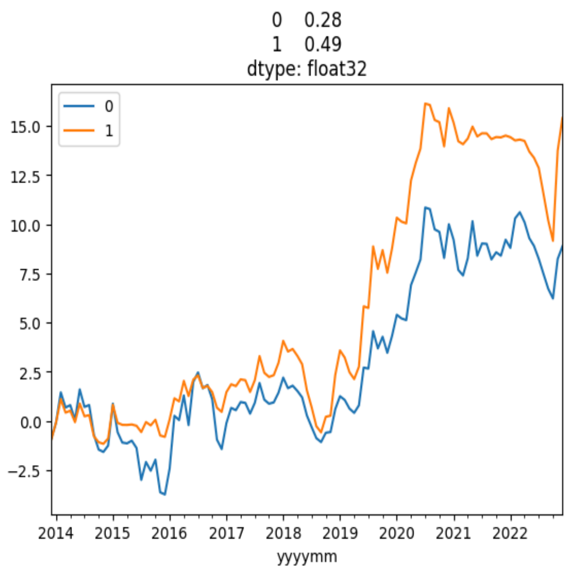
0   0.28
1   -0.26
2   -0.23
3   -0.28
4   -0.30
5   -0.27
6   -0.03
7   0.29
8   0.28
9   0.28
10   0.28
dtype: float64

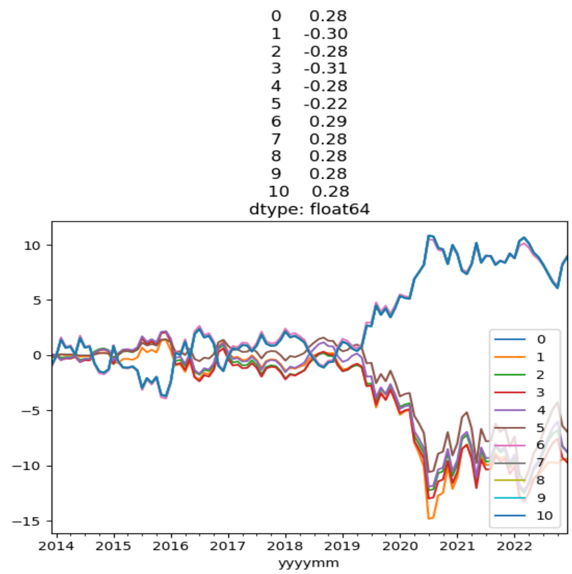


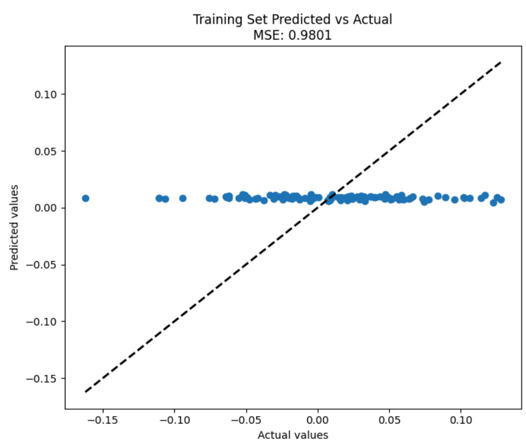


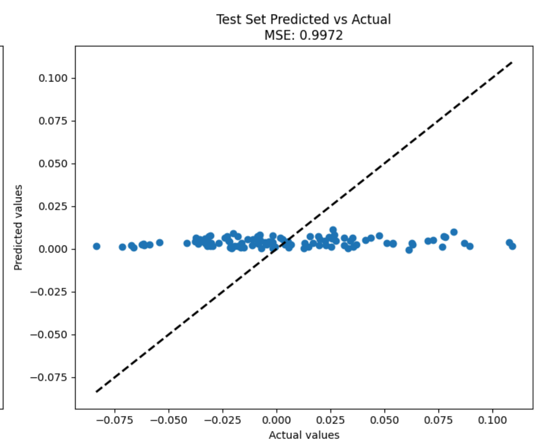


seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.01	300

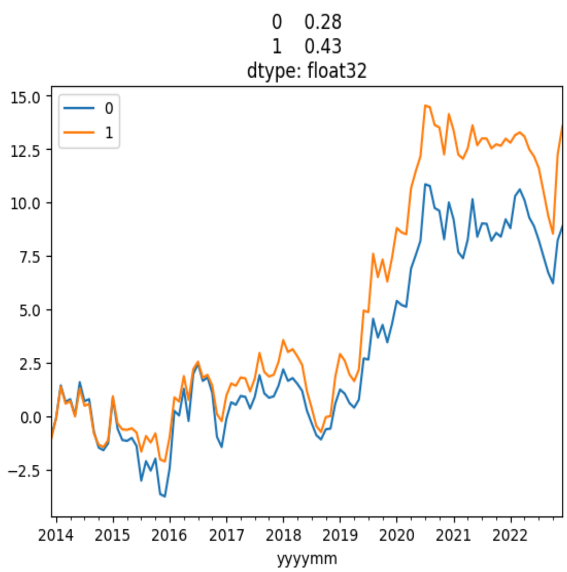


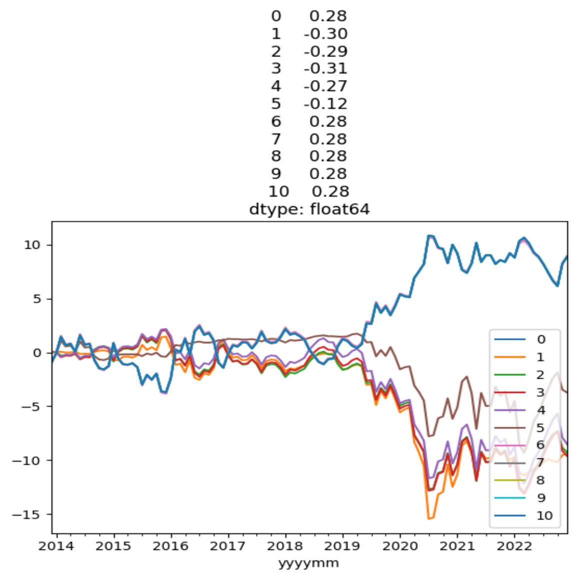


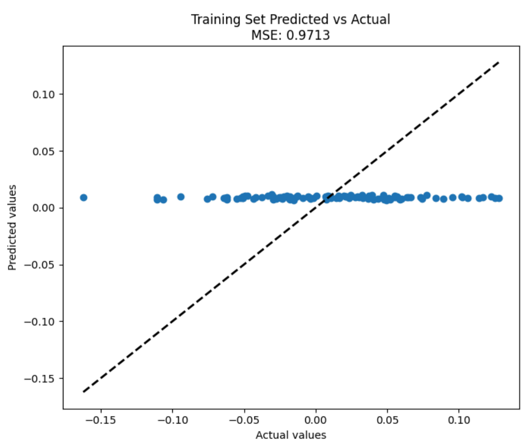


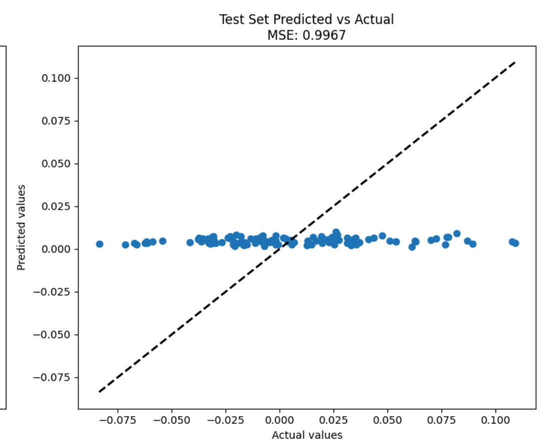


seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.01	350



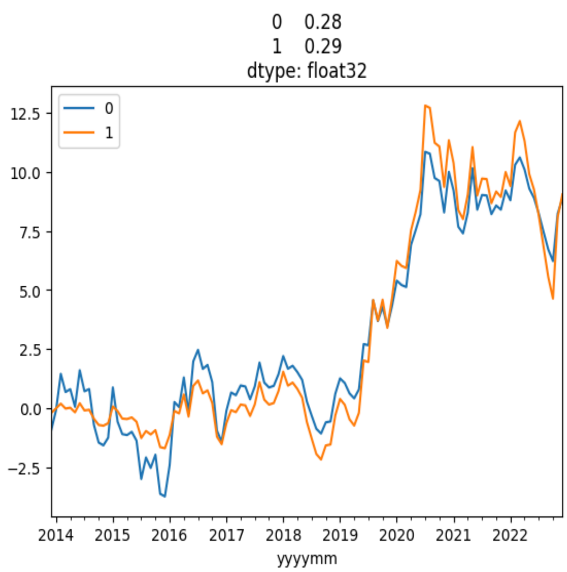


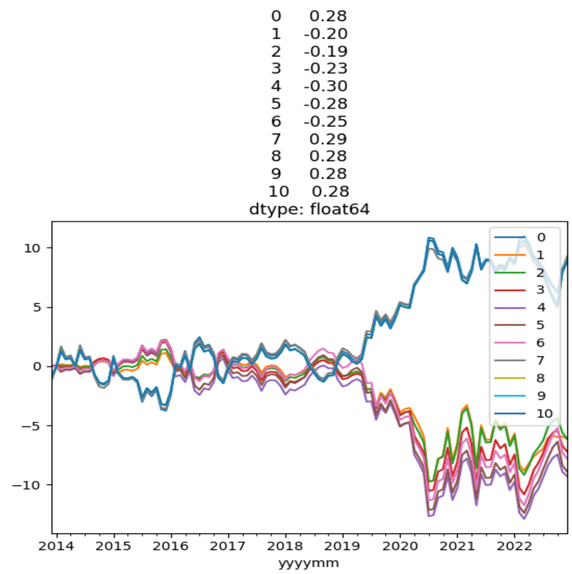


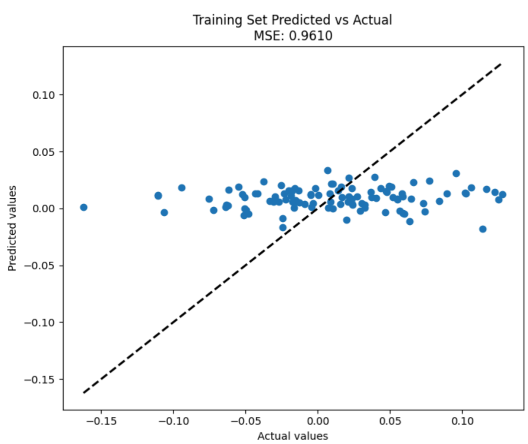


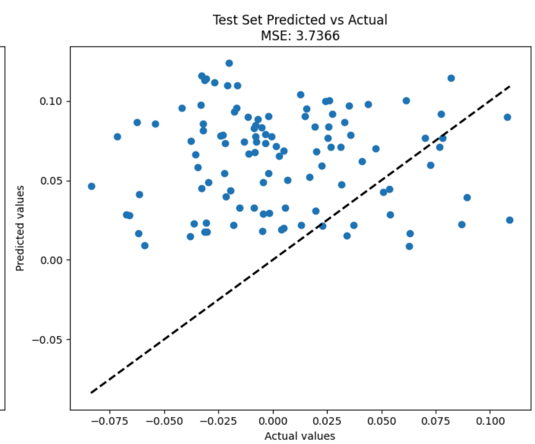
GLD Gradient Descent Learning Rate Variation

seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.001	250

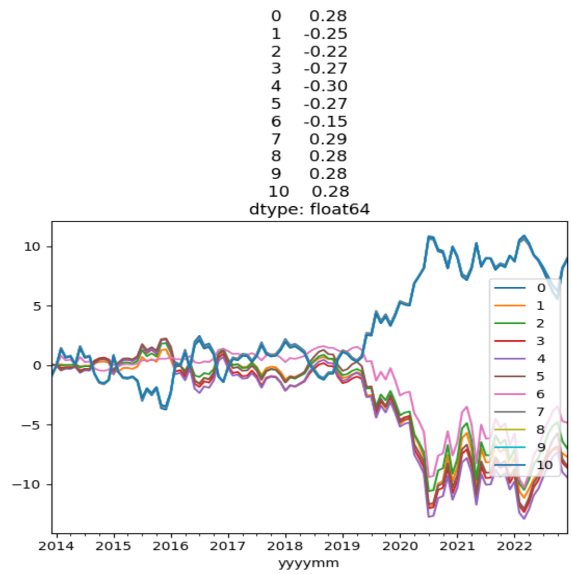


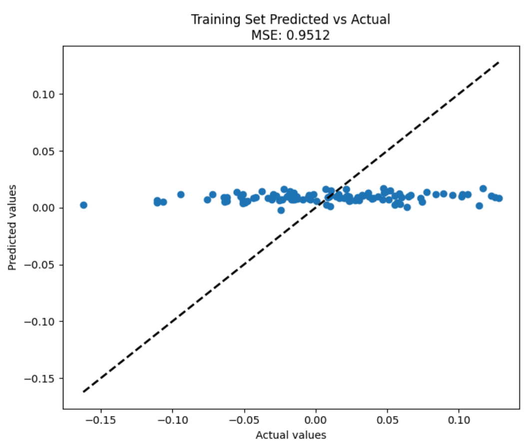


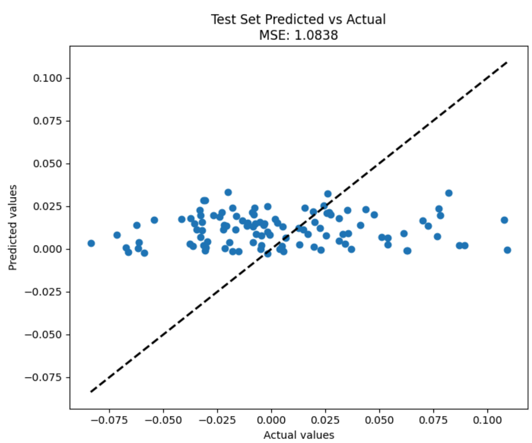




seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.005	250

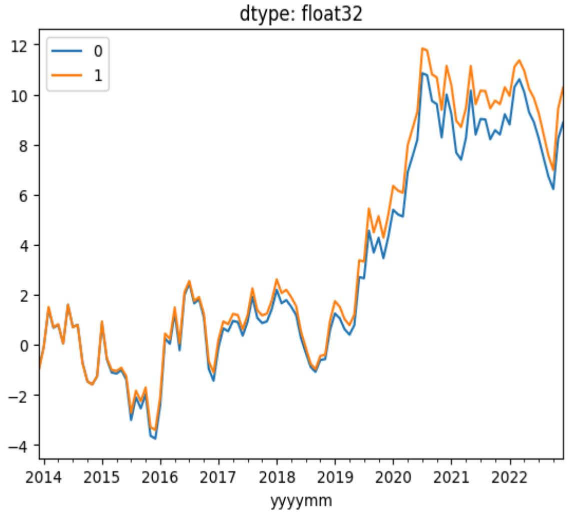




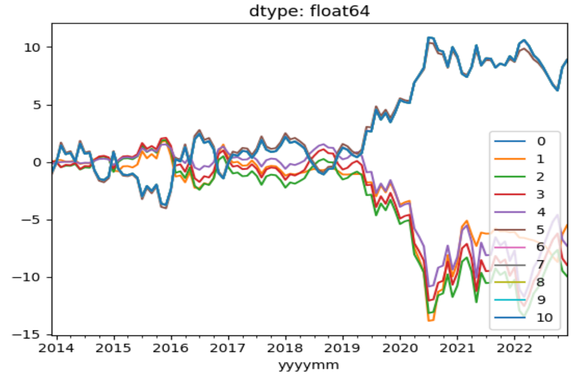


seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.02	250

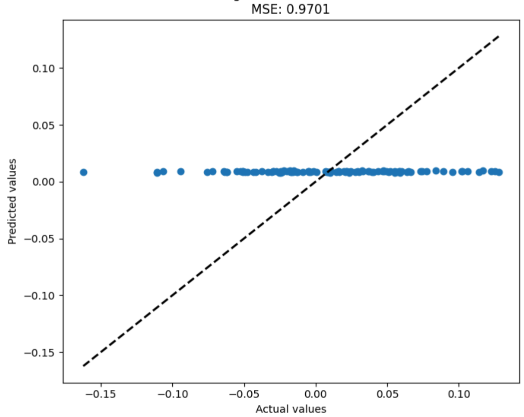
0   0.28
1   0.33
dtype: float32



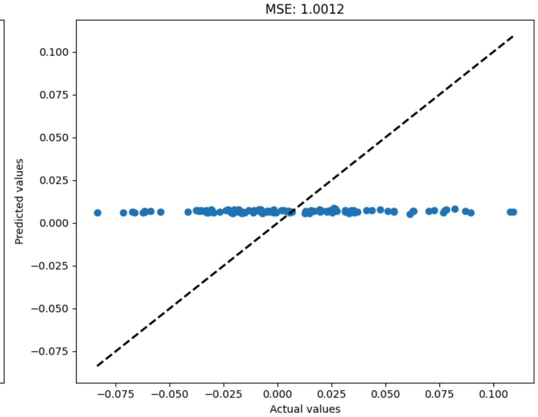
0   0.28
1   -0.17
2   -0.32
3   -0.29
4   -0.23
5   0.28
6   0.28
7   0.28
8   0.28
9   0.28
10   0.28
dtype: float64



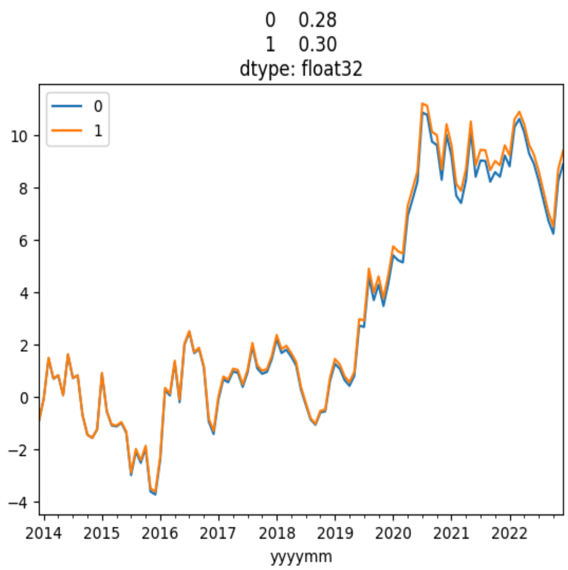
Training Set Predicted vs Actual
MSE: 0.9701

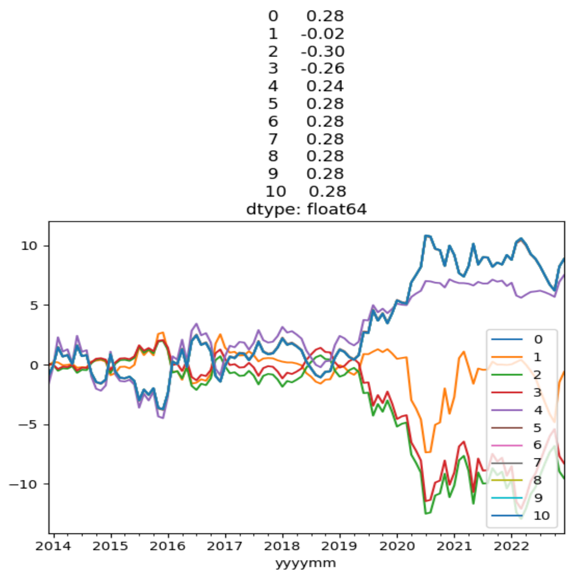


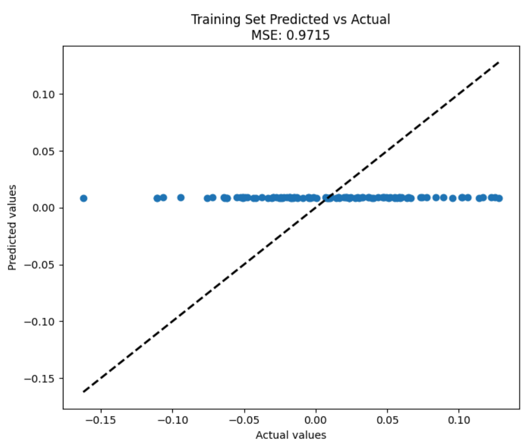
Test Set Predicted vs Actual
MSE: 1.0012

