



X ROBECO

Seminar Group C

Presentation Overview

01 Introduction

02 Methodology

03 Results

04 Conclusion



Introduction

Improving the Prediction of Cross-Sectional Government Bond Returns through ML



Methodology

01 Models

Which models are we using?

OLS (All), OLS (4 factor), Elastic Net, GLM
RF, XGB, SVM, NN



02 Variable Selection

How did we select variables?

Which variables were selected?



03 Feature Importance



How are we planning to measure feature importance?

→ Dive more in the section Results

04 Training methods



Which 5 ways of training and predicting?

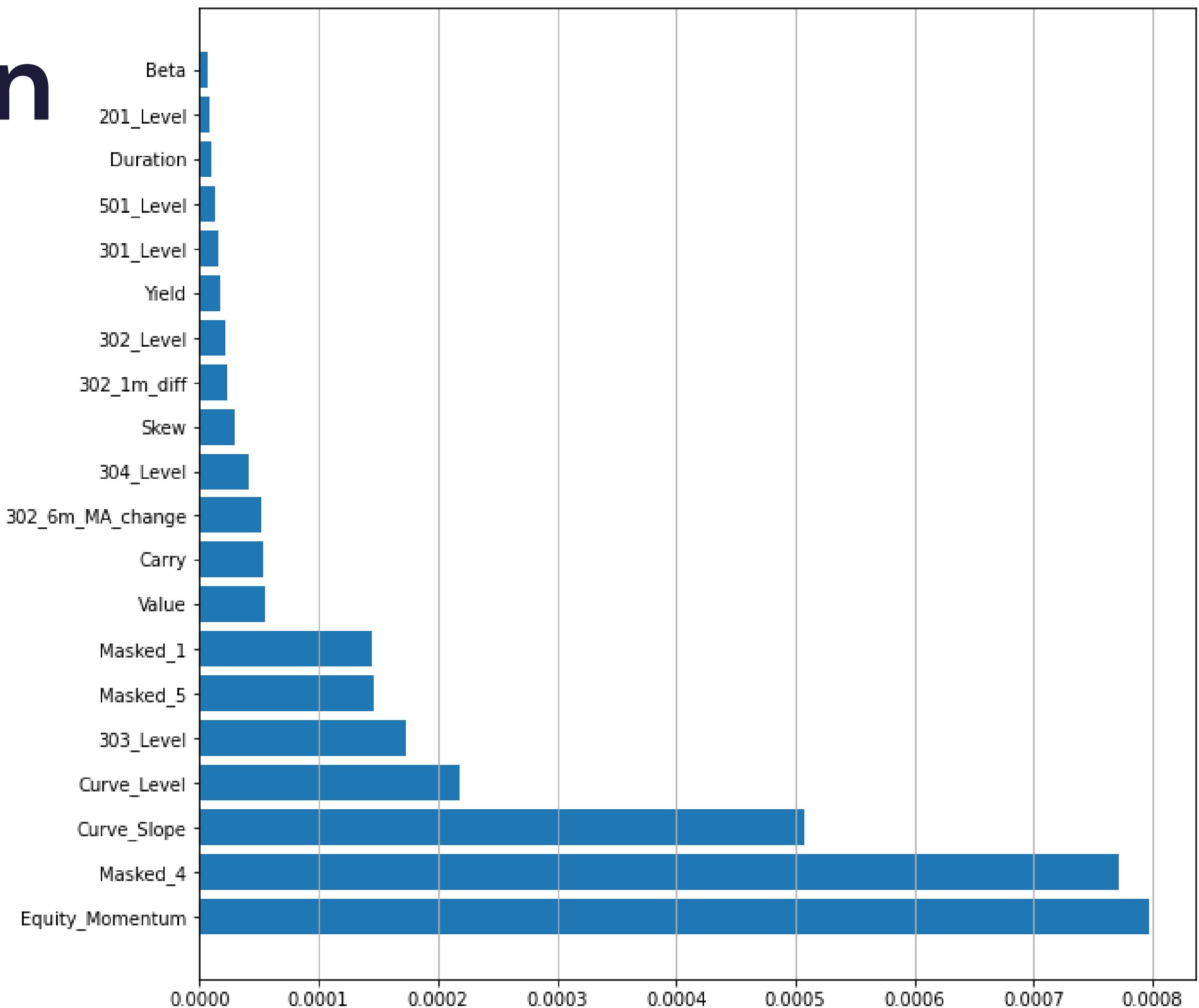


Variable selection

The most important predictors according to elastic net regularization (lambda parameter)

Top 20 features

3 Masked variables



5 different training methods

Static



Training

30%

Validation

50%

Test

Dynamic 4x



Training

30%

Validation

50%

Rehypertune?

Test



— Methodology (Training Methods)

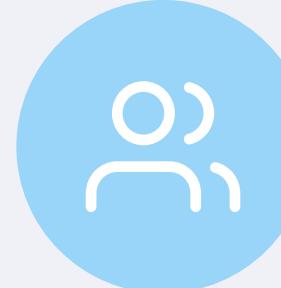
Results

We discuss results in 3 sections



R²

Model evaluation



VIM

Feature Importance



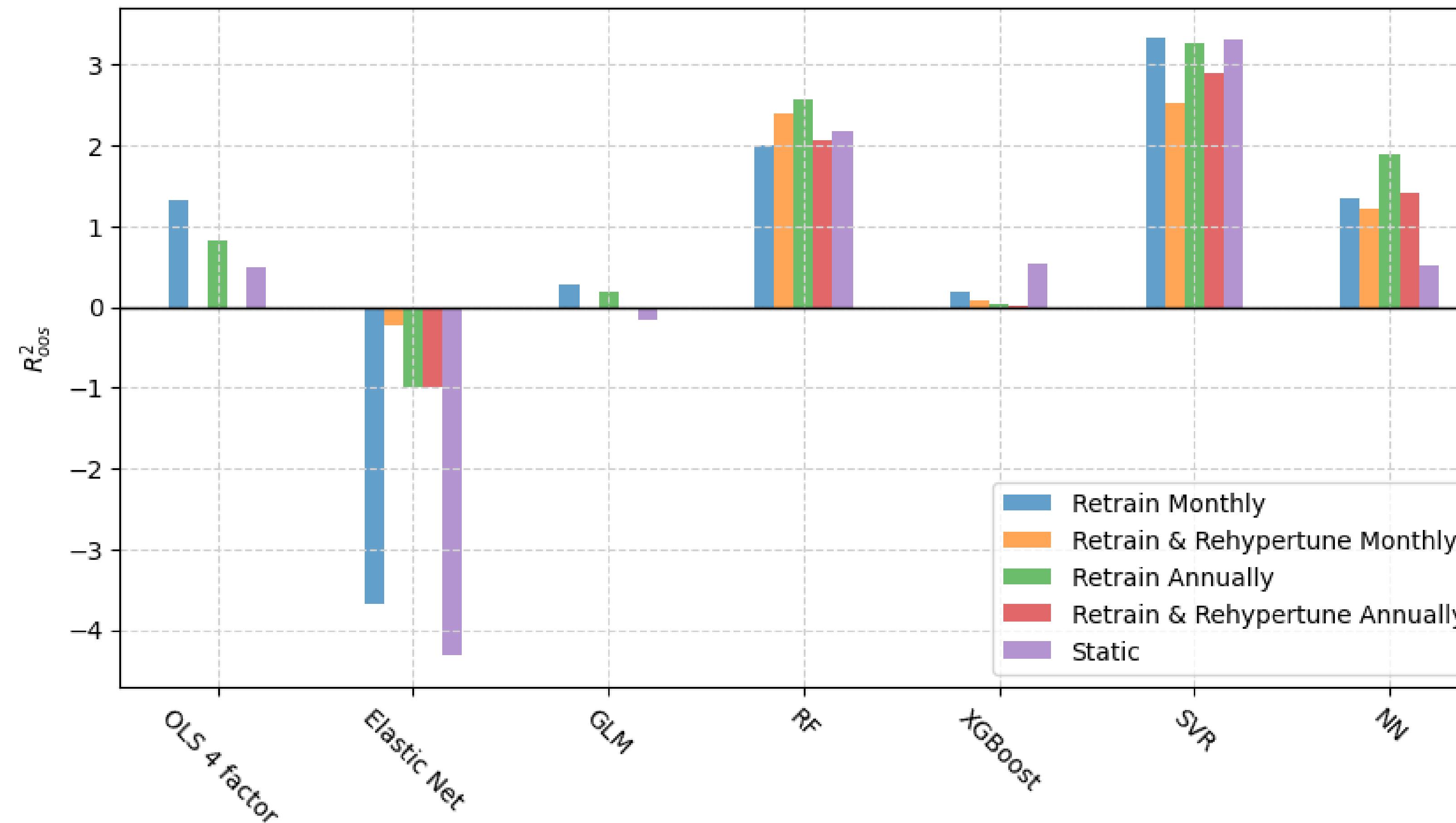
Profit?

