Discussion of

“Machine Learning Mutual Fund Flows”

by Fausch, Frigg, Ruenzi, and Weigert.

Daniele Bianchi
Queen Mary, University of London

2025 FMA Consortium on Asset Management

Overview of the paper

Goal: Forecasting fund flows based on observable characteristics $\mathbf{z}_{i,t}$:

$$\widehat{\text{flow}}_{i,t+h} = \mathbf{z}_{i,t}'\widehat{\boldsymbol{\theta}}$$

Overview of the paper

Goal: Forecasting fund flows based on observable characteristics $\mathbf{z}_{i,t}$:

$$\widehat{\text{flow}}_{i,t+h} = \mathbf{z}'_{i,t}\widehat{\boldsymbol{\theta}}$$

Why? Predicting flows is key for fund managers and investors:

$\hookrightarrow$ Liquidity management (e.g., optimize your cash reserves, reduce needs for costly credit lines).

$\hookrightarrow$ Capital efficiency (e.g., timing investments around expected inflows).

Overview of the paper

Goal: Forecasting fund flows based on observable characteristics $\mathbf{z}_{i,t}$:

$$\widehat{\text{flow}}_{i,t+h} = \mathbf{z}_{i,t}'\widehat{\boldsymbol{\theta}}$$

Why? Predicting flows is key for fund managers and investors:

↪ Liquidity management (e.g., optimize your cash reserves, reduce needs for costly credit lines).

↪ Capital efficiency (e.g., timing investments around expected inflows).

Issues:

↪ $\mathbf{z}_{i,t}$ can be large (variable selection, risk of overfitting).

↪ Is $\mathbf{z}_{i,t}'\widehat{\boldsymbol{\theta}}$, i.e., linearity, a reasonable assumption?

Overview of the paper

Main idea: the investors' decision process might be non-linear, i.e.,

$$\widehat{\text{flow}}_{i,t+h} = \cancel{\mathbf{z}'_{i,t}\widehat{\boldsymbol{\theta}}} \quad \widehat{f(\mathbf{z}_{i,t})}$$

$\hookrightarrow$ A non-linear $\widehat{f(\mathbf{z}_{i,t})}$ produces more accurate $\widehat{\text{flow}}_{i,t+h}$.

Overview of the paper

Main idea: the investors' decision process might be non-linear, i.e.,

$$\widehat{\text{flow}}_{i,t+h} = \cancel{\mathbf{z}_{i,t}'\widehat{\boldsymbol{\theta}}} \quad \widehat{f(\mathbf{z}_{i,t})}$$

$\hookrightarrow$ A non-linear $\widehat{f(\mathbf{z}_{i,t})}$ produces more accurate $\widehat{\text{flow}}_{i,t+h}$.

Focus on interpretability, i.e., $\forall j = 1, \ldots, p$

$$\frac{\partial\ \widehat{\text{flow}}_{i,t+h}}{\partial\ \widehat{f(z_{ij,t})}}$$

How the expected flow $\widehat{\text{flow}}_{i,t+h}$ changes as a function of $z_{ij,t}$?

My comments:

Comments/observations:

#1: Opening the black box (SHAP values).

#2: Economic evaluation (funds sorting strategy).

Comment #1: Opening the black box.

SHAP values and interpretability

Originally introduced as a game-theoretic tool to assign credit to players in cooperative games by considering:

↪ All possible ways the coalition could have been formed.

↪ Each player's marginal contribution in each scenario.

SHAP values and interpretability

Originally introduced as a game-theoretic tool to assign credit to players in cooperative games by considering:

- ↪ All possible ways the coalition could have been formed.
- ↪ Each player's marginal contribution in each scenario.

Key principles:

- ↪ Look at every possible coalition of players.
- ↪ Calculate the marginal/additional contribution each player brings for a given coalition.
- ↪ Average these marginal contributions across all possible coalitions.

SHAP values and interpretability

Originally introduced as a game-theoretic tool to assign credit to players in cooperative games by considering:

- ↪ All possible ways the coalition could have been formed.
- ↪ Each player's marginal contribution in each scenario.

Key principles:

- ↪ Look at every possible coalition of ~~players~~ predictors.
- ↪ Calculate the marginal/additional contribution each ~~player~~ predictor brings for a given ~~coalition~~ forecast.
- ↪ Average these marginal contributions across all possible ~~coalitions~~ forecasts.

SHAP values and interpretability

Consider two predictors: Expense ratio (ER) and total net assets (TNA).

SHAP values and interpretability

Consider two predictors: Expense ratio (ER) and total net assets (TNA).