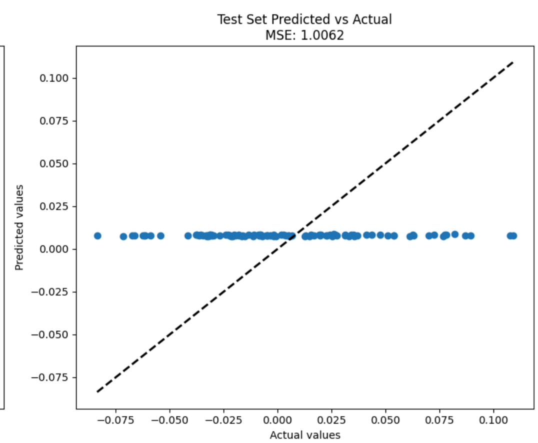


seed	width	ridge penalty	learning rate	epoch
10	200	0.035	0.025	250





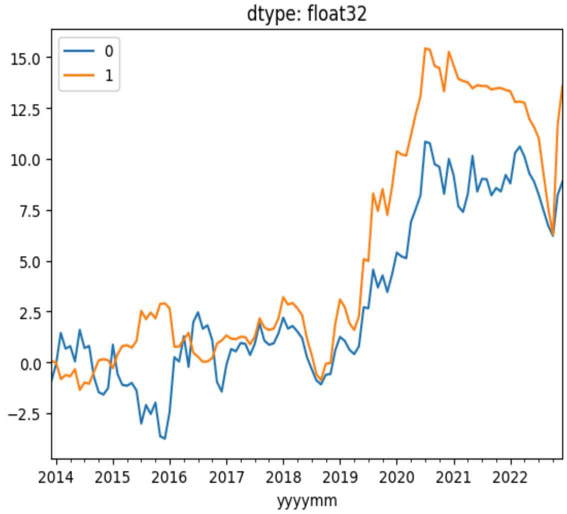




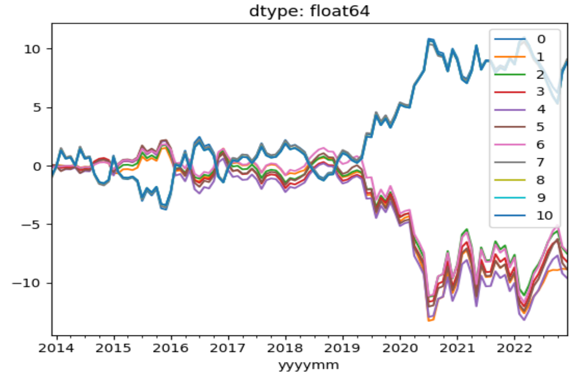
GLD Gradient Descent Ridge Penalty Variation

seed	width	ridge penalty	learning rate	epoch
10	200	0.01	0.01	250

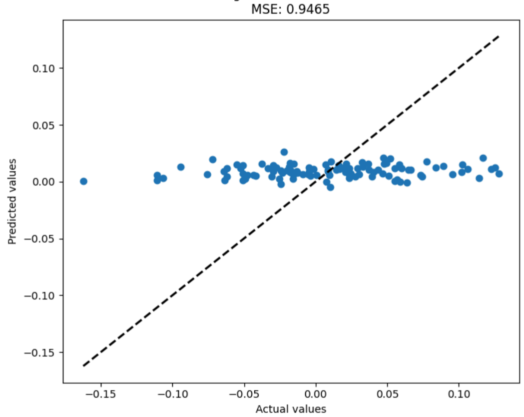
0 0.28
1 0.43
dtype: float32



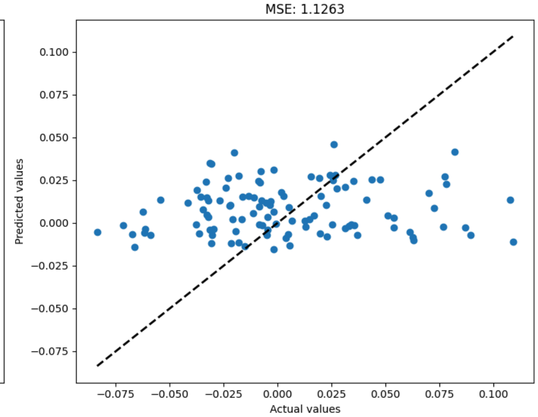
0 0.28
1 -0.28
2 -0.24
3 -0.26
4 -0.31
5 -0.28
6 -0.23
7 0.29
8 0.28
9 0.28
10 0.28
dtype: float64



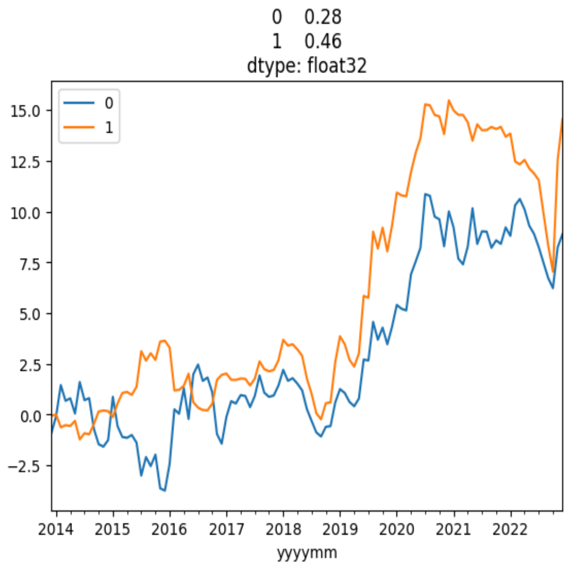
Training Set Predicted vs Actual
MSE: 0.9465

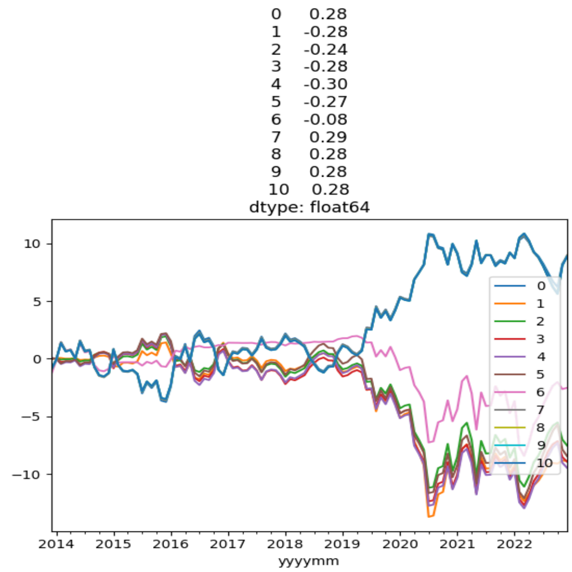


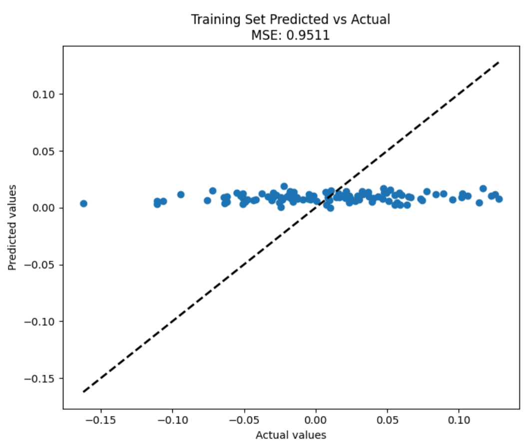
Test Set Predicted vs Actual
MSE: 1.1263

