Machine Learning Basics on Linear Factor Models

5. An Empirical Assessment on US equities

This slide set is largely based on Gu, Kelly, and Xiu (2020) and its Internet Appendix.

Potential Factor Zoo for Return Predictions

- Hou, Xue and Zhang (2018) summarize that 452 “factors” so far are “discovered” in the literature that are shown to be able to predict returns.
- The factors are classified into six groups
  - Momentum
  - Value vs growth
  - Investment (firm’s investment activities)
  - Profitability
  - Intangibles
  - Trading frictions
- They replicate the anomalies, and conclude:

> Our work has important implications. First, anomalies are not created equal. The value and momentum anomalies replicate well, along with the investment and profitability anomalies. Most of these anomalies reside in value-weighted returns, which account for 97% of the aggregate market capitalization. In

Empirical study of US equities: key takeaways

- **Adding many variables impacts model predictability**
- Large scale empirical analysis of ~30k individual stocks from 1957-2016; 94 characteristics for each stock, interactions of each characteristic with 8 aggregate time series variables; 74 industry sector dummy variables, totaling more than 900 baseline signals

Method	Tuning parameter	Performance (OOS R^2 per month)
OLS-3 panel benchmark (size, btm, mom)	No	0.16%
OLS panel 900+ predictors	No	Negative
Elastic net 900+ predictors	Yes	0.11%
PCR / PLS 900+ predictors	Yes	0.26%, 0.27%
Trees 900+ predictors	Yes	0.33%
Neural networks 900+ predictors	Yes	0.40%

By expanding the OLS panel to include 900+ predictors, predictability vanishes immediately (due to inefficiency of OLS)

Elastic net immediately solves OLS inefficiency problem; while PCR / PLS further improves by reducing further dimension of predictor set

Allowing nonlinearites improves predictions substantially

Gu, S., Kelly, B., Xiu, D. Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, Volume 33, Issue 5, May 2020, Pages 2223–2273.

Empirical study of US equities: key takeaways

- **Aggregate portfolio predictability also improves with ML**
- Build bottom-up portfolio-level forecasts from the stock-level forecasts produced by ML models (e.g., bottom-up forecasts of the S&P500 portfolio return)

Method	Tuning parameter	Performance (OOS R^2 per month)	
OLS-3 panel benchmark (size, btm, mom)	No	(0.22%)	Benchmark 3-characteristic OLS model produced negative performance vs. the GLM
Generalized linear model	Yes	0.71%	
Trees 900+ predictors	Yes	1.08%	Trees and neural networks further improve upon this
Neural networks 900+ predictors	Yes	1.80%	

- Individual stock returns of smallest and least liquid stocks in sample behave erratically making it more pronounced the predictive power at the portfolio level vs. the stock level
- Aggregating into portfolio averages out much of the unpredictable stock-level noise and boosts the signal strength, helping ML detecting predictive associations

Empirical study of US equities: methodology

1 Statistical model describing a method's general functional form for risk premium predictions

2 Objective function for estimating model parameters

- All estimates share objective of minimizing mean squared predictions error (MSE)
- Regularization is introduced through variations of the MSE objective (e.g., adding parameterization penalties and robustification against outliers)

3 Design disjoint subsamples for estimation and testing and to introduce the notion of "hyperparameter tuning"

Training subsample

- Estimate the model subject to a specific set of tuning parameters values

→

Validation subsample

- Tune hyperparameters
- Construct forecasts for data points
- Calculate objective function based on forecast errors from validation sample
- Evaluate using lowest MSE
- Simulate OOS test of the model

→

Testing subsample

- Re-train models in combined window (validation + test)
- Evaluate method's predictive performance with best / optimal parameters

Gu, S., Kelly, B., Xiu, D. Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, Volume 33, Issue 5, May 2020, Pages 2223–2273.

Empirical study of US equities: performance evaluation

- $\boldsymbol{R^2_{OOS}}$ **is calculated to assess the predictive performance for individual excess stock return forecasts**

$$R^2_{OOS} = 1 - \frac{\sum_{(1,t)\epsilon\mathcal{T}_3}\left(r_{i,t+1} - \hat{r}_{i,t+1}\right)^2}{\sum_{(1,t)\epsilon\mathcal{T}_3} r^2_{i,t+1}}$$

  - $\mathcal{T}_3$ indicates that fits are only assessed on the testing subsample, whose data never enters into model estimation or tuning
  - R^2_{OOS} pools prediction errors across firms and over time into a grand panel-level assessment of each model