Each predictor combination gives a performance (e.g., R^2, MSE),

- ↪ Empty set: $\{\}$ $\Longrightarrow$ $R^2(\{\})$
- ↪ Single predictors: $\{ER\}$, $\{TNA\}$ $\Longrightarrow$ $R^2(\{ER\}), R^2(\{TNA\})$
- ↪ Full set: $\{ER, TNA\}$ $\Longrightarrow$ $R^2(\{ER, TNA\})$

SHAP values and interpretability

Consider two predictors: Expense ratio (ER) and total net assets (TNA).

Each predictor combination gives a performance (e.g., R^2, MSE),

- ↪ Empty set: $\{\}$ $\Longrightarrow$ $R^2(\{\})$
- ↪ Single predictors: $\{ER\}$, $\{TNA\}$ $\Longrightarrow$ $R^2(\{ER\}), R^2(\{TNA\})$
- ↪ Full set: $\{ER, TNA\}$ $\Longrightarrow$ $R^2(\{ER, TNA\})$

The average marginal contribution (SHAP value) for ER and TNA are:

$$\phi_{ER} = \frac{1}{2}\left[R^2(\{ER\}) - R^2(\{\})\right] + \frac{1}{2}\left[R^2(\{ER, TNA\}) - R^2(\{TNA\})\right]$$

SHAP values and interpretability

Consider two predictors: Expense ratio (ER) and total net assets (TNA).

Each predictor combination gives a performance (e.g., R^2, MSE),

$\hookrightarrow$ Empty set: $\{\}$ $\Longrightarrow$ $R^2(\{\})$

$\hookrightarrow$ Single predictors: $\{ER\}$, $\{TNA\}$ $\Longrightarrow$ $R^2(\{ER\}), R^2(\{TNA\})$

$\hookrightarrow$ Full set: $\{ER, TNA\}$ $\Longrightarrow$ $R^2(\{ER, TNA\})$

The average marginal contribution (SHAP value) for ER and TNA are:

$$\phi_{ER} = \frac{1}{2}\left[R^2(\{ER\}) - R^2(\{\})\right] + \frac{1}{2}\left[R^2(\{ER, TNA\}) - R^2(\{TNA\})\right]$$
$$\phi_{TNA} = \frac{1}{2}\left[R^2(\{TNA\}) - R^2(\{\})\right] + \frac{1}{2}\left[R^2(\{ER, TNA\}) - R^2(\{ER\})\right]$$

SHAP values and interpretability

Consider two predictors: Expense ratio (ER) and total net assets (TNA).

Each predictor combination gives a performance (e.g., R^2, MSE),

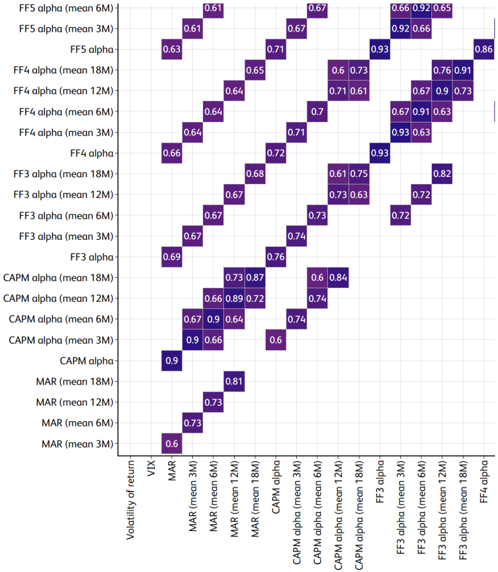
- $\hookrightarrow$ Empty set: $\{\}$ $\Longrightarrow$ 0.01
- $\hookrightarrow$ Single predictors: $\{ER\}$, $\{TNA\}$ $\Longrightarrow$ $0.03, 0.04$
- $\hookrightarrow$ Full set: $\{ER, TNA\}$ $\Longrightarrow$ 0.05

The average marginal contribution (SHAP value) for ER and TNA are:

$$\phi_{ER} = \frac{1}{2}[0.03 - 0.01] + \frac{1}{2}[0.05 - 0.04] = 0.015$$

$$\phi_{TNA} = \frac{1}{2}[0.04 - 0.01] + \frac{1}{2}[0.05 - 0.03] = 0.025 \qquad \phi_{TNA} > \phi_{ER}$$

Comment #1a: SHAP and correlated predictors



Comment #1a: SHAP and correlated predictors

When predictors are correlated, their SHAP values become more ambiguous

↪ The importance gets "shared" between correlated features.