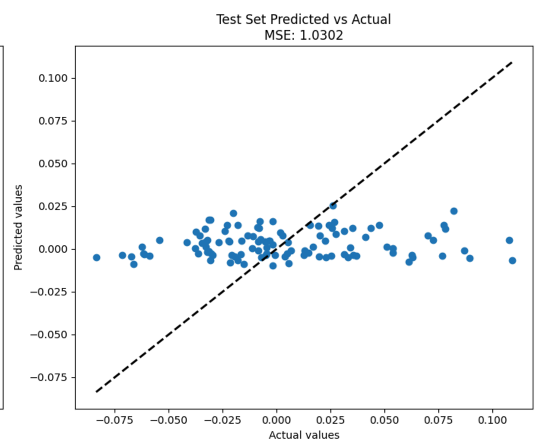


seed	width	ridge penalty	learning rate	epoch
10	200	0.02	0.01	250

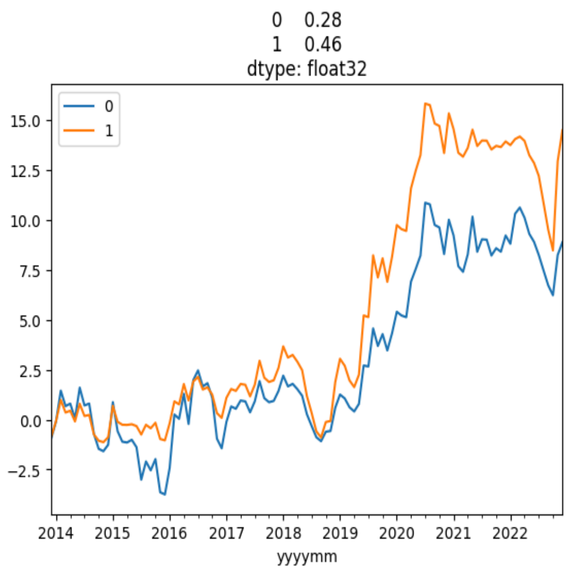


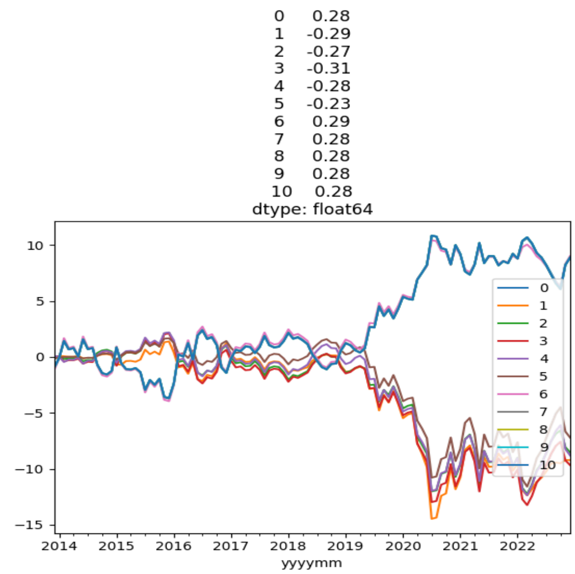


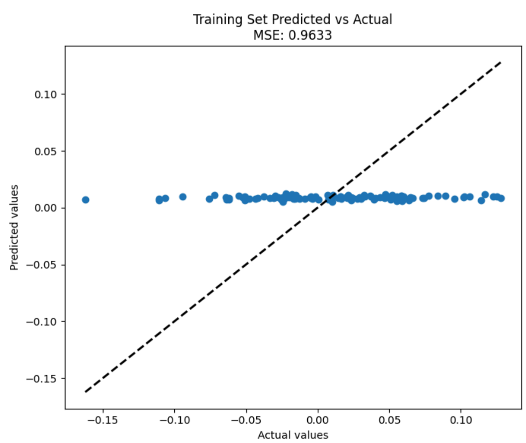


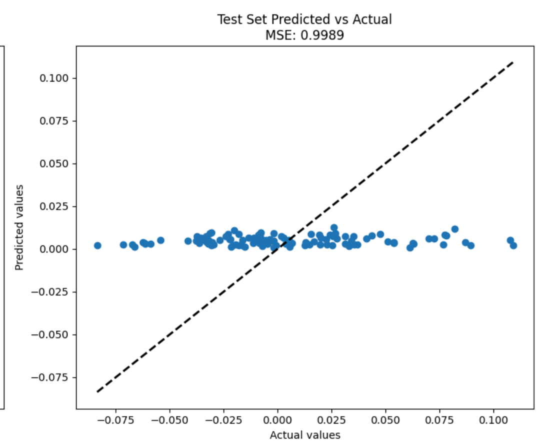


seed	width	ridge penalty	learning rate	epoch
10	200	0.04	0.01	250

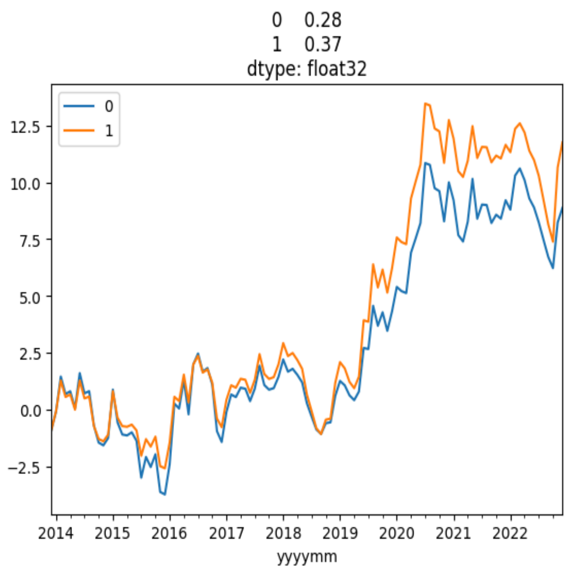


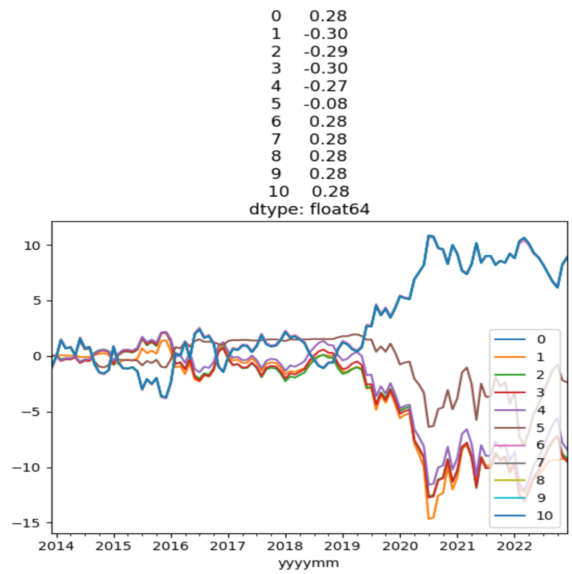


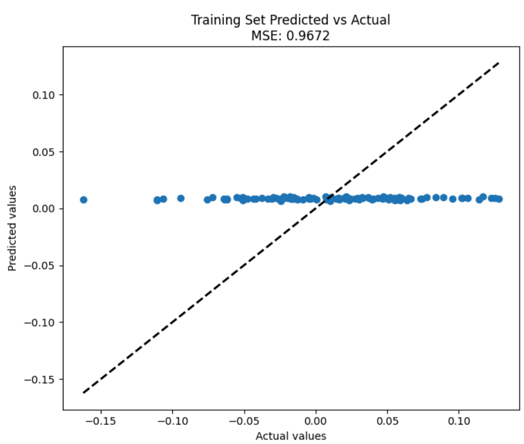


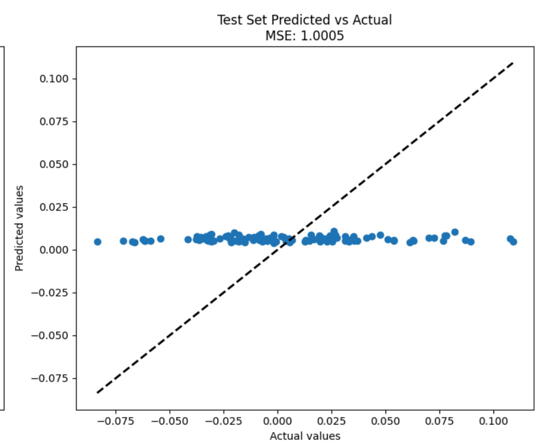


seed	width	ridge penalty	learning rate	epoch
10	200	0.05	0.01	250



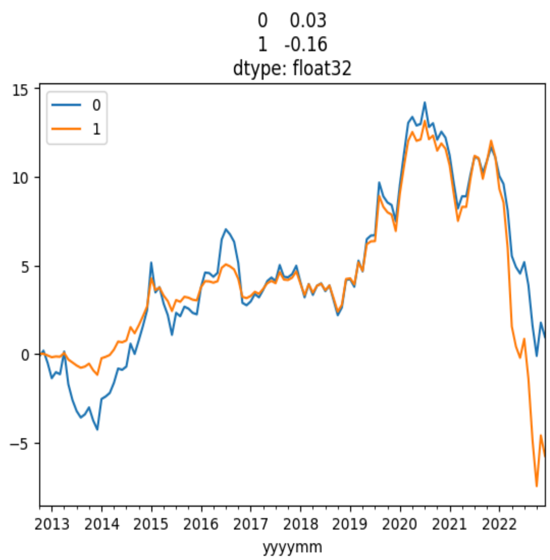


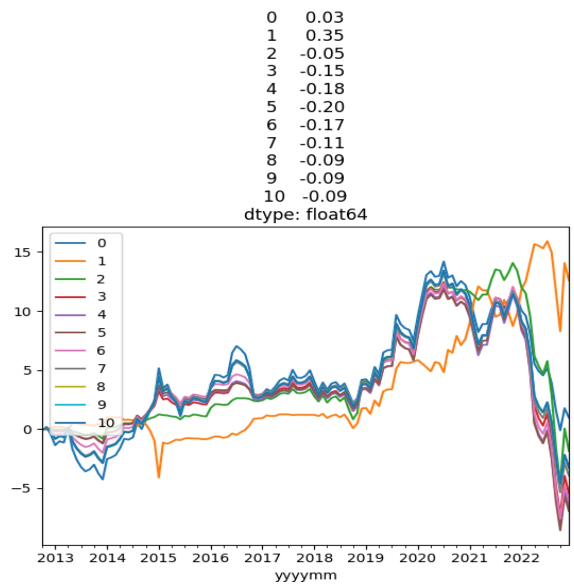


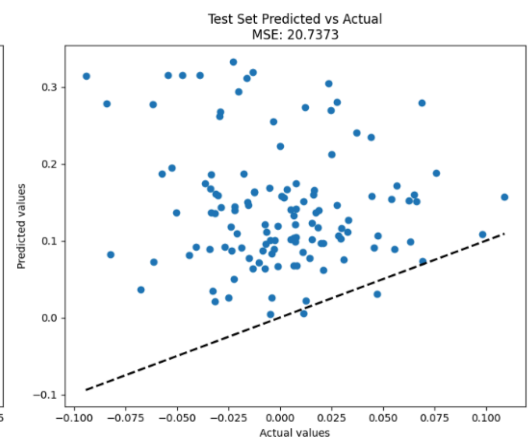


TLT Gradient Descent Epoch Variation

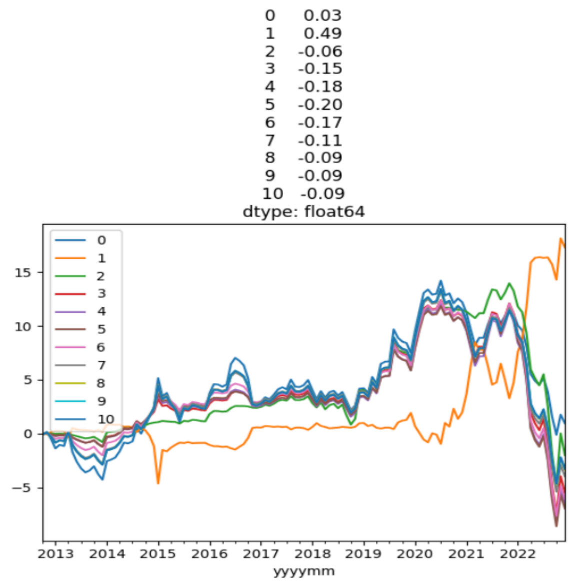
seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	200