- Caution when using R^2_{OOS}…

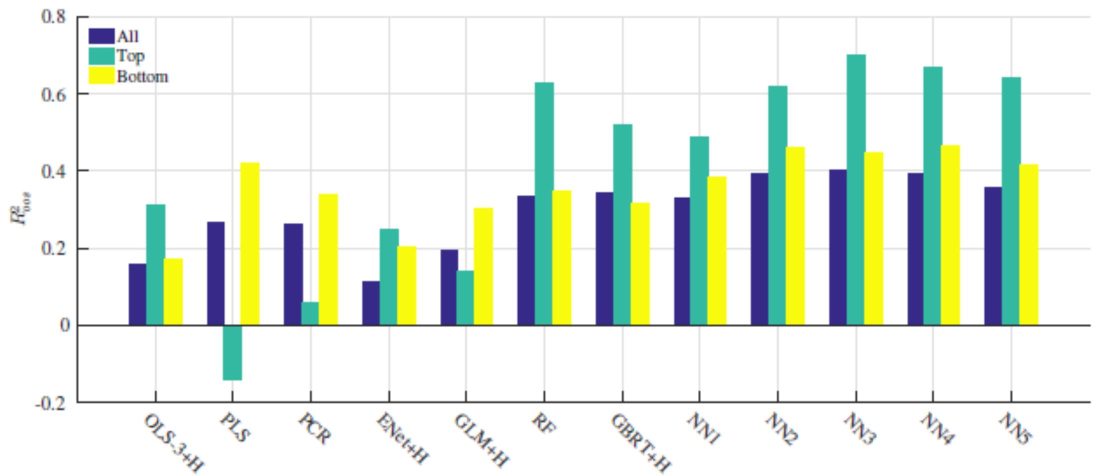
A subtle but important aspect of our R^2 metric is that the denominator is the sum of squared excess returns *without demeaning*. In many out-of-sample forecasting applications, predictions are compared against historical mean returns. While this approach is sensible for the aggregate index or long-short portfolios, for example, it is flawed when it comes to analyzing individual stock returns. Predicting future excess stock returns with historical averages typically *underperforms* a naive forecast of zero by a large margin. That is, the historical mean stock return is so noisy that it artificially lowers the bar for "good" forecasting performance. We avoid this pitfall by benchmarking our R^2 against a forecast value of zero. To give an indication of the importance of this choice, when we benchmark model predictions against historical mean stock returns, the out-of-sample monthly R^2 of all methods rises by roughly three percentage points.

Empirical study of US equities: the cross section of individual stocks

- **Comparison of ML techniques in terms of their R^2_{OOS}**

Table 1: Monthly Out-of-sample Stock-level Prediction Performance (Percentage R^2_{oos})

	OLS +H	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
All	-3.46	0.16	0.27	0.26	0.11	0.19	0.33	0.34	0.33	0.39	0.40	0.39	0.36
Top 1000	-11.28	0.31	-0.14	0.06	0.25	0.14	0.63	0.52	0.49	0.62	0.70	0.67	0.64
Bottom 1000	-1.30	0.17	0.42	0.34	0.20	0.30	0.35	0.32	0.38	0.46	0.45	0.47	0.42



Note: In this table, we report monthly R^2_{oos} for the entire panel of stocks using OLS with all variables (OLS), OLS using only size, book-to-market, and momentum (OLS-3), PLS, PCR, elastic net (ENet), generalize linear model (GLM), random forest (RF), gradient boosted regression trees (GBRT), and neural networks with one to five layers (NN1–NN5). "+H" indicates the use of Huber loss instead of the l_2 loss. We also report these R^2_{oos} within subsamples that include only the top 1,000 stocks or bottom 1,000 stocks by market value. The lower panel provides a visual comparison of the R^2_{oos} statistics in the table (omitting OLS due to its large negative values).

- First row reports entire pooled sample
  - OLS model produces negative R^2_{OOS}
  - Restricting OLS to a sparse parametrization generates substantial improvements
  - Linear model via dimension reduction (PLS / PCR) further improves predictions and compete with boosted trees / RF
  - GLM with group lasso fails to improve as spline functions may not be capturing interaction among features
  - NNs are the best performing nonlinear method and best predictor overall as they capture complex predictor interactions, but benefits of "deep" learning may be limited to 3 layers…
- 2nd and 3rd rows breakout predictability for large and small stocks subsamples
  - General patterns carry
  - Tree methods and NNs are especially successful among large stocks

Empirical study of US equities: the cross section of individual stocks

- **Statistical significance of differences among models at the monthly frequency**

Table 3: Comparison of Monthly Out-of-Sample Prediction using Diebold-Mariano Tests

	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
OLS+H	**3.26***	**3.29***	**3.35***	**3.29***	**3.28***	**3.29***	**3.26***	**3.34***	**3.40***	**3.38***	**3.37***	**3.38***
OLS-3+H		1.42	**1.87**	-0.27	0.62	**1.64**	1.28	1.25	**2.13**	**2.13**	**2.36**	**2.11**
PLS			-0.19	-1.18	-1.47	0.87	0.67	0.63	1.32	1.37	**1.66**	1.08
PCR				-1.10	-1.37	0.85	0.75	0.58	1.17	1.19	1.34	1.00
ENet+H					0.64	**1.90**	1.40	**1.73**	**1.97**	**2.07**	**1.98**	**1.85**
GLM+H						**1.76**	1.22	1.29	**2.28**	**2.17**	**2.68***	**2.37**
RF							0.07	-0.03	0.31	0.37	0.34	0.00
GBRT+H								-0.06	0.16	0.21	0.17	-0.04
NN1									0.56	0.59	0.45	0.04
NN2										0.32	-0.03	-0.88
NN3											-0.32	-0.92
NN4												-1.04

Note: This table reports pairwise Diebold-Mariano test statistics comparing the out-of-sample stock-level prediction performance among thirteen models. Positive numbers indicate the column model outperforms the row model. Bold font indicates the difference is significant at 5% level or better for individual tests, while an asterisk indicates significance at the 5% level for 12-way comparisons via our conservative Bonferroni adjustment.