Portfolio formation



— Results

4 different training methods



Model Evaluation

Evaluated all ML-models using R-squared

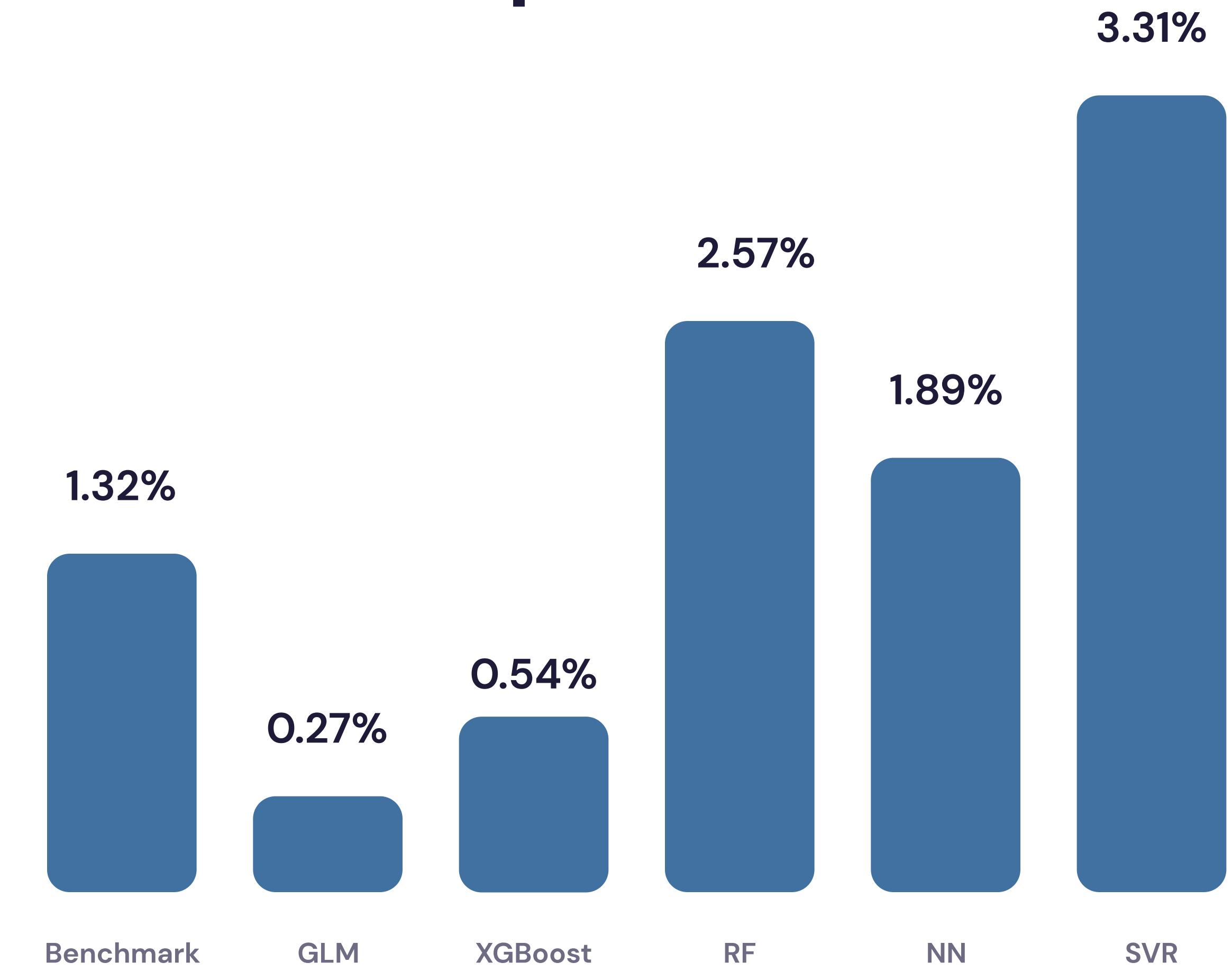


— Results (Model Evaluation)



Best performing

Model comparison

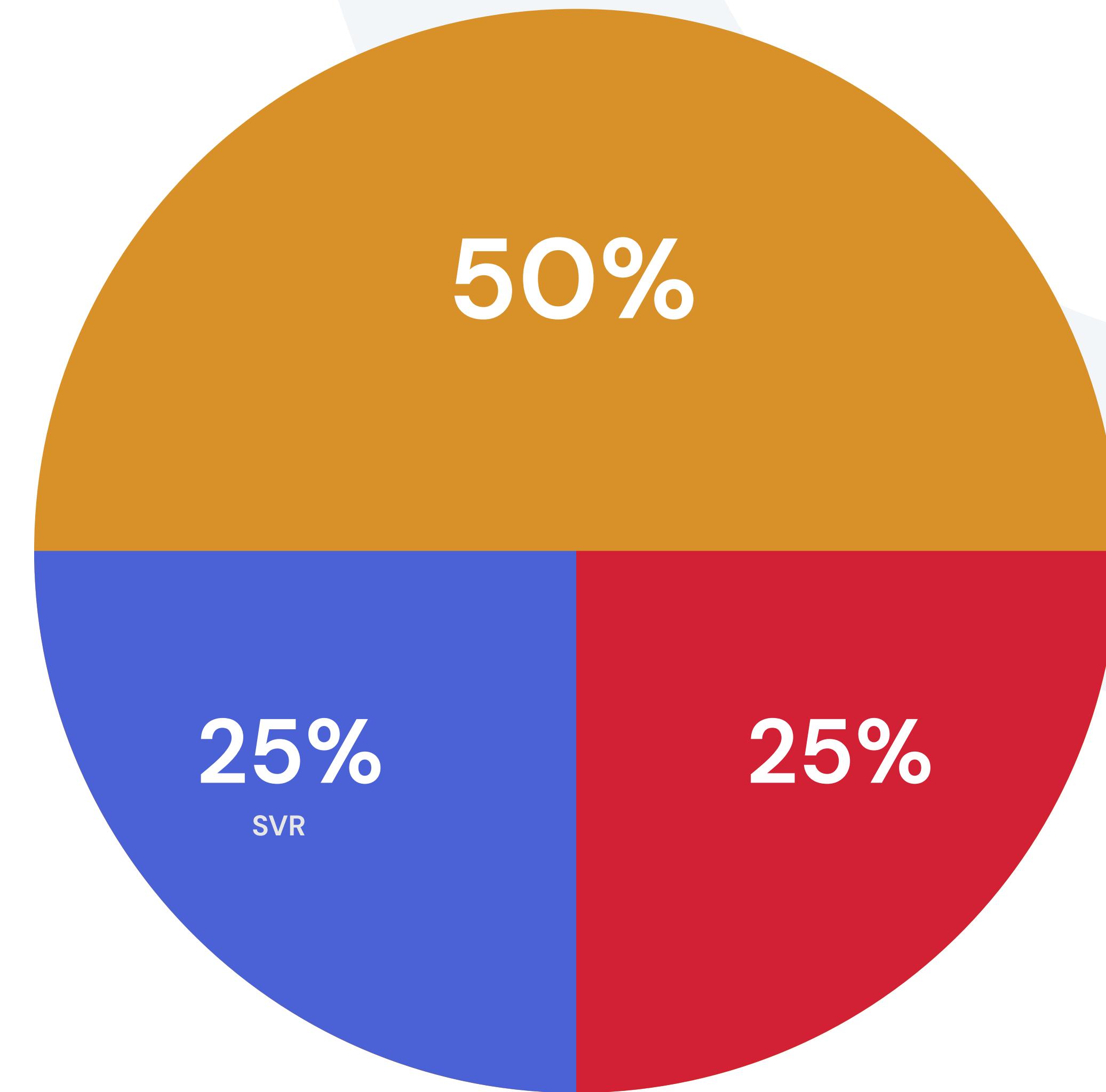


— Results (Model comparison)