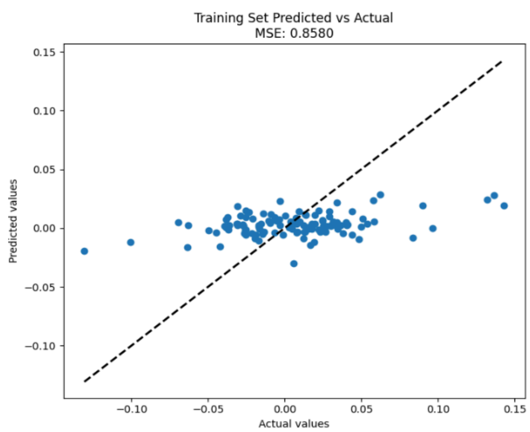


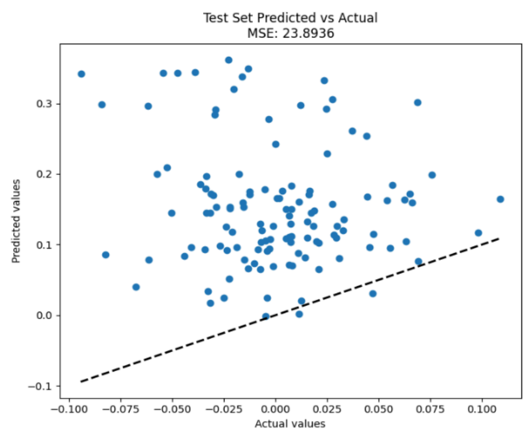




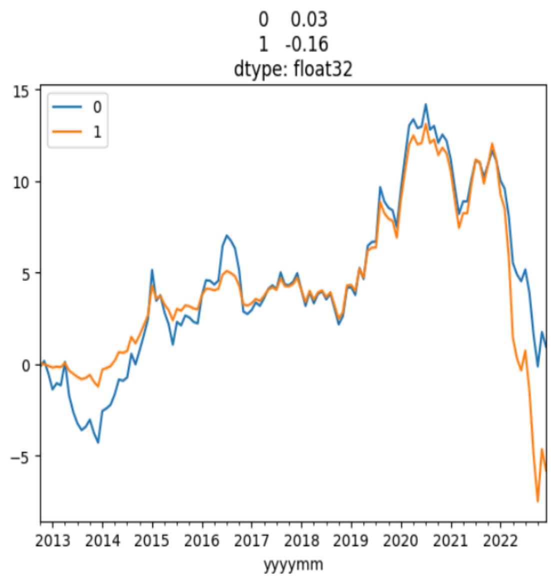
seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	300

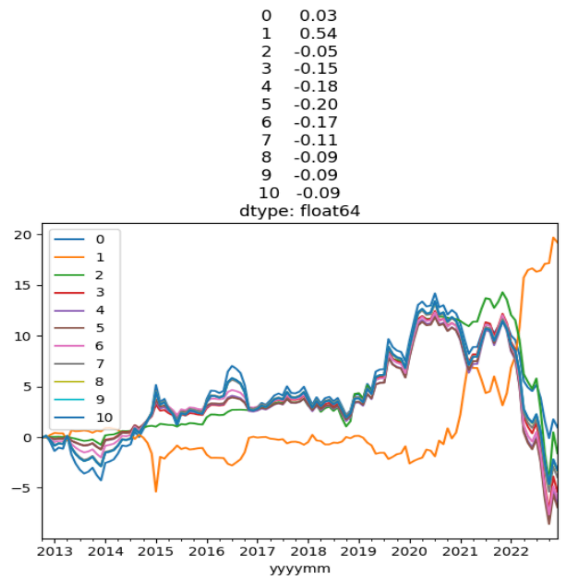


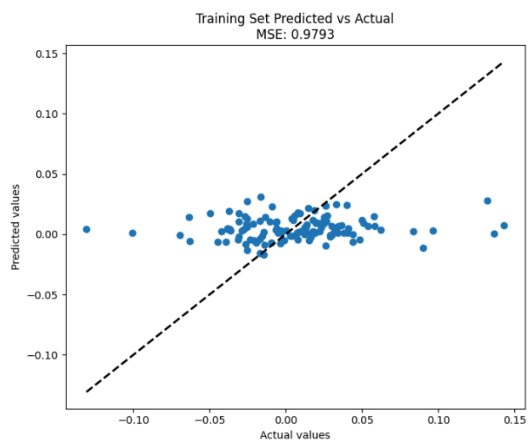


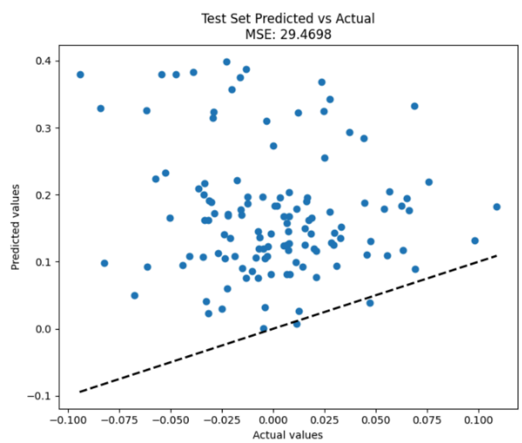


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	400

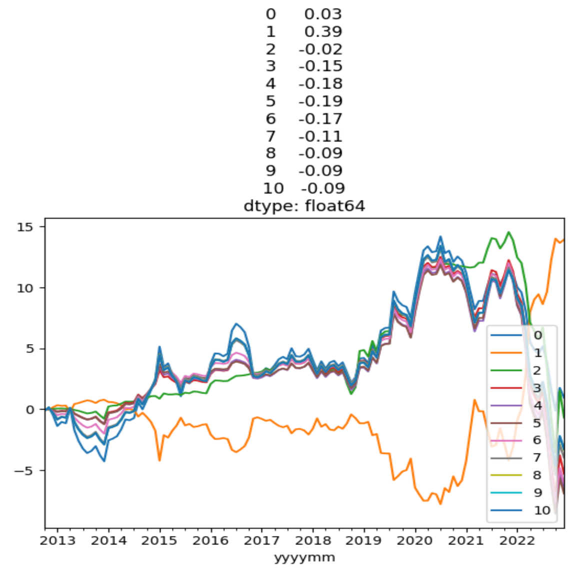


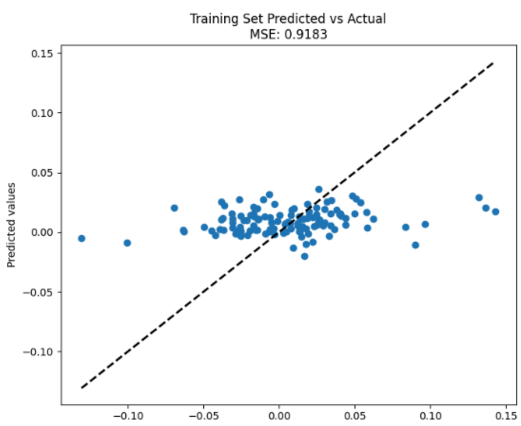


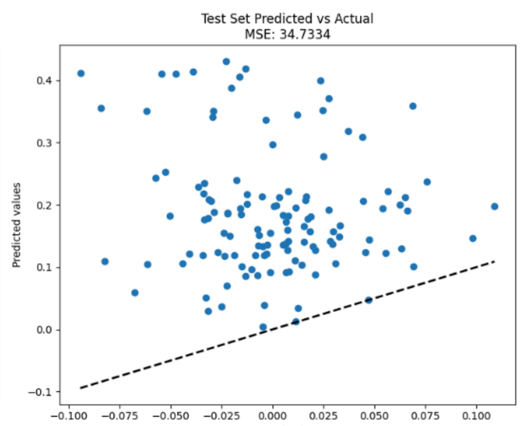




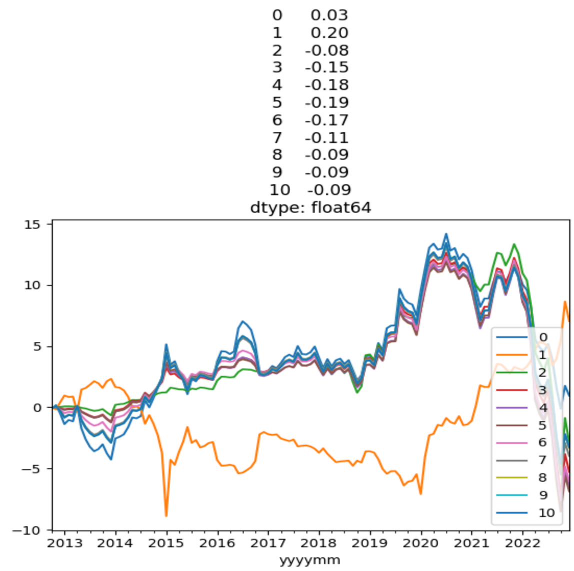
seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	500