  - Constrained linear models produce statistically significant improvements over the unconstrained OLS model, although there's little difference in the performance of penalized linear methods and dimension reduction methods
  - Tree methods uniformly improve over linear models, but improvements are marginally significant
  - **NNs are the only models that product large and significant statistical improvements over linear and generalized linear models**

Empirical study of US equities: which covariates matter?

- **Models are in close agreement regarding the most influential stock-level predictors**, which can be grouped into four categories:

1 Recent price trends
- ST reversal
- Stock momentum
- Momentum change
- Industry momentum
- Recent maximum return
- LT reversal

2 Liquidity variables
- Turnover and turnover volatility
- Log market equity
- Dollar volume
- Amihud illiquidity
- Number of zero trading days
- Bid-ask spread

3 Risk measures
- Total and idiosyncratic return volatility
- Market beta
- Beta-squared

4 Valuation ratios and fundamental signals
- Earnings to price
- Sales to price
- Asset growth
- Number of earnings increases

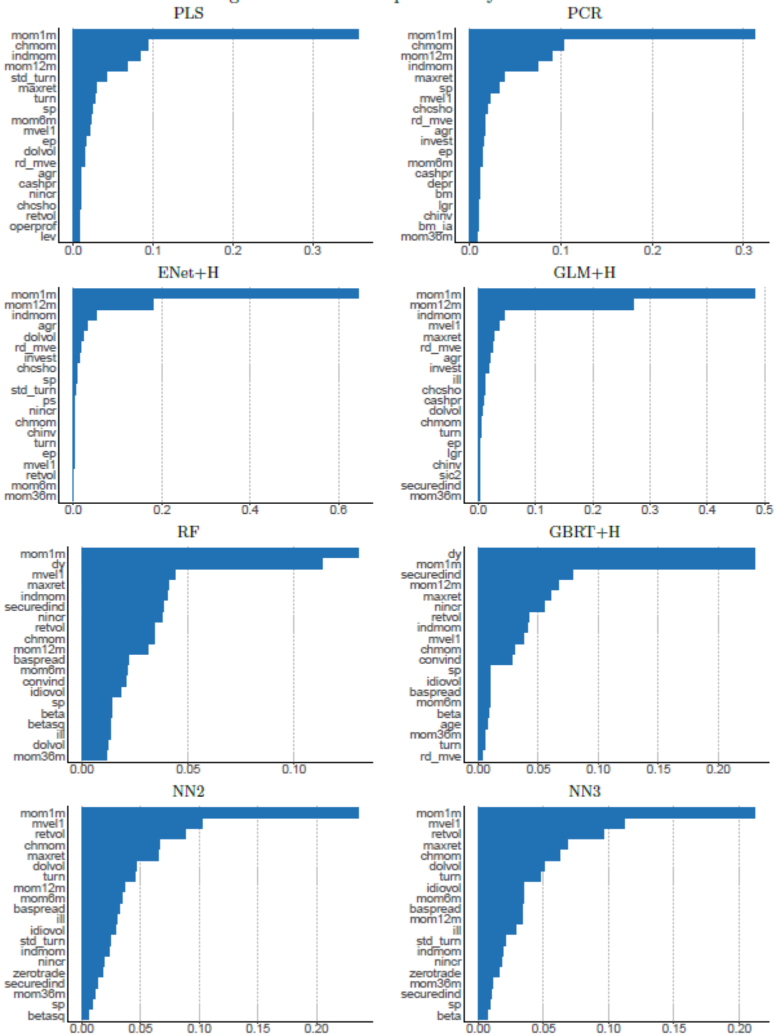
Momentum in Hou, Xue and Zhang (2018)

Value & Investment

Gu, S., Kelly, B., Xiu, D. Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, Volume 33, Issue 5, May 2020, Pages 2223–2273.

Empirical study of US equities: which covariates matter?