Training methods comparison

Among the 4 non-linear models

- Retraining annually
- Retraining monthly
- Static



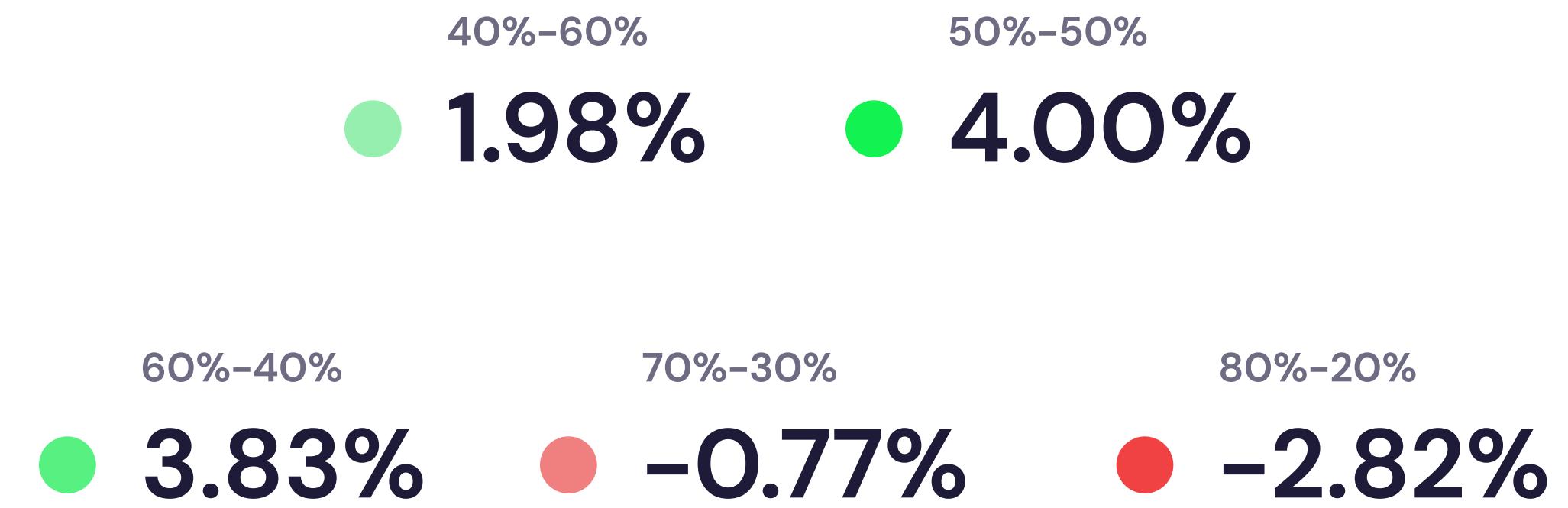
— Results (Model comparison)



*Note: This result was not carried over the other results.

Robustness testing

Robustness testing using 10 different sample splits and bigger hyperparameter grid, where the difference in result was only caused by changing the test set.



Diebold-Mario tests

In order to evaluate the significance of a model relative to its peers, the Diebold-Mariano test is performed on the predictions from monthly retraining and no re-hypertuning.

Table 2: Diebold-Mariano tests between pairs of models

	OLS 4 factor	Elastic Net	GLM	RF	XGBoost	SVR	NN
OLS All	1.98	1.93	2.00	1.99	1.98	2.01	1.99
OLS 4		-1.60	0.64	0.30	-0.11	0.87	0.30
Elastic Net			1.55	1.28	0.91	1.76	1.48
GLM				-0.74	-1.55	1.04	-1.12
RF					-2.46	1.50	-0.24
XGBoost						2.18	0.66
SVR							-1.47