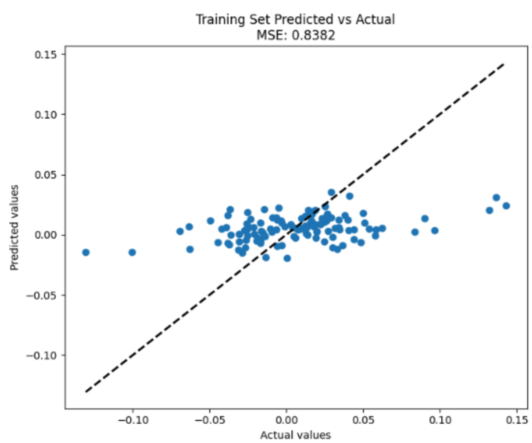


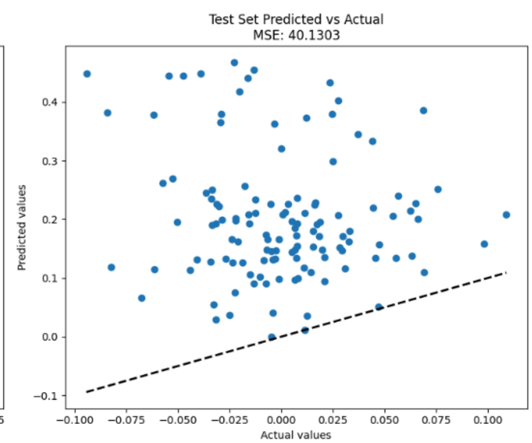




seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	800

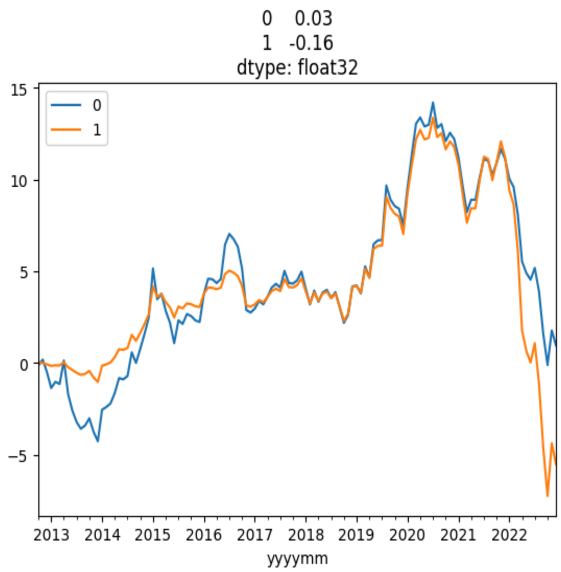


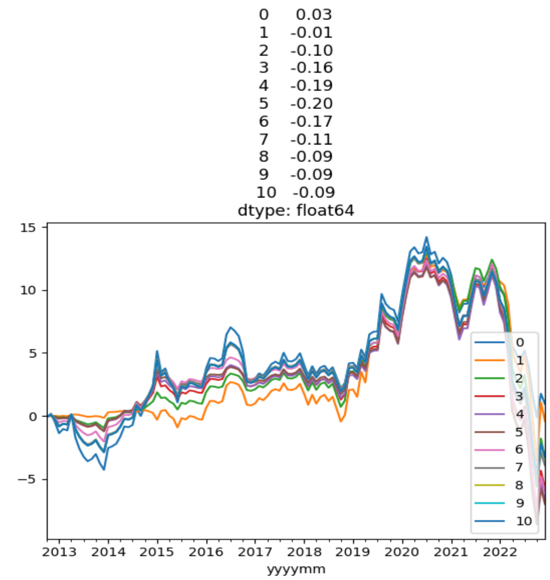


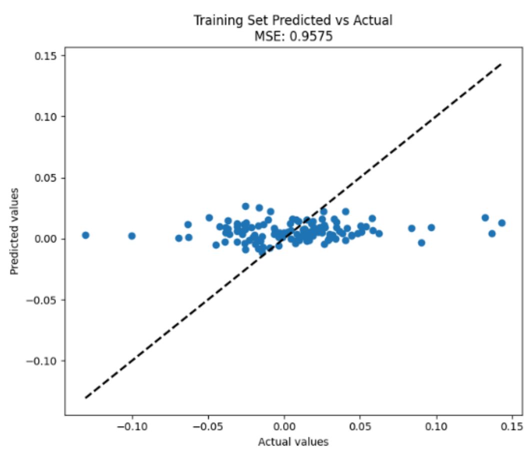


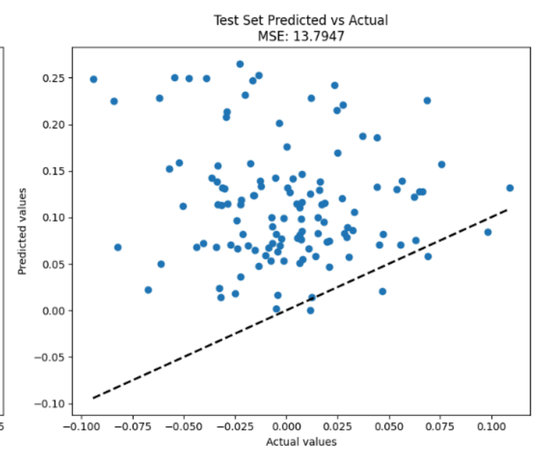
TLT Learning Rate Variation

seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.03	400

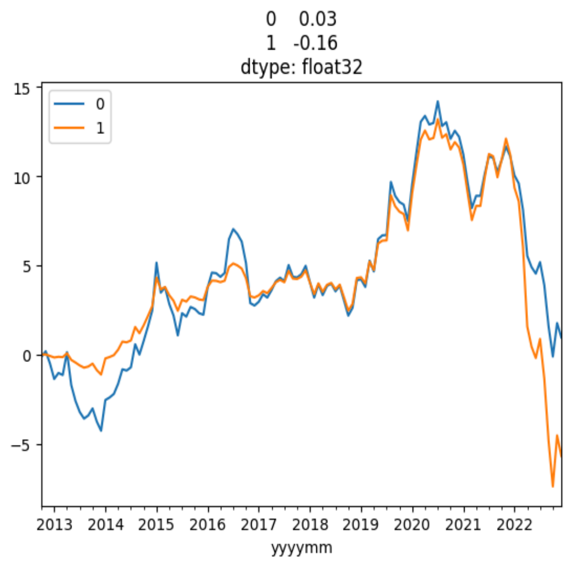


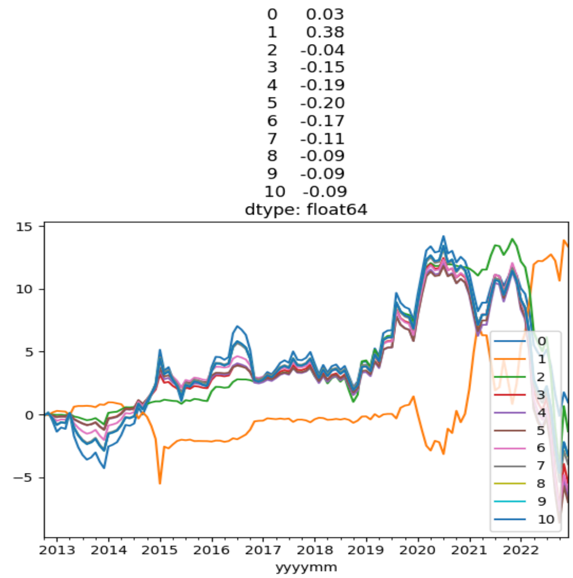


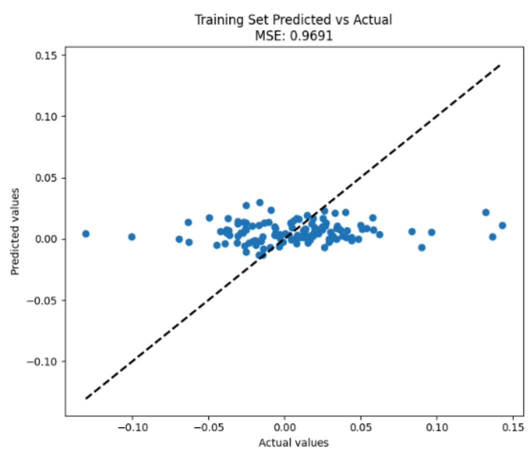


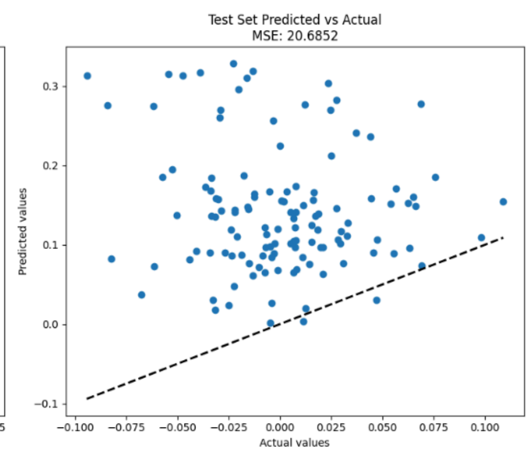


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.06	400

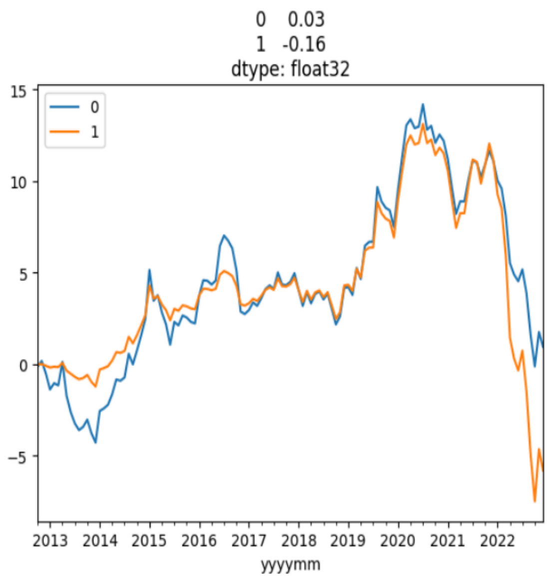


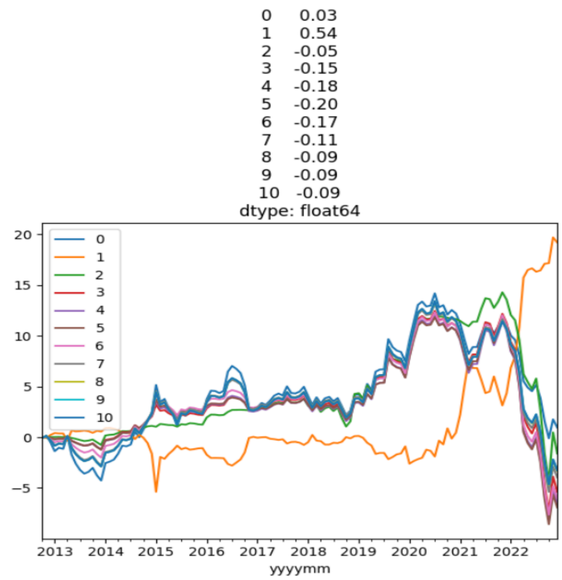


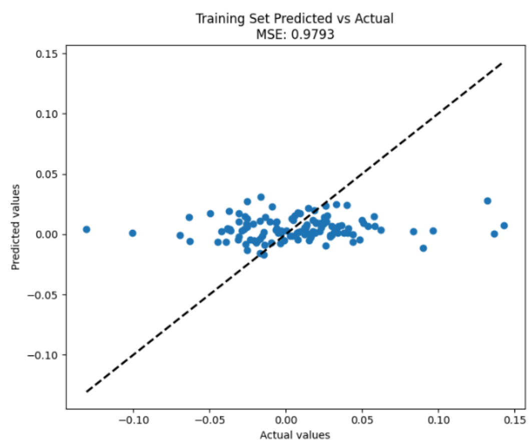


