



## Motivation & data universe

- FOMC decisions create **cross-sectional return dispersion**; we predict which stocks outperform cap-group peers over the next 60 trading days.
- Inputs combine Fed / macro variables, FinBERT sentiment, and stock-level signals via five ML models.

498 U.S. equities, 5 sectors, 3 cap buckets, Jul 2024–Oct 2025. Label:  $Y_{i,t} = 1$  if  $R_{i,t:t+60} > \bar{R}_{cap(i),t:t+60}$ . Walk-forward folds: 2024Q4 / 2025Q1 / 2025Q2;

2025Q3 held-out test.

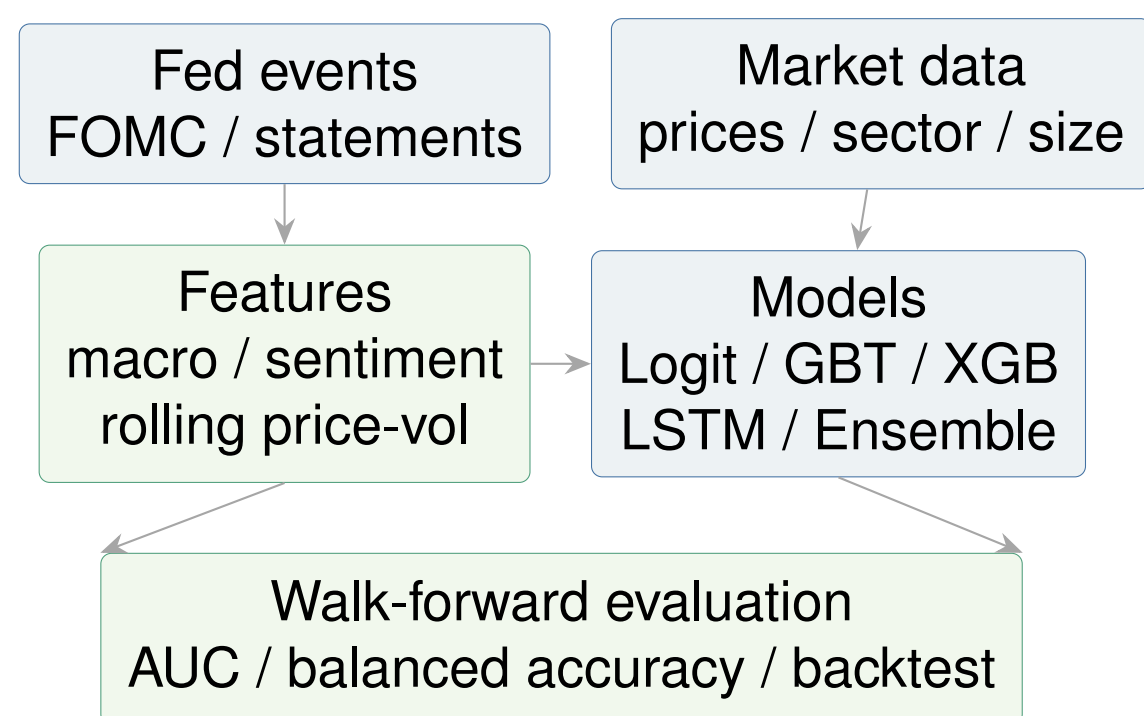
Group	Key features
Fed / macro / sentiment	FOMC surprise, FinBERT tone, VIX, yield-curve slope, credit spread
Price / volume	Rolling returns (5/20/60d), volatility, relative volume, momentum
Structural	Sector dummies, cap-bucket indicator, days-since-event

## Ensemble AUC peaks at 0.715 in favorable regimes

**Key finding:** the predictive signal is real (ensemble AUC up to 0.715) but regime-dependent—walk-forward validation and diagnostic transparency matter more than a headline metric.