Figure 4: Variable Importance By Model

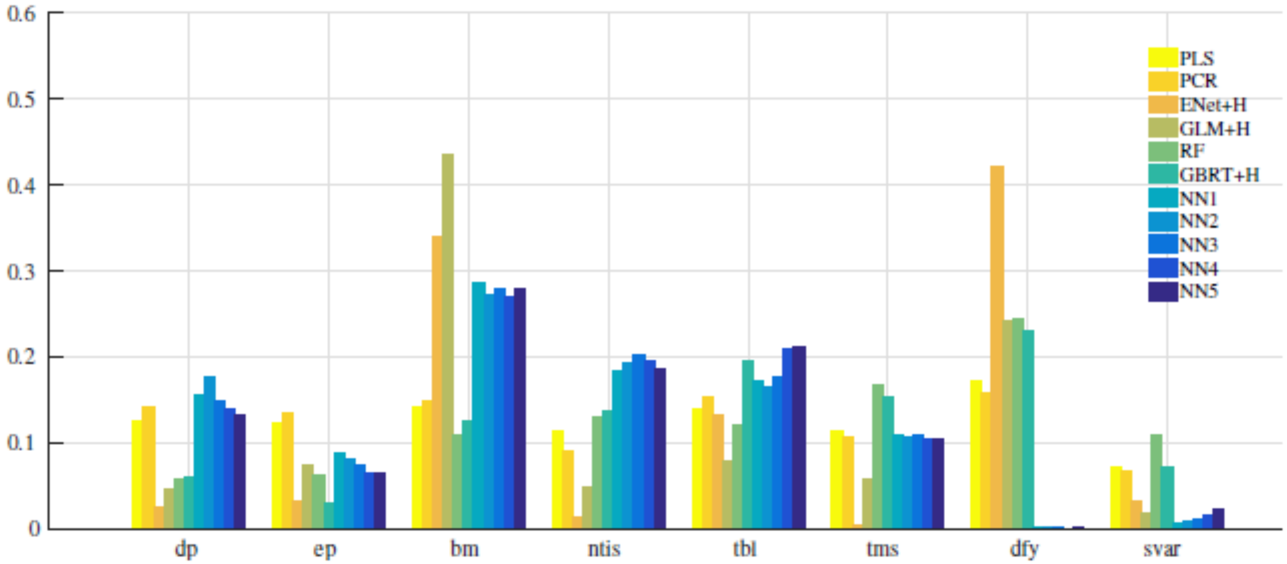


Note: Variable importance for the top 20 most influential variables in each model. Variable importance is an average over all training samples. Variable importances within each model are normalized to sum to one.

- **Reports importance of the top 20 stock-level characteristics for each method**
  - Indicators importance magnitude for penalized linear models and dimension reduction models are highly skewed toward momentum and reversal
  - **Trees and neural networks are more democratic, drawing predictive information from a broader set of characteristics**

Empirical study of US equities: which covariates matter?

- Shows the R^2**-based importance measure for each predictor variable** (again normalized to sum to one within a given model).



Note: Variable importance for eight macroeconomic variables in each model. Variable importance is an average over all training samples. Variable importances within each model are normalized to sum to one. The lower panel provides a complementary visual comparison of macroeconomic variable importances.

  - All models agree that the aggregate book-to-market ratio is a critical predictor, whereas market volatility has little role in any model.
  - Linear and generalized linear models strongly favor bond market variables including the default spread and treasury rate. Nonlinear methods (trees and neural networks) place great emphasis on exactly those predictors ignored by linear methods, such as term spreads and issuance activity

Gu, S., Kelly, B., Xiu, D. Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, Volume 33, Issue 5, May 2020, Pages 2223–2273.

Empirical study of US equities: portfolio level performance

Table 5: Monthly Portfolio-level Out-of-Sample Predictive R^2

	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
S&P 500	-0.22	-0.86	-1.55	0.75	0.71	1.37	1.40	1.08	1.13	1.80	1.63	1.17
Big Value	0.10	0.00	-0.33	0.25	0.59	1.31	1.06	0.85	0.87	1.46	1.21	0.99
Big Growth	-0.33	-1.26	-1.62	0.70	0.51	1.32	1.19	1.00	1.10	1.50	1.24	1.11
Big Neutral	-0.17	-1.09	-1.51	0.80	0.36	1.31	1.28	1.43	1.24	1.70	1.81	1.40
Small Value	0.30	1.66	1.05	0.64	0.85	1.24	0.52	1.59	1.37	1.54	1.40	1.30
Small Growth	-0.16	0.14	-0.18	-0.33	-0.12	0.71	1.24	0.05	0.42	0.48	0.41	0.50
Small Neutral	-0.27	0.60	0.19	0.21	0.28	0.88	0.36	0.58	0.62	0.70	0.58	0.68
Big Conservative	-0.57	-0.10	-1.06	1.02	0.46	1.11	0.55	1.15	1.13	1.59	1.37	1.07
Big Aggressive	0.20	-0.80	-1.15	0.30	0.67	1.75	2.00	1.33	1.51	1.78	1.55	1.42
Big Neutral	-0.29	-1.75	-1.96	0.83	0.48	1.13	0.77	0.85	0.85	1.51	1.45	1.16
Small Conservative	-0.05	1.17	0.71	-0.02	0.34	0.96	0.56	0.82	0.87	0.96	0.90	0.83
Small Aggressive	-0.10	0.51	0.01	-0.09	0.14	1.00	1.46	0.34	0.64	0.75	0.62	0.71
Small Neutral	-0.30	0.45	0.12	0.42	0.35	0.76	-0.01	0.70	0.69	0.83	0.66	0.72
Big Robust	-1.02	-1.08	-2.06	0.55	0.35	1.10	0.33	0.74	0.79	1.28	1.03	0.74
Big Weak	-0.12	1.42	1.07	0.89	1.10	1.33	1.77	1.79	1.79	2.05	1.66	1.60
Big Neutral	0.86	-1.22	-1.26	0.41	0.13	1.10	0.91	0.84	0.94	1.19	1.15	0.99
Small Robust	-0.71	0.35	-0.38	-0.04	-0.42	0.70	0.19	0.24	0.50	0.63	0.53	0.55
Small Weak	0.05	1.06	0.59	-0.13	0.44	1.05	1.42	0.71	0.92	0.99	0.90	0.89
Small Neutral	-0.51	0.07	-0.47	-0.33	-0.32	0.60	-0.08	0.10	0.25	0.38	0.32	0.41
Big Up	0.20	-0.25	-1.24	0.66	1.17	1.18	0.90	0.80	0.76	1.13	1.12	0.93
Big Down	-1.54	-1.63	-1.55	0.44	-0.33	1.14	0.71	0.36	0.70	1.07	0.90	0.84
Big Medium	-0.04	-1.51	-1.94	0.81	-0.08	1.57	1.80	1.29	1.32	1.71	1.55	1.23
Small Up	0.07	0.78	0.56	-0.07	0.25	0.62	-0.03	0.06	0.07	0.21	0.19	0.25
Small Down	-0.21	0.15	-0.20	0.15	-0.01	1.51	1.38	0.74	0.82	1.02	0.91	0.96
Small Medium	0.07	0.82	0.20	0.59	0.37	1.22	1.06	1.09	1.09	1.18	1.00	1.03
SMB	0.81	2.09	0.39	1.72	2.36	0.57	0.35	1.40	1.16	1.31	1.20	1.27
HML	0.66	0.50	1.21	0.46	0.84	0.98	0.21	1.22	1.31	1.06	1.25	1.24
RMW	-2.35	1.19	0.41	-1.07	-0.06	-0.54	-0.92	0.68	0.47	0.84	0.53	0.54
CMA	0.80	-0.44	0.03	-1.07	1.24	-0.11	-1.04	1.88	1.60	1.06	1.84	1.31
UMD	-0.90	-1.09	-0.47	0.47	-0.37	1.37	-0.25	-0.56	-0.26	0.19	0.27	0.35