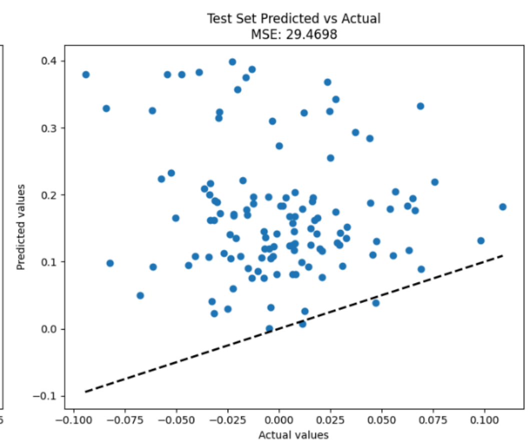


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	400

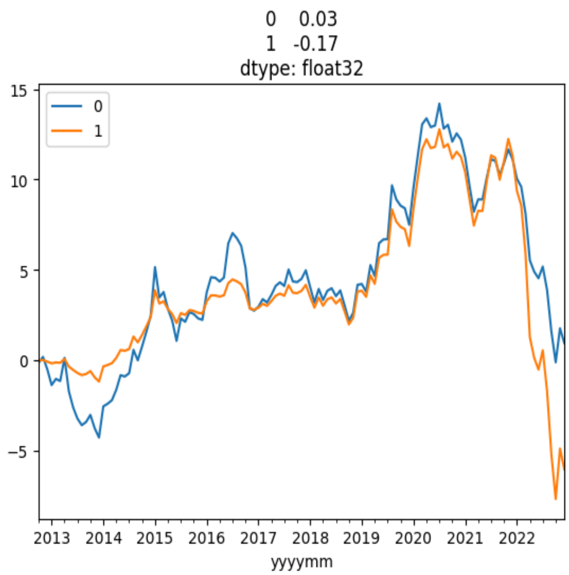


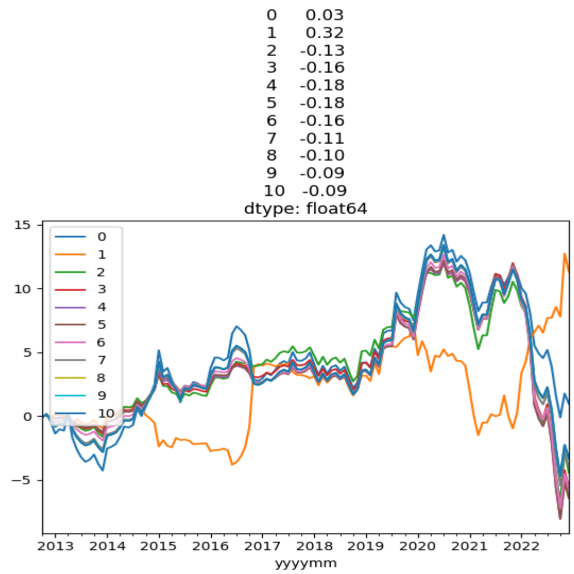


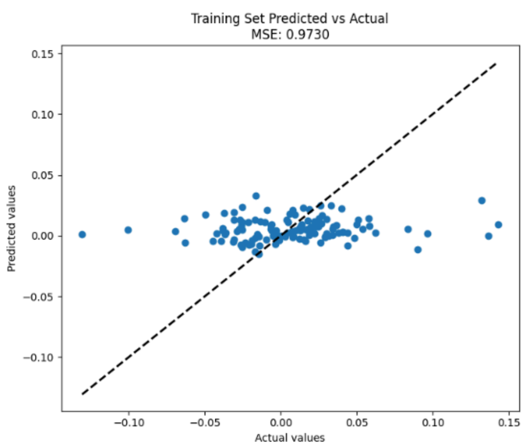


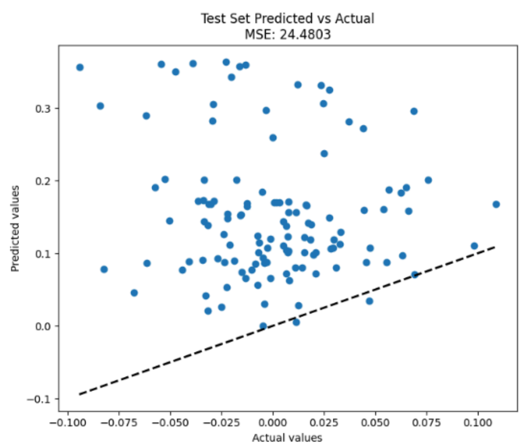


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.24	400

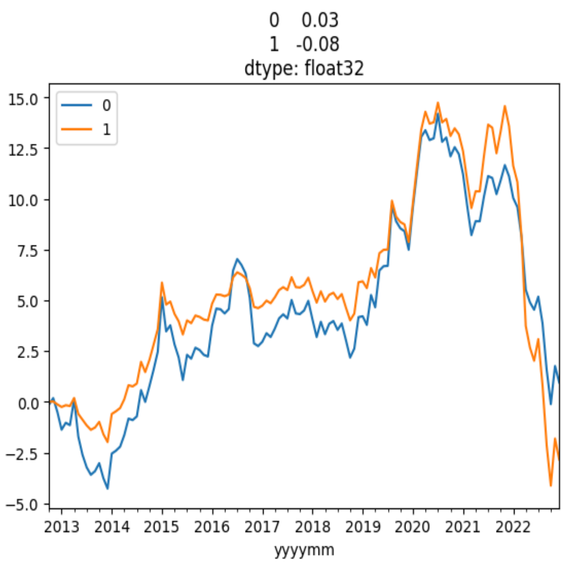


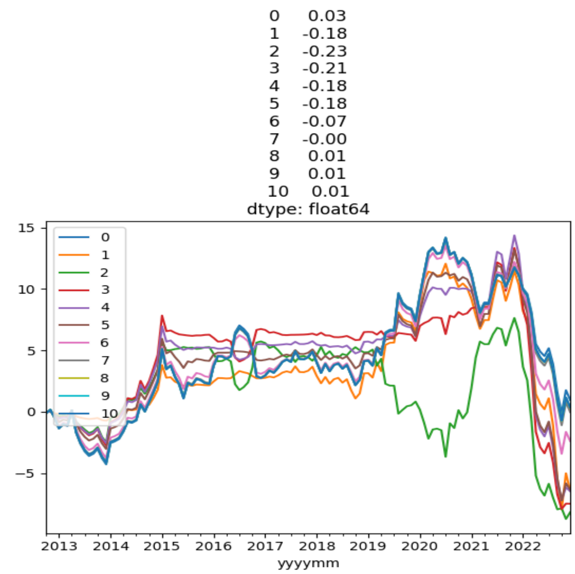


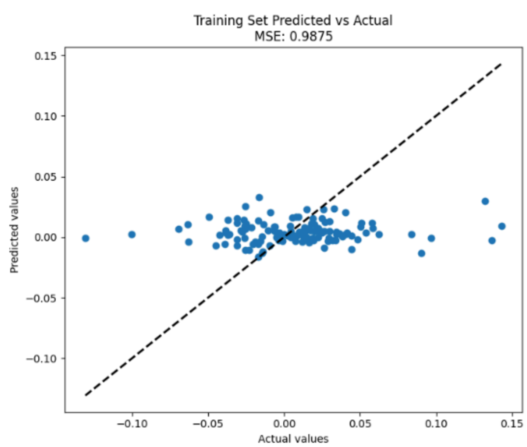


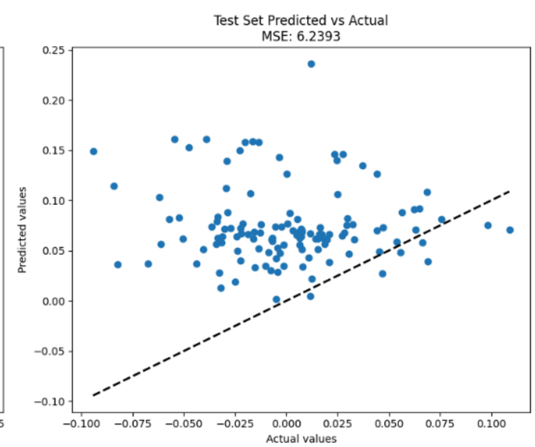


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.48	400



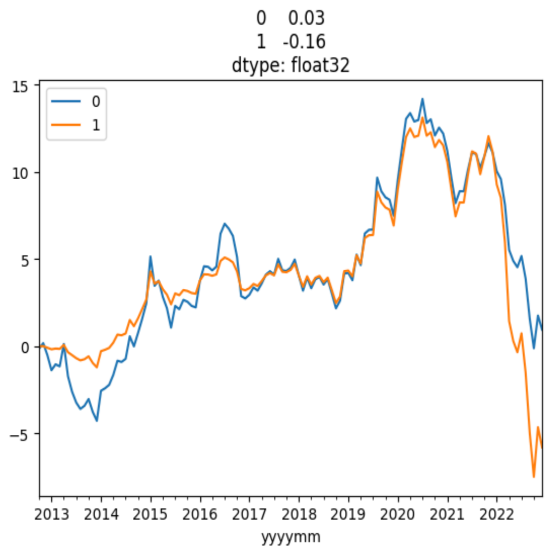


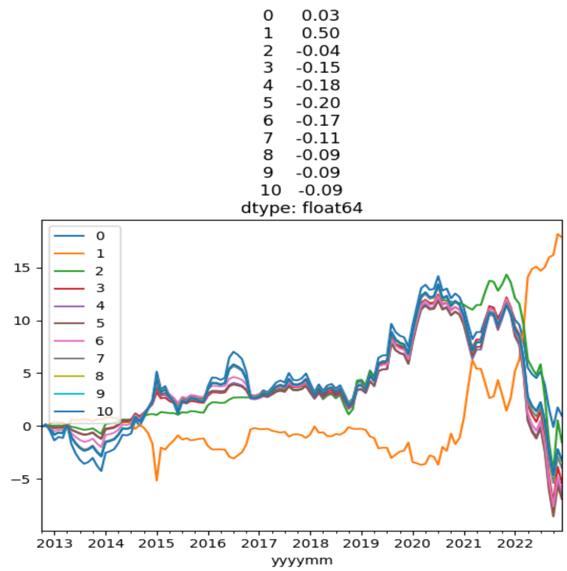


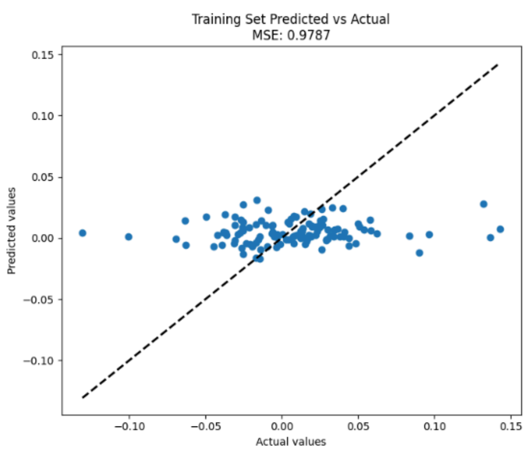


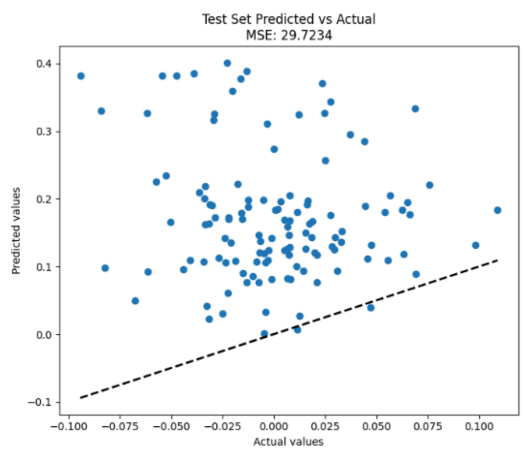
TLT Ridge Penalty Variation