Comment #1a: SHAP and correlated predictors

When predictors are correlated, their SHAP values become more ambiguous

$\hookrightarrow$ The importance gets "shared" between correlated features.

Recall that, for e.g.,:

$$\phi_{ER} = \frac{1}{2}[R^2(\{ER\}) - R^2(\{\})] + \frac{1}{2}[R^2(\{ER, TNA\}) - R^2(\{TNA\})]$$

Comment #1a: SHAP and correlated predictors

When predictors are correlated, their SHAP values become more ambiguous

↪ The importance gets "shared" between correlated features.

Recall that, for e.g.,:

$$\phi_{ER} = \frac{1}{2}[R^2(\{ER\}) - R^2(\{\})] + \frac{1}{2}[R^2(\{ER, TNA\}) - R^2(\{TNA\})]$$

↪ Assume ER and TNA are positively correlated.

↪ $R^2(\{ER, TNA\}) < R^2(\{ER\}) + R^2(\{TNA\})$ (redundant information).

↪ The larger the correlation, the more ϕ_{ER} converges to ϕ_{TNA}.

Comment #1b: SHAP are "true to the model", not "true to the data"

Model-agnostic?

$\hookrightarrow$ SHAP decompose the model prediction, not underlying data relationships.

Comment #1b: SHAP are "true to the model", not "true to the data"

Model-agnostic?

↪ SHAP decompose the model prediction, not underlying data relationships.

"True to the Model":

↪ Explain how the model uses each predictor.

↪ Depends on the model's forecasts.

Comment #1b: SHAP are "true to the model", not "true to the data"

Model-agnostic?

$\hookrightarrow$ SHAP decompose the model prediction, not underlying data relationships.

"True to the Model":

$\hookrightarrow$ Explain how the model uses each predictor.

$\hookrightarrow$ Depends on the model's forecasts.

Not "True to the Data":

$\hookrightarrow$ Don't capture true DGP ($R^2 << 1$).

$\hookrightarrow$ Vary across model choices.

Comment #1b: SHAP are "true to the model", not "true to the data"

Model-agnostic?

$\hookrightarrow$ SHAP decompose the model prediction, not underlying data relationships.

"True to the Model":

$\hookrightarrow$ Explain how the model uses each predictor.

$\hookrightarrow$ Depends on the model's forecasts.

Not "True to the Data":

$\hookrightarrow$ Don't capture true DGP ($R^2 << 1$).

$\hookrightarrow$ Vary across model choices.

E.g., for a linear model $\Longrightarrow \widehat{\text{flow}}_{i,t+1} = \widehat{\beta}_{ER} ER_{i,t} + \widehat{\beta}_{TNA} TNA_{i,t}$,

$\hookrightarrow$ SHAP decompose the linear forecast $\widehat{\beta}_{ER} ER_{i,t} + \widehat{\beta}_{TNA} TNA_{i,t}$. Even if the true relationship is nonlinear.

Comment #1b: SHAP are “true to the model”, not “true to the data”

Model-agnostic?

↪ SHAP decompose the model prediction, not underlying data relationships.

“True to the Model”:

↪ Explain how the model uses each predictor.

↪ Depends on the model’s forecasts.

Not “True to the Data”:

↪ Don’t capture true DGP ($R^2 << 1$).

↪ Vary across model choices.

What’s special about non-linear methods?

↪ It would be interesting to understand if non-linear methods pin down different predictors for fund flows.

Comment #2: Economic evaluation of forecasts.

2a) Sorting forecasts and long-short portfolios

To what extent are these performances really different from each other?