Note: In this table, we report the out-of-sample predictive R^2s for 25 portfolios using OLS with size, book-to-market, and momentum, OLS-3, PLS, PCR, elastic net (ENet), generalized linear model with group lasso (GLM), random forest (RF), gradient boosted regression trees (GBRT), and five architectures of neural networks (NN1,...,NN5), respectively. "+H" indicates the use of Huber loss instead of the l_2 loss. The 25 portfolios are 3×2 size double-sorted portfolios used in the construction of the Fama-French value, investment, and profitability factors, as well as momentum.

- Form **bottom-up forecasts** for 25 of the most well known portfolios in the empirical finance literature, including the S&P 500, six size and value portfolios, six size and investment portfolios, six size and profitability portfolios, and six size and momentum portfolios. All of these portfolios are long-only and therefore have substantial overlap with the market index. We also study long-short size, value, profitability, investment, and momentum factor portfolios (SMB, HML, RMW, CMA, UMD)
- Reports monthly R^2_{OOS} over our 30-year testing sample
- Regularized linear methods fail to outperform naïve constant forecasts of the S&P500
- **All nonlinear models have substantial positive predictive performance**
- Patterns in S&P500 forecasting performance across models carry over to long-only characteristic-sorted portfolios and long-short factor portfolios. **Nonlinear methods excel**

Gu, S., Kelly, B., Xiu, D. Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, Volume 33, Issue 5, May 2020, Pages 2223–2273.