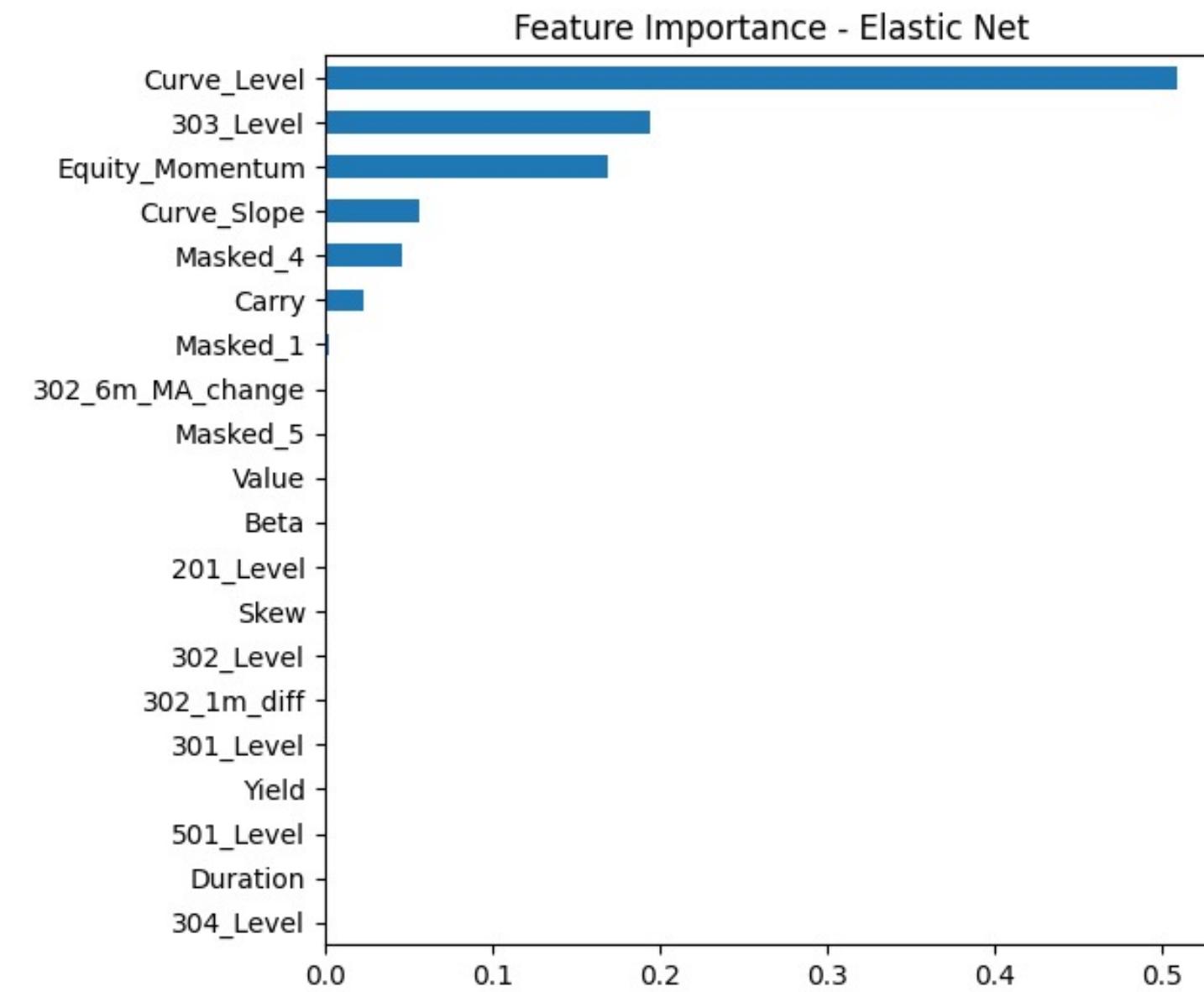
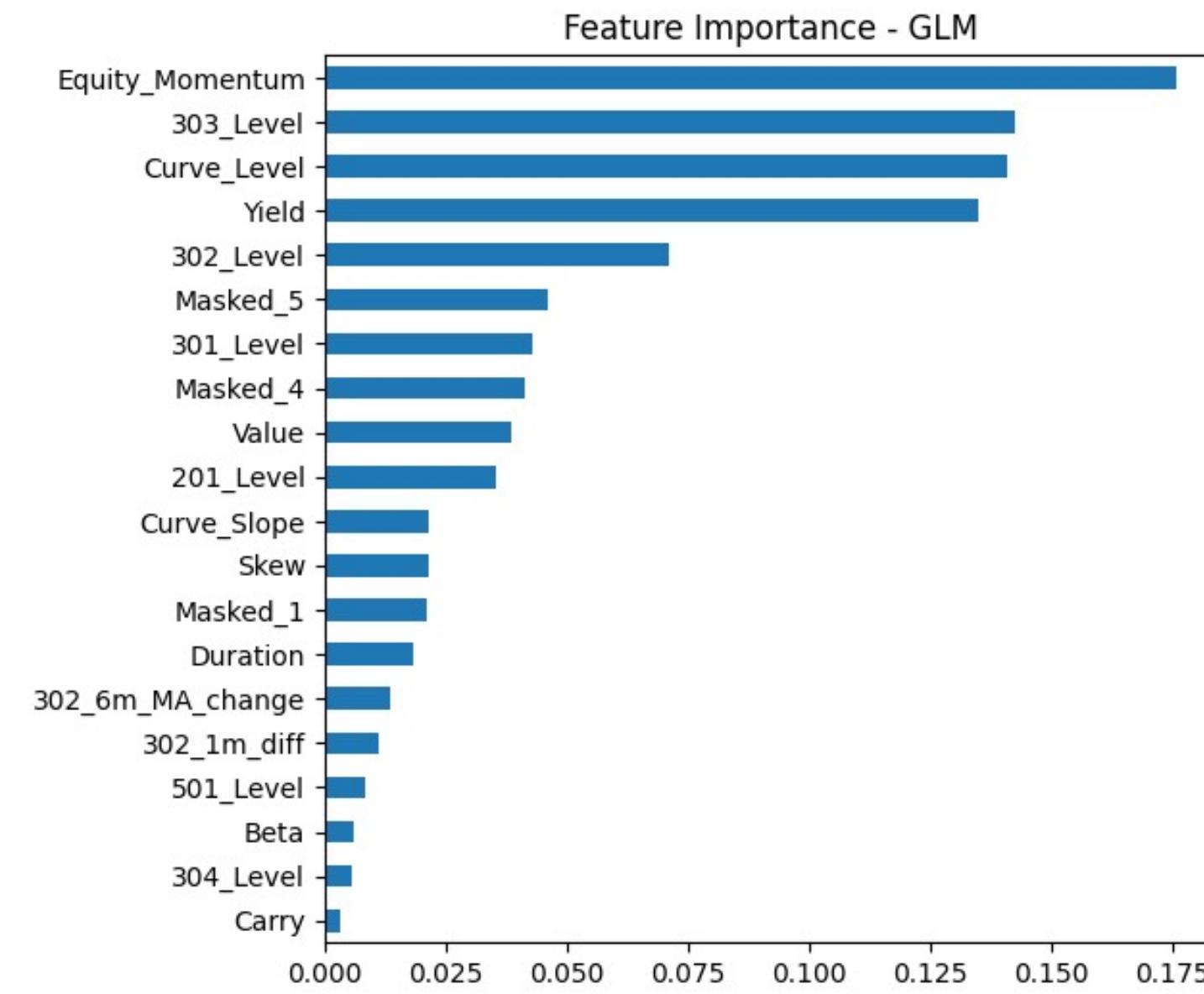
Note: The table reports the Diebold-Mariano (adjusted) test statistic between pairs of models. Positive values indicate that the model in the column outperforms the one in the row. Bold font indicates the value is significant at a 5% level.