Panel A - top and bottom decile portfolios												
	Top 10%:				Bottom 10%:				Long-Short:			
	FF4	FF5	FF6	FF7	FF4	FF5	FF6	FF7	FF4	FF5	FF6	FF7
Equally weighted	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*				
	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)				
OLS - FF4	-0.03	0.02	0.02	-0.01	-0.19*	-0.17**	-0.17**	-0.18**	0.16	0.19*	0.19**	0.17*
	(0.05)	(0.08)	(0.06)	(0.07)	(0.10)	(0.08)	(0.08)	(0.08)	(0.11)	(0.10)	(0.08)	(0.09)
Elastic Net	-0.04	0.02	0.02	-0.02	-0.21*	-0.20**	-0.20**	-0.21***	0.17	0.22*	0.22**	0.17*
	(0.06)	(0.08)	(0.07)	(0.07)	(0.11)	(0.09)	(0.08)	(0.08)	(0.13)	(0.12)	(0.09)	(0.10)
Decision Tree	-0.05	-0.02	-0.02	-0.04	-0.12	-0.13**	-0.13**	-0.13**	0.07	0.11	0.11*	0.09
	(0.05)	(0.06)	(0.05)	(0.06)	(0.08)	(0.06)	(0.06)	(0.06)	(0.08)	(0.07)	(0.06)	(0.07)
Random Forest	0.00	0.01	0.01	0.01	-0.16*	-0.19***	-0.19***	-0.19***	0.12	0.16	0.16**	0.16**
	(0.05)	(0.06)	(0.06)	(0.05)	(0.09)	(0.07)	(0.07)	(0.06)	(0.10)	(0.09)	(0.08)	(0.07)
Gradient Boosting	0.00	0.02	0.02	0.00	-0.18*	-0.19***	-0.19***	-0.20***	0.18*	0.21**	0.21***	0.18**
	(0.05)	(0.06)	(0.06)	(0.06)	(0.10)	(0.07)	(0.07)	(0.06)	(0.11)	(0.09)	(0.08)	(0.08)
Neural Network (I)	-0.02	0.01	0.01	-0.02	-0.18*	-0.19***	-0.19***	-0.20***	0.16	0.20*	0.20**	0.16*
	(0.05)	(0.07)	(0.06)	(0.06)	(0.09)	(0.07)	(0.07)	(0.07)	(0.11)	(0.11)	(0.09)	(0.09)

2a) Sorting forecasts and long-short portfolios

Performance comes mainly from shorting the bottom 10%. To what extent can share classes of managed funds be shorted? How expensive is it?

	Panel A - top and bottom decile portfolios											
	Top 10%:				**Bottom 10%:**				**Long-Short:**			
	FF4	FF5	FF6	FF7	FF4	FF5	FF6	FF7	FF4	FF5	FF6	FF7
Equally weighted	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*	-0.08*				
	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)				
OLS - FF4	-0.03	0.02	0.02	-0.01	-0.19*	-0.17**	-0.17**	-0.18**	0.16	0.19*	0.19**	0.17*
	(0.05)	(0.08)	(0.06)	(0.07)	(0.10)	(0.08)	(0.08)	(0.08)	(0.11)	(0.10)	(0.08)	(0.09)
Elastic Net	-0.04	0.02	0.02	-0.02	-0.21*	-0.20**	-0.20**	-0.21***	0.17	0.22*	0.22**	0.17*
	(0.06)	(0.08)	(0.07)	(0.07)	(0.11)	(0.09)	(0.08)	(0.08)	(0.13)	(0.12)	(0.09)	(0.10)
Decision Tree	-0.05	-0.02	-0.02	-0.04	-0.12	-0.13**	-0.13**	-0.13**	0.07	0.11	0.11*	0.09
	(0.05)	(0.06)	(0.05)	(0.06)	(0.08)	(0.06)	(0.06)	(0.06)	(0.08)	(0.07)	(0.06)	(0.07)
Random Forest	0.00	0.01	0.01	0.01	-0.16*	-0.19***	-0.19***	-0.19***	0.12	0.16	0.16**	0.16**
	(0.05)	(0.06)	(0.06)	(0.05)	(0.09)	(0.07)	(0.07)	(0.06)	(0.10)	(0.09)	(0.08)	(0.07)
Gradient Boosting	0.00	0.02	0.02	0.00	-0.18*	-0.19***	-0.19***	-0.20***	0.18*	0.21**	0.21***	0.18**
	(0.05)	(0.06)	(0.06)	(0.06)	(0.10)	(0.07)	(0.07)	(0.06)	(0.11)	(0.09)	(0.08)	(0.08)
Neural Network (I)	-0.02	0.01	0.01	-0.02	-0.18*	-0.19***	-0.19***	-0.20***	0.16	0.20*	0.20**	0.16*
	(0.05)	(0.07)	(0.06)	(0.06)	(0.09)	(0.07)	(0.07)	(0.07)	(0.11)	(0.11)	(0.09)	(0.09)

2b) Forecasting accuracy and fund performance

"If ML models can improve the accuracy of flow prediction for a given fund compared to more basic methods [...], then the liquidity management of that fund can potentially be improved by using ML methods."