Empirical study of US equities: portfolio level performance

Table A.7: Implied Sharpe Ratio Improvements

	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
S&P 500	-	-	-	0.08	0.08	0.14	0.15	0.11	0.12	0.20	0.17	0.12
Big Value	0.05	0.00	-	0.03	0.07	0.14	0.12	0.09	0.10	0.15	0.13	0.11
Big Growth	-	-	-	0.08	0.06	0.14	0.13	0.11	0.12	0.16	0.13	0.12
Big Neutral	0.01	-	-	0.08	0.04	0.13	0.13	0.15	0.13	0.17	0.18	0.14
Small Value	0.02	0.15	0.10	0.06	0.08	0.11	0.05	0.14	0.13	0.14	0.13	0.12
Small Growth	-	0.03	-	-	-	0.14	0.21	0.01	0.09	0.10	0.08	0.10
Small Neutral	-	0.06	0.02	0.02	0.03	0.09	0.04	0.06	0.06	0.07	0.06	0.07
Big Conservative	-	-	-	0.09	0.04	0.10	0.05	0.11	0.10	0.14	0.12	0.10
Big Aggressive	-	-	-	0.04	0.09	0.20	0.23	0.16	0.18	0.21	0.18	0.17
Big Neutral	-	-	-	0.08	0.05	0.11	0.08	0.08	0.08	0.14	0.14	0.11
Small Conservative	-	0.12	0.08	0.00	0.04	0.10	0.06	0.09	0.09	0.10	0.09	0.09
Small Aggressive	-	0.09	0.00	-	0.03	0.16	0.22	0.06	0.11	0.12	0.10	0.12
Small Neutral	-	0.04	0.01	0.04	0.03	0.07	0.00	0.06	0.06	0.08	0.06	0.07
Big Robust	-	-	-	0.06	0.04	0.11	0.03	0.08	0.08	0.13	0.10	0.08
Big Weak	0.03	0.15	0.12	0.10	0.12	0.14	0.19	0.19	0.19	0.21	0.17	0.17
Big Neutral	-	-	-	0.06	0.02	0.14	0.12	0.11	0.13	0.15	0.15	0.13
Small Robust	-	0.04	-	0.00	-	0.07	0.02	0.02	0.05	0.06	0.05	0.05
Small Weak	0.04	0.17	0.11	-	0.08	0.17	0.22	0.13	0.15	0.16	0.15	0.15
Small Neutral	-	0.01	-	-	-	0.06	-	0.01	0.03	0.04	0.03	0.04
Big Up	-	-	-	0.06	0.11	0.11	0.08	0.07	0.07	0.10	0.10	0.09
Big Down	-	-	-	0.05	-	0.13	0.08	0.04	0.08	0.12	0.10	0.10
Big Medium	-	-	-	0.13	-	0.22	0.25	0.19	0.20	0.24	0.22	0.18
Small Up	-	0.08	0.06	-	0.03	0.07	0.00	0.01	0.01	0.02	0.02	0.03
Small Down	-	0.03	-	0.03	0.00	0.23	0.22	0.13	0.14	0.17	0.15	0.16
Small Medium	0.01	0.08	0.02	0.06	0.04	0.12	0.11	0.11	0.11	0.12	0.10	0.10

Note: Improvement in annualized Sharpe ratio ($SR^{*} - SR$) implied by the full sample Sharpe ratio of each portfolio together with machine learning predictive R^2_{oos} from Table 5. Cases with negative R^2_{oos} imply a Sharpe ratio deterioration and are omitted.

- Form **bottom-up forecasts** of S&P 500, six size and value portfolios, six size and investment portfolios, six size and profitability portfolios, and six size and momentum portfolios, plus long-short SMB, HML, RMW, CMA, UMD
- Reports **improvement in annualized Sharpe ratio** for an investor exploiting ML predictions for portfolio timing in order to **assess the economic magnitudes of portfolio predictability**
- NN3 for example improves the buy-and-hold Sharpe ratio of the S&P500 by 0.20
- For characteristic-based portfolios, **nonlinear ML help improve Sharpe ratios** by anywhere from few percentage points to over 24 percentage points

Empirical study of US equities: portfolio level performance

Table 6: Market Timing Sharpe Ratio Gains

	OLS-3 +H	PLS	PCR	ENet +H	GLM +H	RF	GBRT +H	NN1	NN2	NN3	NN4	NN5
S&P 500	0.07	0.05	-	0.12	0.19	0.18	0.19	0.22	0.20	0.26	0.22	0.19
Big Value	-	0.06	-	0.09	0.06	0.09	0.08	0.11	0.11	0.13	0.10	0.11
Big Growth	0.08	-	-	0.10	0.17	0.20	0.21	0.22	0.20	0.26	0.22	0.21
Big Neutral	0.06	0.03	-	0.11	0.16	0.13	0.17	0.23	0.21	0.23	0.23	0.21
Small Value	-	0.15	0.09	0.01	0.08	0.07	0.08	0.11	0.11	0.10	0.11	0.13
Small Growth	0.00	0.03	-	-	-	0.04	0.05	0.02	0.03	0.03	0.02	0.02
Small Neutral	0.02	0.09	0.05	0.03	0.04	0.11	0.11	0.09	0.08	0.10	0.09	0.11
Big Conservative	0.08	0.02	-	0.08	0.15	0.09	0.13	0.17	0.14	0.19	0.16	0.14
Big Aggressive	0.08	-	-	0.01	0.13	0.22	0.18	0.21	0.19	0.23	0.20	0.20
Big Neutral	0.04	-	-	0.09	0.11	0.09	0.11	0.13	0.12	0.18	0.18	0.16
Small Conservative	0.04	0.17	0.12	0.02	0.05	0.17	0.15	0.11	0.11	0.14	0.13	0.15
Small Aggressive	0.01	0.05	-	-	-	0.08	0.06	0.02	0.05	0.06	0.04	0.05
Small Neutral	0.01	0.06	0.03	0.01	0.04	0.08	0.09	0.07	0.06	0.08	0.07	0.09
Big Robust	0.10	0.07	-	0.11	0.18	0.17	0.18	0.18	0.16	0.22	0.19	0.16
Big Weak	0.05	0.12	0.05	0.09	0.12	0.21	0.17	0.22	0.20	0.21	0.18	0.19
Big Neutral	0.09	0.00	-	0.09	0.20	0.19	0.17	0.22	0.21	0.24	0.21	0.20
Small Robust	0.09	0.04	-	0.00	0.00	0.10	0.07	0.04	0.05	0.08	0.08	0.08
Small Weak	-0.03	0.09	0.00	-	-	0.07	0.07	0.06	0.06	0.06	0.05	0.06
Small Neutral	0.04	0.04	-	0.00	0.01	0.11	0.09	0.04	0.04	0.07	0.07	0.08
Big Up	0.10	0.05	-	0.10	0.21	0.16	0.14	0.17	0.14	0.17	0.18	0.17
Big Down	-	0.09	-	-	0.02	0.08	0.10	0.10	0.07	0.12	0.11	0.09
Big Medium	-	0.04	-	0.14	0.09	0.17	0.20	0.22	0.21	0.25	0.22	0.19
Small Up	0.08	0.13	0.10	0.05	0.07	0.16	0.12	0.07	0.06	0.08	0.07	0.10
Small Down	-	0.04	-	-	-	0.06	0.04	0.01	0.01	0.02	0.01	0.01
Small Medium	0.05	0.11	0.07	0.08	0.09	0.13	0.15	0.13	0.12	0.14	0.13	0.15
SMB	0.06	0.17	0.09	0.24	0.26	0.00	-	0.21	0.18	0.15	0.09	0.11
HML	0.00	0.01	0.04	-	-	0.04	0.02	0.04	0.06	0.04	0.02	0.01
RMW	-	-	-	-	-	-	-	-	-	0.01	0.01	-
CMA	0.02	0.02	0.00	-	-	0.08	-	0.00	0.01	0.05	0.04	0.06
UMD	-	-	-	-	-	-	-	-	-	-	-	-