seed	width	ridge penalty	learning rate	epoch
10	200	0.0000025	0.12	400

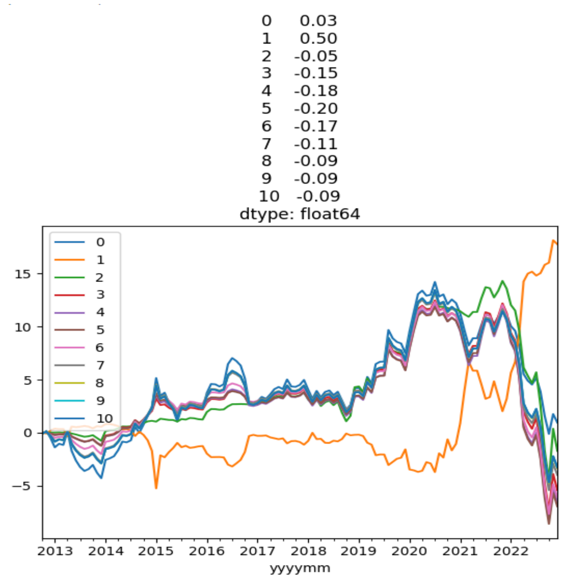


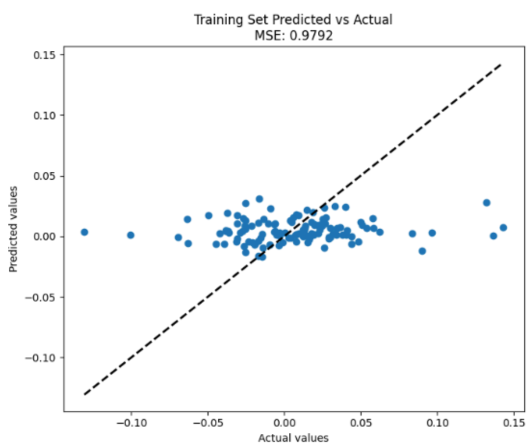


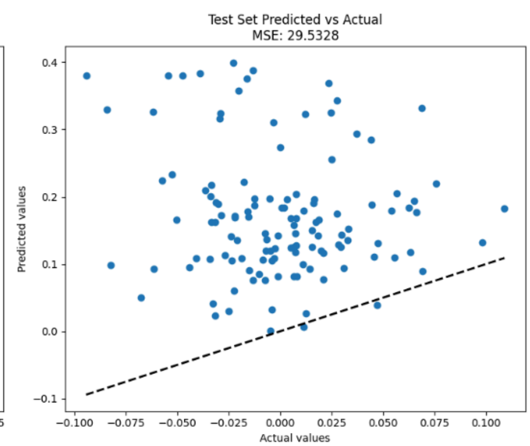




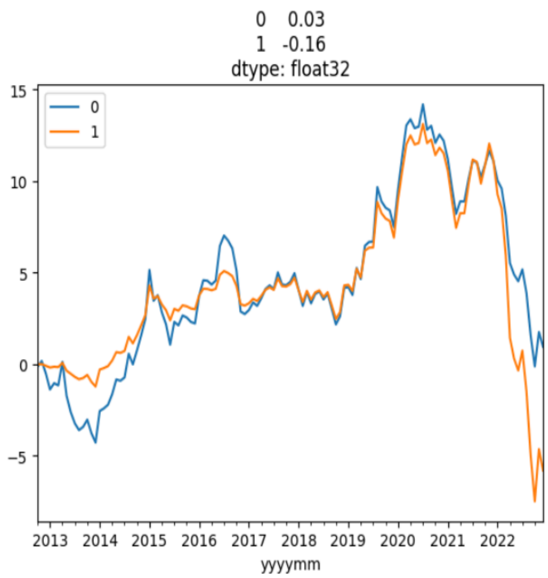
seed	width	ridge penalty	learning rate	epoch
10	200	0.000005	0.12	400

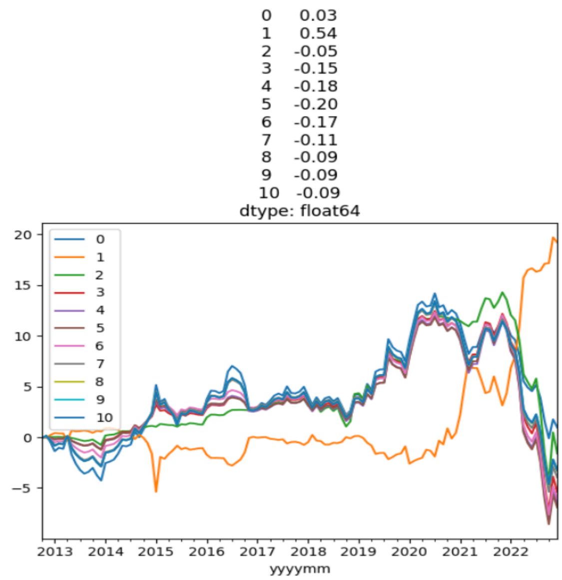


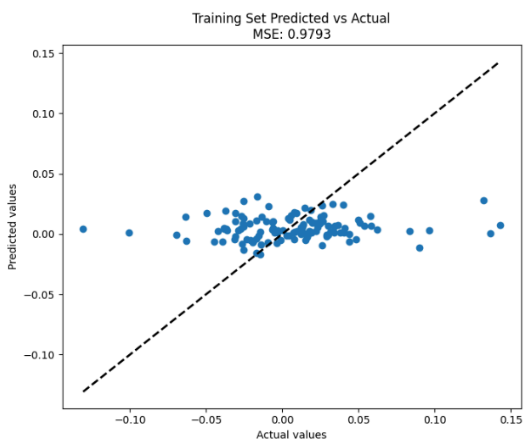


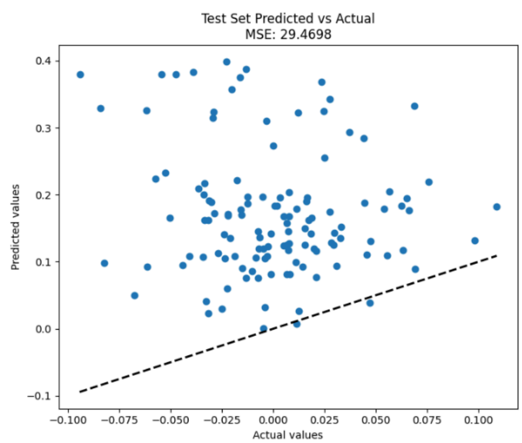


seed	width	ridge penalty	learning rate	epoch
10	200	0.00001	0.12	400

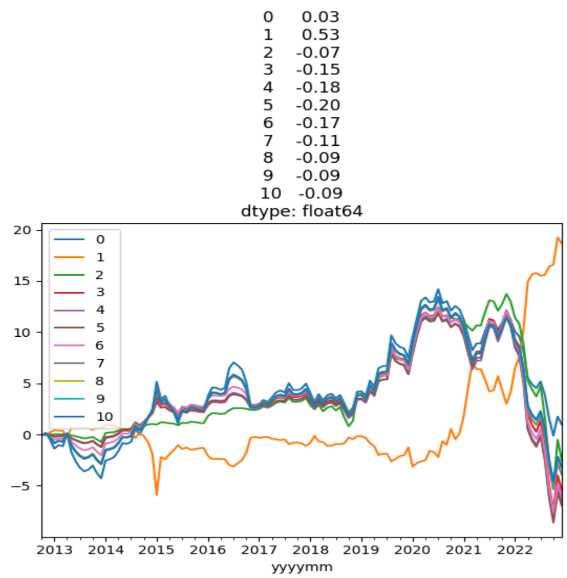


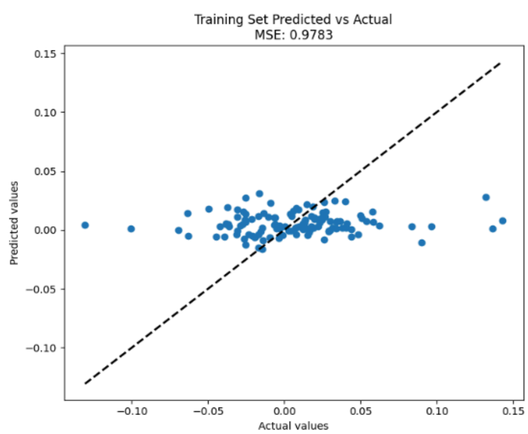


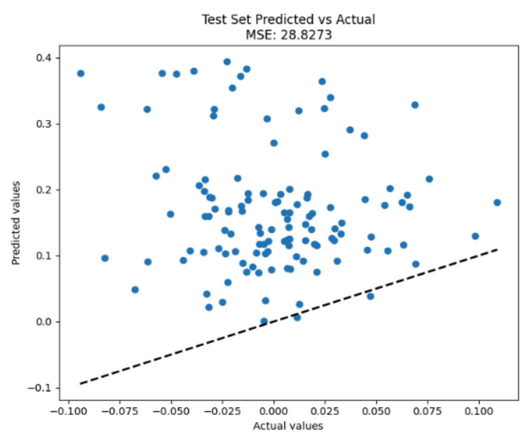




seed	width	ridge penalty	learning rate	epoch
10	200	0.00005	0.12	400







seed	width	ridge penalty	learning rate	epoch
10	200	0.0001	0.12	400