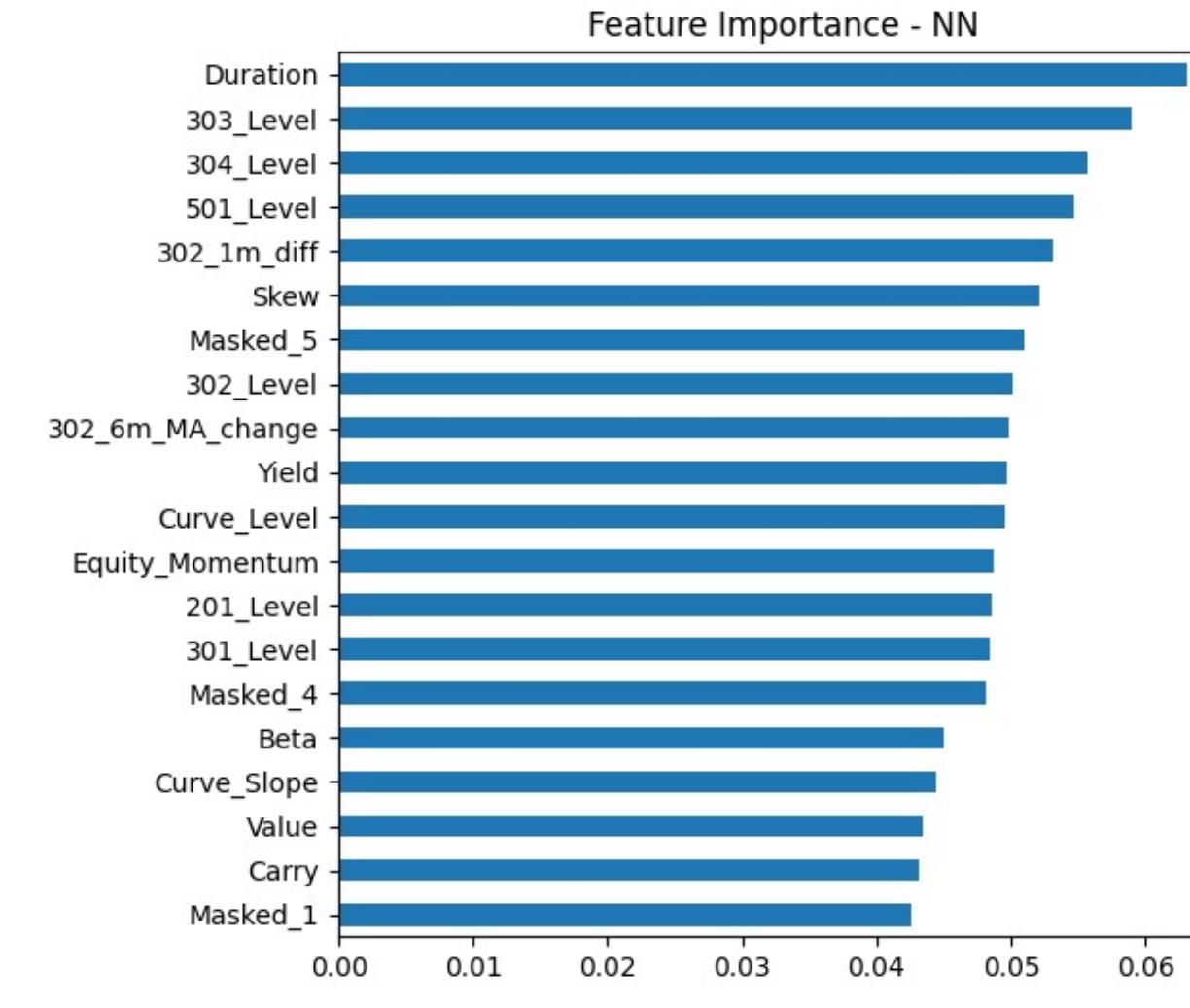
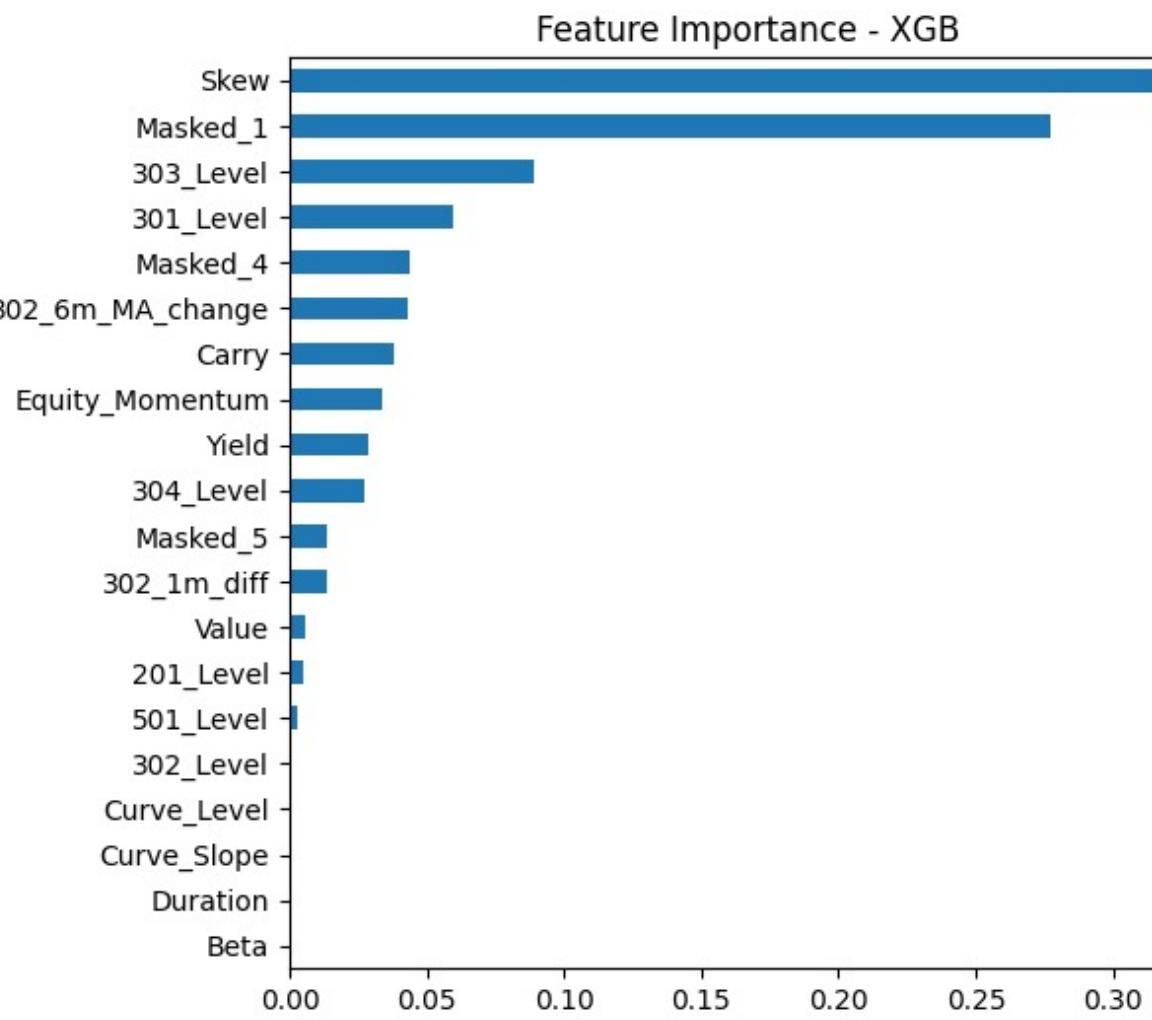
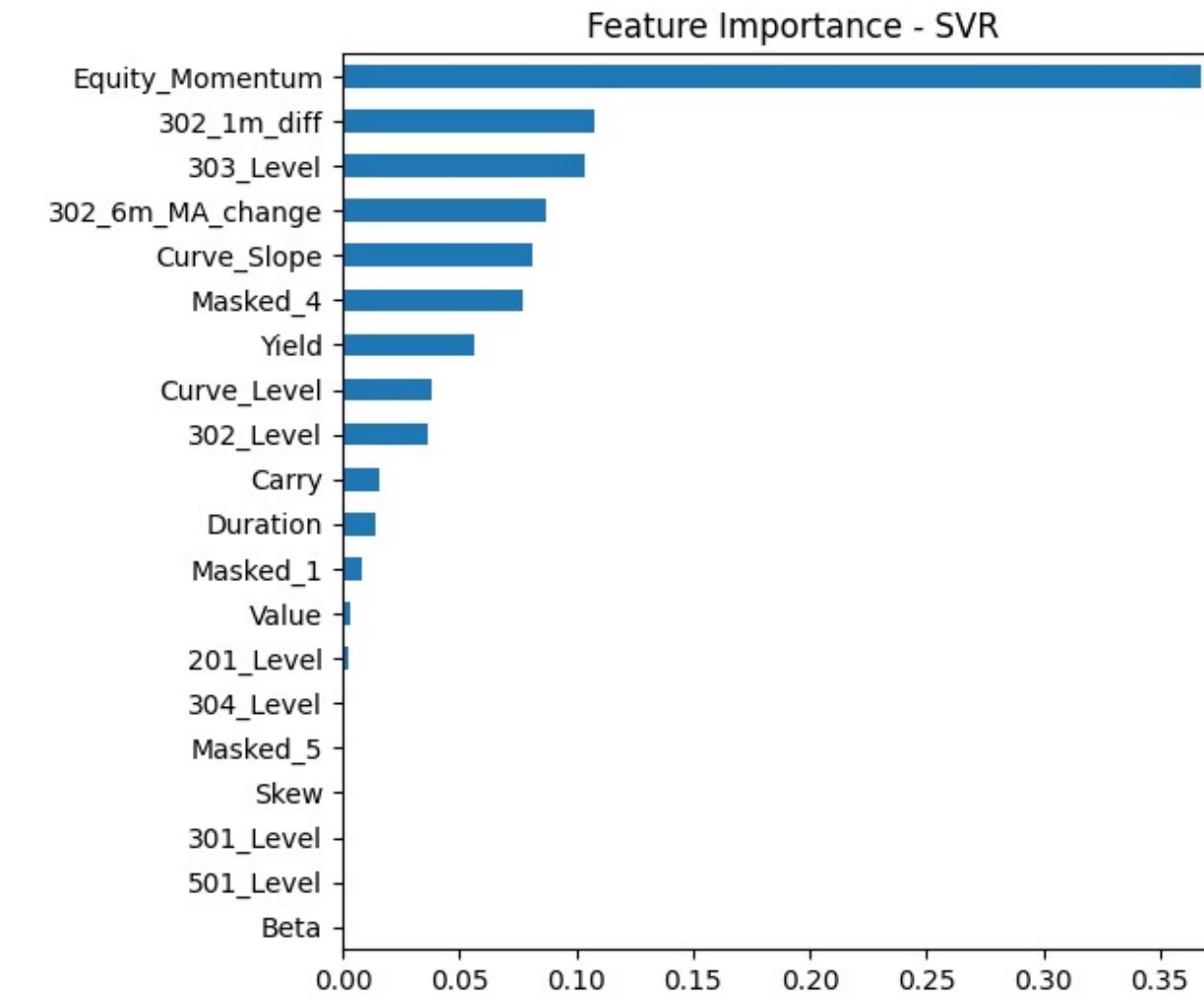
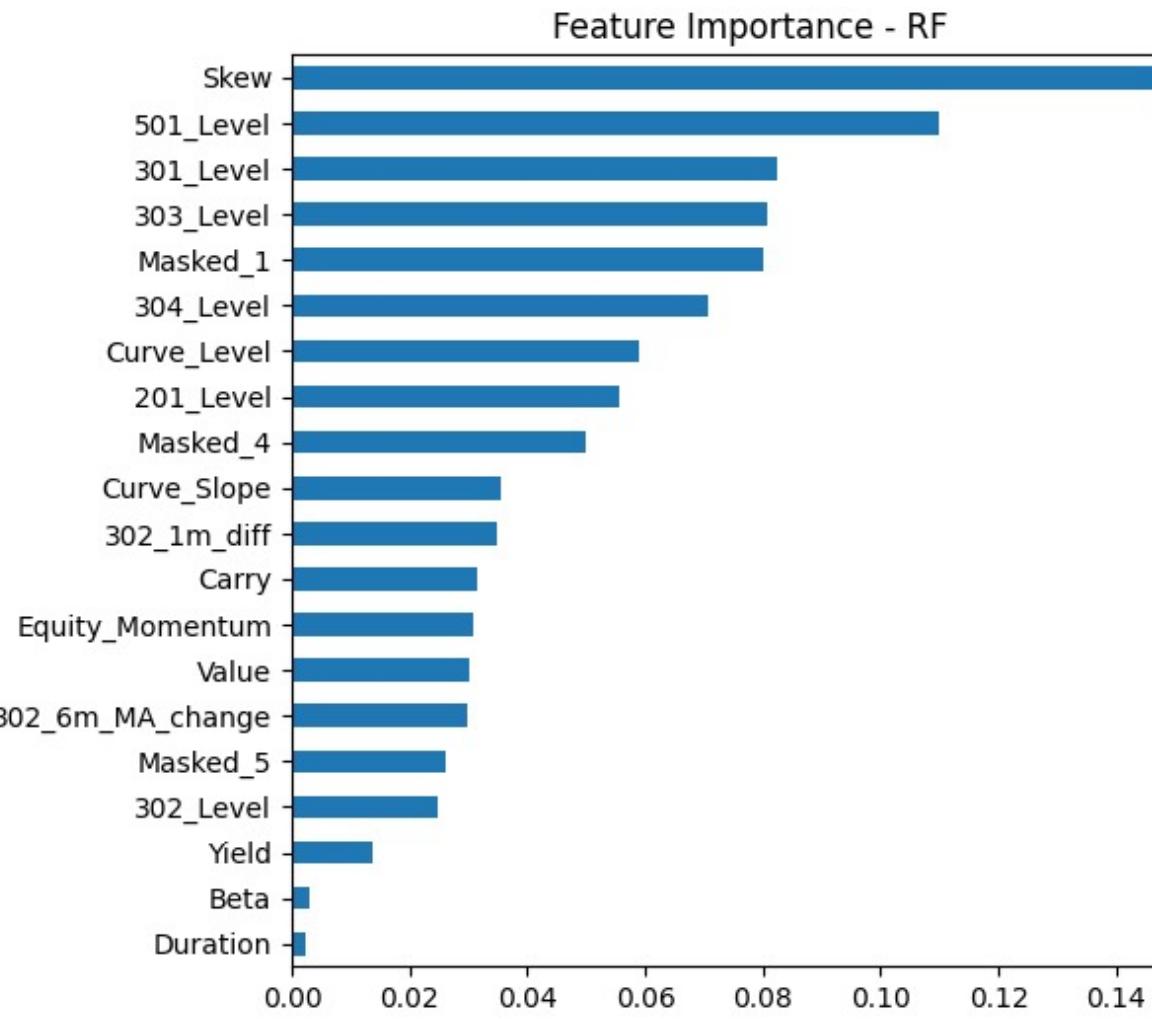
Feature Importance