Note: Improvement in annualized Sharpe ratio ($SR^* - SR$). We compute the SR^* by weighting the portfolios based on a market timing strategy Campbell and Thompson (2007). Cases with Sharpe ratio deterioration are omitted.

- Form **bottom-up forecasts** of S&P 500, six size and value portfolios, six size and investment portfolios, six size and profitability portfolios, and six size and momentum portfolios, plus long-short SMB, HML, RMW, CMA, UMD
- We can also **assess the economic magnitudes of portfolio predictability with a market timing trading strategy**
- Table reports annualized Sharpe ratio gains (relative to a buy-and-hold strategy) for timing strategies based on ML
- The **strongest and most consistent trading strategies are those based on nonlinear models, with NNs the best overall**

Simulation of 3 factor model

Table A.1: Comparison of Predictive R^2s for Machine Learning Algorithms in Simulations

Model	(a)				(b)			
Parameter	$P_c = 50$		$P_c = 100$		$P_c = 50$		$P_c = 100$	
R^2(%)	IS	OOS	IS	OOS	IS	OOS	IS	OOS
OLS	7.50	1.14	8.19	-1.35	3.44	-4.72	4.39	-7.75
OLS+H	7.48	1.25	8.16	-1.15	3.43	-4.60	4.36	-7.54
PCR	2.69	0.90	1.70	0.43	0.65	0.02	0.41	-0.01
PLS	6.24	3.48	6.19	2.82	1.02	-0.08	0.99	-0.17
Lasso	6.04	4.26	6.08	4.25	1.36	0.58	1.36	0.61
Lasso+H	6.00	4.26	6.03	4.25	1.32	0.59	1.31	0.61
Ridge	6.46	3.89	6.67	3.39	1.66	0.34	1.76	0.23
Ridge+H	6.42	3.91	6.61	3.42	1.63	0.35	1.73	0.25
ENet	6.04	4.26	6.08	4.25	1.35	0.58	1.35	0.61
ENet+H	6.00	4.26	6.03	4.25	1.32	0.59	1.31	0.61
GLM	5.91	4.11	5.94	4.08	3.38	1.22	3.31	1.17
GLM+H	5.85	4.12	5.88	4.09	3.32	1.24	3.24	1.20
RF	8.34	3.35	8.23	3.30	8.05	3.07	8.22	3.02
GBRT	7.08	3.35	7.02	3.33	6.51	2.76	6.42	2.84
GBRT+H	7.16	3.45	7.11	3.37	6.47	3.12	6.37	3.22
NN1	6.53	4.37	6.72	4.28	5.61	2.78	5.80	2.59
NN2	6.55	4.42	6.72	4.26	6.22	3.13	6.33	2.91
NN3	6.47	4.34	6.67	4.27	6.03	2.96	6.09	2.68
NN4	6.47	4.31	6.66	4.24	5.94	2.81	6.04	2.51
NN5	6.41	4.27	6.55	4.14	5.81	2.72	5.70	2.20
Oracle	6.22	5.52	6.22	5.52	5.86	5.40	5.86	5.40

Note: In this table, we report the average in-sample (IS) and out-of-sample (OOS) R^2 for models (*a*) and (*b*) using Ridge, Lasso, Elastic Net (ENet), generalized linear model with group lasso (GLM), random forest (RF), gradient boosted regression trees (GBRT), and five architectures of neural networks (NN1,...,NN5), respectively. "+H" indicates the use of Huber loss instead of the l_2 loss. "Oracle" stands for using the true covariates in a pooled-OLS regression. We fix $N = 200$, $T = 180$, and $P_x = 2$, comparing $P_c = 100$ with $P_c = 50$. The number of Monte Carlo repetitions is 100.