	Long portfolio:					**Short portfolio:**					**Long-Short:**				
	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6
Mean	-0.017	-0.045	-0.050	-0.048	-0.048	-0.048	-0.072*	-0.072*	-0.078**	-0.078**	0.030*	0.028	0.021	0.030	0.030*
	(0.066)	(0.045)	(0.044)	(0.048)	(0.046)	(0.063)	(0.042)	(0.043)	(0.038)	(0.038)	(0.016)	(0.018)	(0.016)	(0.022)	(0.018)
OLS	-0.022	-0.049	-0.050	-0.048	-0.048	-0.050	-0.074*	-0.076*	-0.083**	-0.083**	0.030***	0.028***	0.028***	0.037***	0.037***
	(0.065)	(0.042)	(0.042)	(0.040)	(0.040)	(0.063)	(0.042)	(0.043)	(0.041)	(0.041)	(0.011)	(0.01)	(0.010)	(0.011)	(0.011)

2b) Forecasting accuracy and fund performance

"If ML models can improve the accuracy of flow prediction for a given fund compared to more basic methods [...], then the liquidity management of that fund can potentially be improved by using ML methods."

	Long portfolio:					**Short portfolio:**					**Long-Short:**				
	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6
Mean	-0.017	-0.045	-0.050	-0.048	-0.048	-0.048	-0.072*	-0.072*	-0.078**	-0.078**	0.030*	0.028	0.021	0.030	0.030*
	(0.066)	(0.045)	(0.044)	(0.048)	(0.046)	(0.063)	(0.042)	(0.043)	(0.038)	(0.038)	(0.016)	(0.018)	(0.016)	(0.022)	(0.018)
OLS	-0.022	-0.049	-0.050	-0.048	-0.048	-0.050	-0.074*	-0.076*	-0.083**	-0.083**	0.030***	0.028***	0.028***	0.037***	0.037***
	(0.065)	(0.042)	(0.042)	(0.040)	(0.040)	(0.063)	(0.042)	(0.043)	(0.041)	(0.041)	(0.011)	(0.01)	(0.010)	(0.011)	(0.011)

A negative alpha on the long leg?

2b) Forecasting accuracy and fund performance

"If ML models can improve the accuracy of flow prediction for a given fund compared to more basic methods [...], then the liquidity management of that fund can potentially be improved by using ML methods."

	Long portfolio:					Short portfolio:					Long-Short:				
	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6
Mean	-0.017	-0.045	-0.050	-0.048	-0.048	-0.048	-0.072*	-0.072*	-0.078**	-0.078**	0.030*	0.028	0.021	0.030	0.030*
	(0.066)	(0.045)	(0.044)	(0.048)	(0.046)	(0.063)	(0.042)	(0.043)	(0.038)	(0.038)	(0.016)	(0.018)	(0.016)	(0.022)	(0.018)
OLS	-0.022	-0.049	-0.050	-0.048	-0.048	-0.050	-0.074*	-0.076*	-0.083**	-0.083**	0.030***	0.028***	0.028***	0.037***	0.037***
	(0.065)	(0.042)	(0.042)	(0.040)	(0.040)	(0.063)	(0.042)	(0.043)	(0.041)	(0.041)	(0.011)	(0.01)	(0.010)	(0.011)	(0.011)

A negative alpha on the long leg? Value of forecast w.r.t to naive predictions?

2b) Forecasting accuracy and fund performance

"If ML models can improve the accuracy of flow prediction for a given fund compared to more basic methods [...], then the liquidity management of that fund can potentially be improved by using ML methods."

	Long portfolio:					Short portfolio:					Long-Short:				
	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6	CAPM	FF3	FF4	FF5	FF6
Mean	-0.017	-0.045	-0.050	-0.048	-0.048	-0.048	-0.072*	-0.072*	-0.078**	-0.078**	0.030*	0.028	0.021	0.030	0.030*
	(0.066)	(0.045)	(0.044)	(0.048)	(0.046)	(0.063)	(0.042)	(0.043)	(0.038)	(0.038)	(0.016)	(0.018)	(0.016)	(0.022)	(0.018)
OLS	-0.022	-0.049	-0.050	-0.048	-0.048	-0.050	-0.074*	-0.076*	-0.083**	-0.083**	0.030***	0.028***	0.028***	0.037***	0.037***
	(0.065)	(0.042)	(0.042)	(0.040)	(0.040)	(0.063)	(0.042)	(0.043)	(0.041)	(0.041)	(0.011)	(0.01)	(0.010)	(0.011)	(0.011)

Perhaps one can look at the realized alphas based on predictability, i.e.,

i) Sort funds into quintiles based on R_j^2 each month.

ii) Compare average performance across quintiles.

iii) $Performance_{j,t+1} = \alpha + \beta R^2_{j,t} + \gamma \text{Controls}_{j,t} + \epsilon_{j,t+1}$.

Summary

Highly recommended reading! I learned a lot...