Method

1. After fitting a model, the respective in-sample R² is computed
2. For each feature, the same metric is computed after setting the respective feature equal to its mean.



Feature Importance



Portfolio

- OOS forecasting period: May 2006 – May 2022 (17 years)
- Equal weight portfolio
 - strong empirical properties
 - monthly rebalance
- 30-40-30 split; total 35 bonds cross-sectionally
- Long-short portfolio (H-L); net zero investment
- Performance mis-alignment (individual vs portfolio)
- Portfolio performance influenced by cross-correlations: Gu et. al. (2020)
- 73% of OOS correlations > 0.5
- Top tier credit rated nations (AAA, AA+ etc.)
- Typical government bond characteristics observed

Table 2: Performance of Machine Learning Portfolios

	OLS-MF				OLS-All			
	Pred	Avg	Std	SR	Pred	Avg	Std	SR
Low 30% (L)	0.01	0.03	0.13	0.69	-0.11	0.03	0.16	0.56
Mid 40%	0.02	0.03	0.16	0.62	0.02	0.03	0.15	0.62
High 30% (H)	0.03	0.03	0.19	0.59	0.14	0.04	0.19	0.67
H - L	0.03	0.01	0.12	0.17	0.25	0.01	0.12	0.30
	ENet				GLM			
	Pred	Avg	Std	SR	Pred	Avg	Std	SR
Low 30% (L)	-0.05	0.02	0.11	0.51	0.04	0.02	0.13	0.45
Mid 40%	-0.03	0.03	0.17	0.67	0.05	0.03	0.17	0.66
High 30% (H)	-0.01	0.04	0.20	0.62	0.05	0.04	0.19	0.69
H - L	-0.03	0.03	0.17	0.67	0.01	0.02	0.12	0.56
	RF				XGBoost			
	Pred	Avg	Std	SR	Pred	Avg	Std	SR
Low 30% (L)	0.04	0.02	0.16	0.54	0.04	0.03	0.16	0.60
Mid 40%	0.05	0.03	0.16	0.64	0.06	0.03	0.16	0.64
High 30% (H)	0.07	0.03	0.17	0.70	0.08	0.03	0.16	0.64
H - L	0.03	0.01	0.10	0.29	0.05	0.00	0.11	0.04
	SVR				NN			
	Pred	Avg	Std	SR	Pred	Avg	Std	SR
Low 30% (L)	0.03	0.03	0.16	0.56	0.00	0.03	0.14	0.65
Mid 40%	0.04	0.03	0.16	0.70	0.04	0.03	0.16	0.63
High 30% (H)	0.05	0.03	0.16	0.61	0.07	0.03	0.17	0.62
H - L	0.02	0.00	0.10	0.10	0.07	0.01	0.10	0.17