- **Simulated data from 2 different data generating processes** that produce data from a high dimensional predictor set
  - (a) individual predictors enter only linearly and additively
  - (b) predictors can enter through nonlinear transformations and via pairwise interactions
- Applied ML repertoire to the simulated datasets
- **Reports monthly R_S^2** both IS and OOS **for each model and each method over 50 and 100 MC repetitions**
  - **Linear and generalized linear methods dominate in the linear and uninteracted setting**
  - **Tree-based methods and neural networks significantly outperform in the the nonlinear and interactive setting**

Finally, an AI-powered ETF

- AIEQ ETF, powered by IBM Watson

- How EquBot is beating the market with AIEQ, the AI-powered ETF
  - https://www.ibm.com/blog/equbot-aieq-ai-powered-etf/

ChatGPT

The AI-powered ETF (exchange-traded fund) "AIEQ" (AI Powered Equity ETF) selects stocks through a process driven by artificial intelligence algorithms. Here's a simplified overview of how it works:

1. **Data Collection**: AIEQ's AI system gathers vast amounts of data from various sources, including financial statements, market news, social media, and other relevant information.
2. **Data Processing**: The AI algorithms analyze the collected data to identify patterns, trends, and correlations. They look for factors that may indicate potential opportunities for investment.
3. **Machine Learning**: AIEQ utilizes machine learning techniques to continuously improve its investment strategy. It learns from historical data and adjusts its models based on changing market conditions and the performance of past investment decisions.
4. **Stock Selection**: Based on the analysis and insights generated by the AI algorithms, AIEQ selects a portfolio of stocks that it believes have the potential to outperform the market. The selection process may involve various criteria, such as fundamental indicators, technical analysis, sentiment analysis, and risk management considerations.
5. **Portfolio Management**: AIEQ's portfolio is dynamically managed, meaning that the composition of stocks may change over time as market conditions and the AI's analysis evolve. The AI algorithms continuously monitor the performance of the portfolio and make adjustments as needed to optimize returns and manage risk.

Overall, AIEQ's approach relies on the power of artificial intelligence to sift through vast amounts of data, identify investment opportunities, and make data-driven decisions in real-time.

Elaborate on machine learning techniques used by AIEQ

ChatGPT

AIEQ utilizes various machine learning techniques to enhance its investment strategy and continuously improve its performance. Here are some of the key machine learning techniques that AIEQ may employ:

1. **Supervised Learning**: AIEQ may use supervised learning algorithms to train models on historical data with known outcomes. For example, it can train models to predict stock price movements or classify stocks into different categories based on their performance. By learning from past data, the algorithms can identify patterns and relationships that help inform investment decisions.
2. **Unsupervised Learning**: Unsupervised learning techniques enable AIEQ to discover hidden patterns or structures within the data without explicit supervision. For instance, clustering algorithms can group stocks based on similarities in their characteristics or price movements, allowing AIEQ to identify potential investment opportunities or diversify its portfolio effectively.
3. **Reinforcement Learning**: Reinforcement learning involves training algorithms to make sequential decisions in an environment to maximize cumulative rewards. AIEQ can use reinforcement learning to optimize its trading strategies over time by learning from the outcomes of previous decisions. For example, the algorithm may adjust its investment allocations based on the performance of past trades and market conditions.
4. **Deep Learning**: Deep learning, a subset of machine learning, involves training artificial neural networks with multiple layers to learn complex patterns and representations from data. AIEQ may utilize deep learning techniques, such as deep neural networks, to analyze unstructured data like news articles, social media sentiment, or alternative data sources. These models can extract valuable insights and signals that inform investment decisions.
5. **Natural Language Processing (NLP)**: NLP techniques enable AIEQ to analyze and understand textual data, such as financial news, earnings reports, or social media discussions. AIEQ's algorithms can extract sentiment, identify key events or trends, and assess the impact of news on stock prices. By incorporating NLP, AIEQ can react quickly to new information and adjust its investment strategy accordingly.
6. **Ensemble Learning**: Ensemble learning combines multiple machine learning models to improve predictive accuracy and robustness. AIEQ may employ ensemble methods, such as random forests or gradient boosting, to aggregate predictions from different algorithms and sources. By leveraging the diversity of individual models, AIEQ can make more accurate forecasts and reduce the risk of relying on any single model.

Overall, AIEQ's use of machine learning techniques allows it to analyze vast amounts of data, extract valuable insights, and make informed investment decisions in real-time. These techniques enable AIEQ to adapt to changing market conditions, optimize its portfolio, and potentially outperform traditional investment strategies.

Performance vs S&P 500

HOME > AIEQ • NYSEARCA

Amplify Ai Powered Equity ETF

$36.19 ↑ 39.60% +10.27 5 Y

13 Mar, 18:30:00 UTC-4 · USD · NYSEARCA · Disclaimer

1 D | 5 D | 1 M | 6 M | YTD | 1 Y | 5 Y | MAX

Amplify Ai Powere...	$36.19	+$10.27	↑ 39.60%
S&P 500	5,165.31	+2,342.83	↑ 83.01%

Lecture 5 Exercise

- Compare different neural network methods vs. linear regression and select based on OOS R2 performance and portfolio performance.