Separate long-short

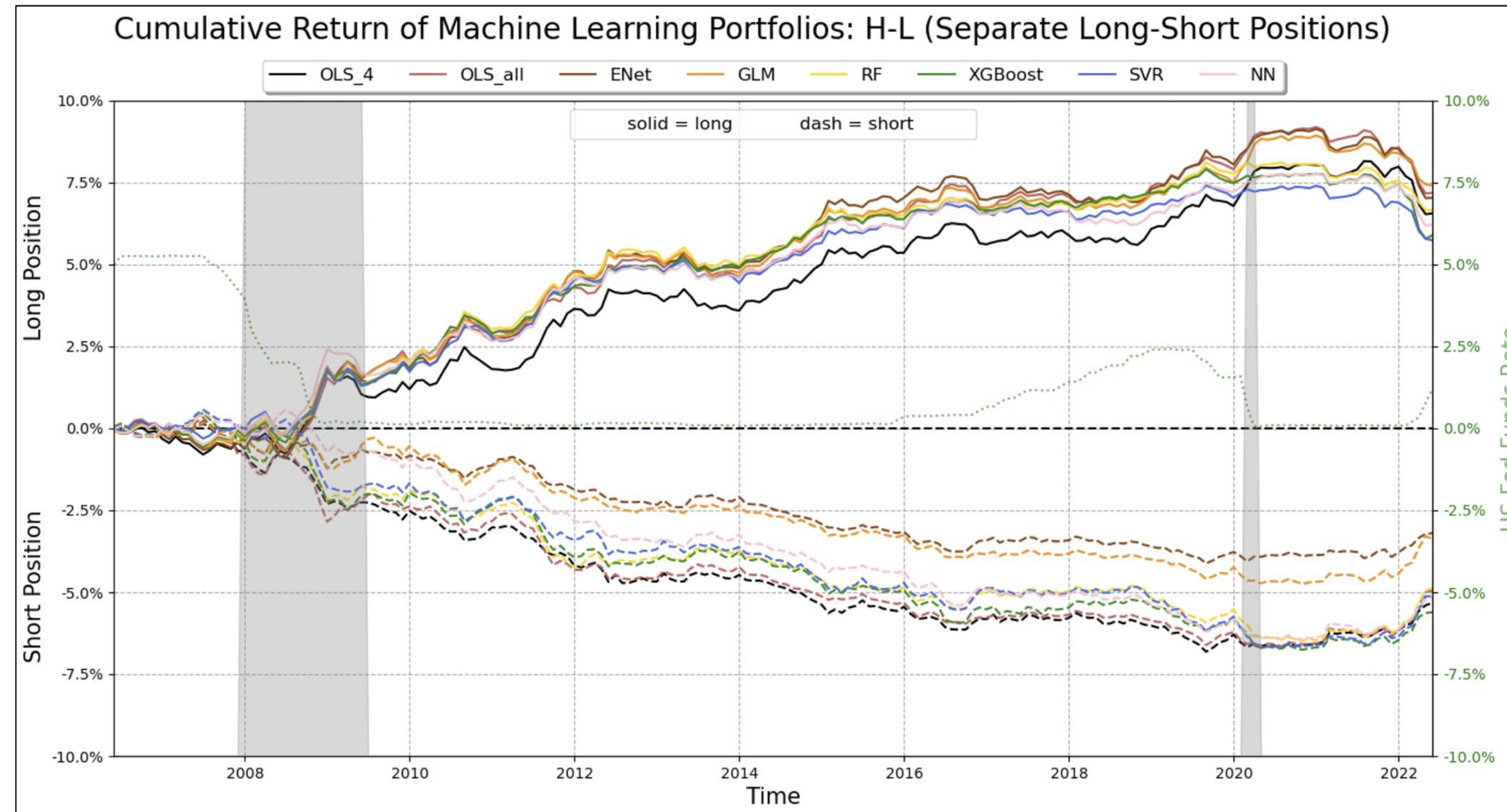
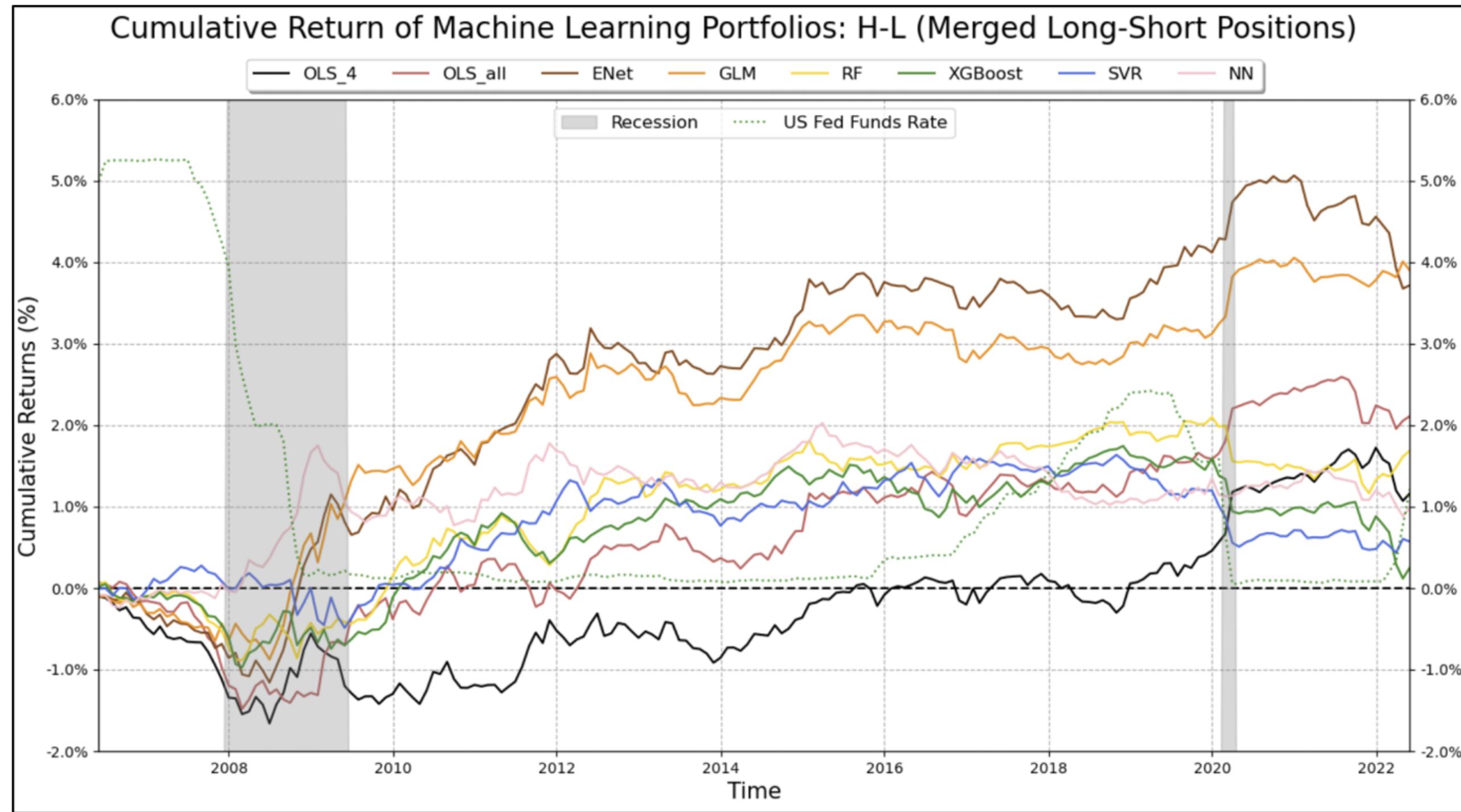
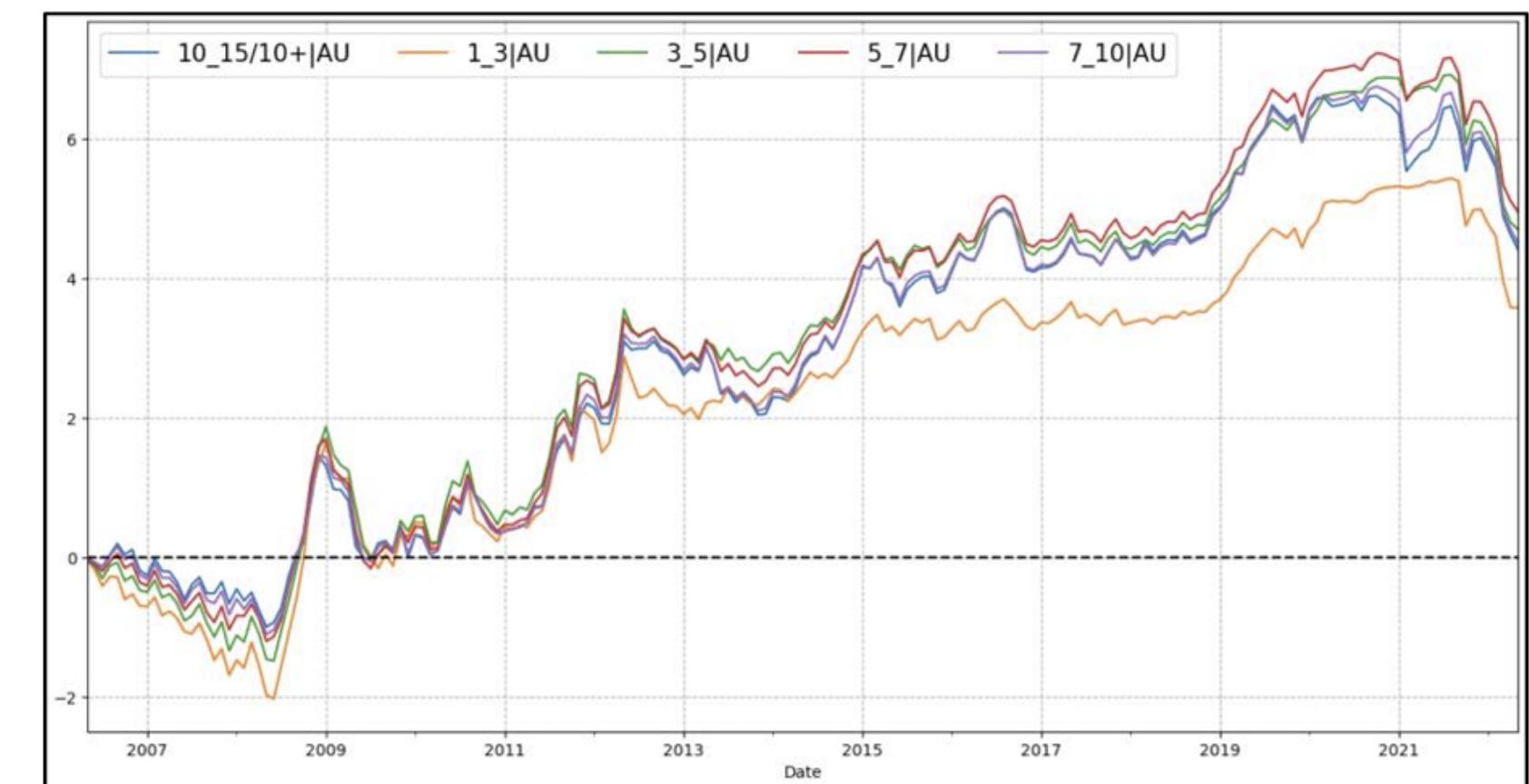


Table 3: Key Performance Metrics of Long-Short Machine Learning (Bond) Portfolios (Equal Weighted)

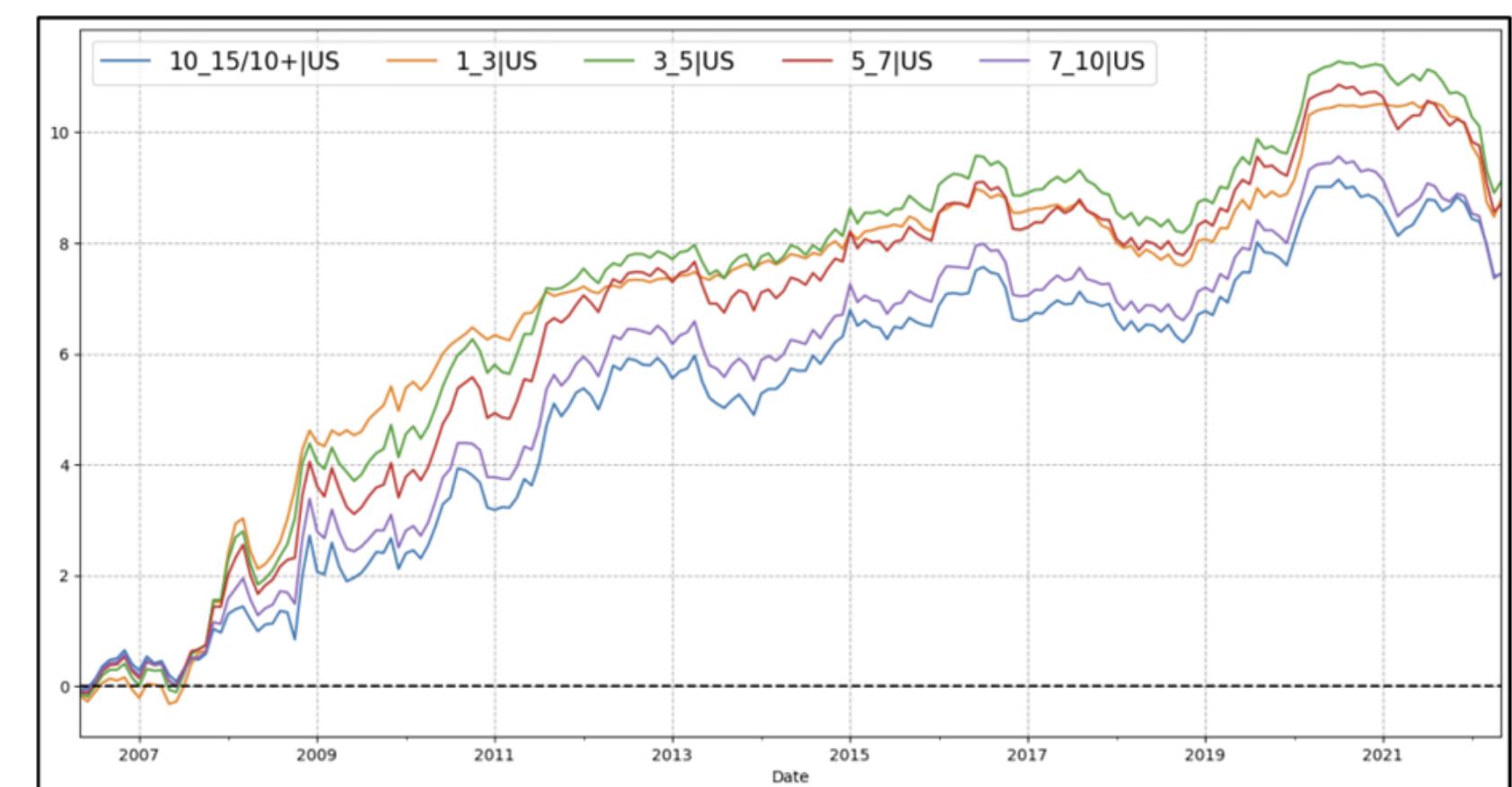
	OLS-MF	OLS-All	ENet	GLM	RF	XGBoost	SVR	NN
Max DD (%)	1.56	1.56	1.39	0.78	0.96	1.62	1.20	1.14
Max 1M Loss (%)	0.33	0.37	0.41	0.35	0.43	0.40	0.42	0.35
Turnover (%)	41.9	111.5	59.3	71.3	139.3	138.8	118.1	100.1
R^2_{oos}	-0.05	-9.92	0.01	-0.01	-0.11	-0.21	-0.01	-0.40
Information Ratio	0.17	0.30	0.48	0.56	0.29	0.04	0.10	0.17
Appraisal Ratio	-0.69	0.55	0.30	1.08	0.92	0.28	0.31	-0.25



Actual return evolution OOS (AUS)



Actual return evolution OOS (US)





SKILLS



X ROBECO

Thank you! Questions?