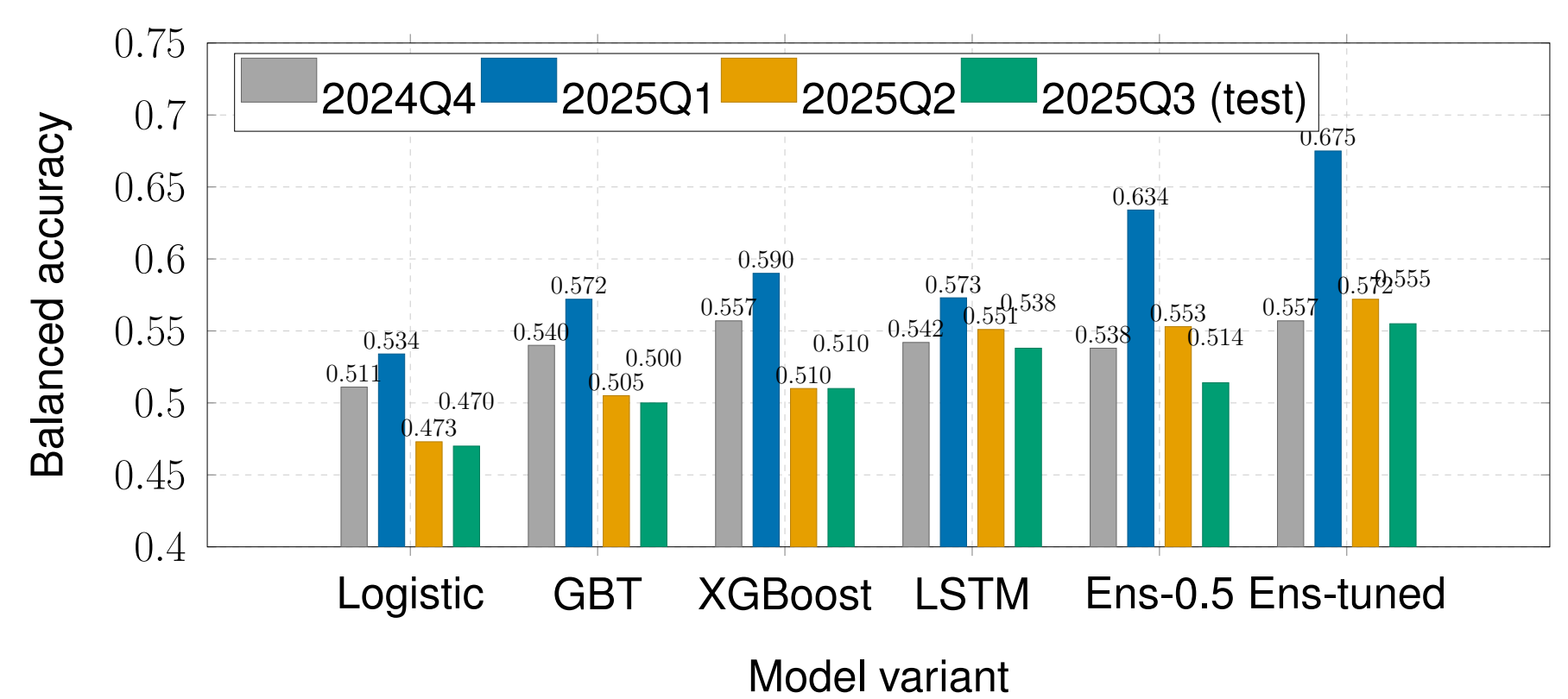
Stage	Best AUC	Ens. AUC	Excess Ret.
2024Q4	0.578	0.577	+2.23%
2025Q1	0.626	<b>0.715</b>	+4.04%
2025Q2	0.588	0.599	-2.87%
2025Q3 test	—	≈0.57	+6.87%

## Walk-forward pipeline prevents data leakage



Training uses only past data; standardization, thresholding, and model selection are based on training/validation only.

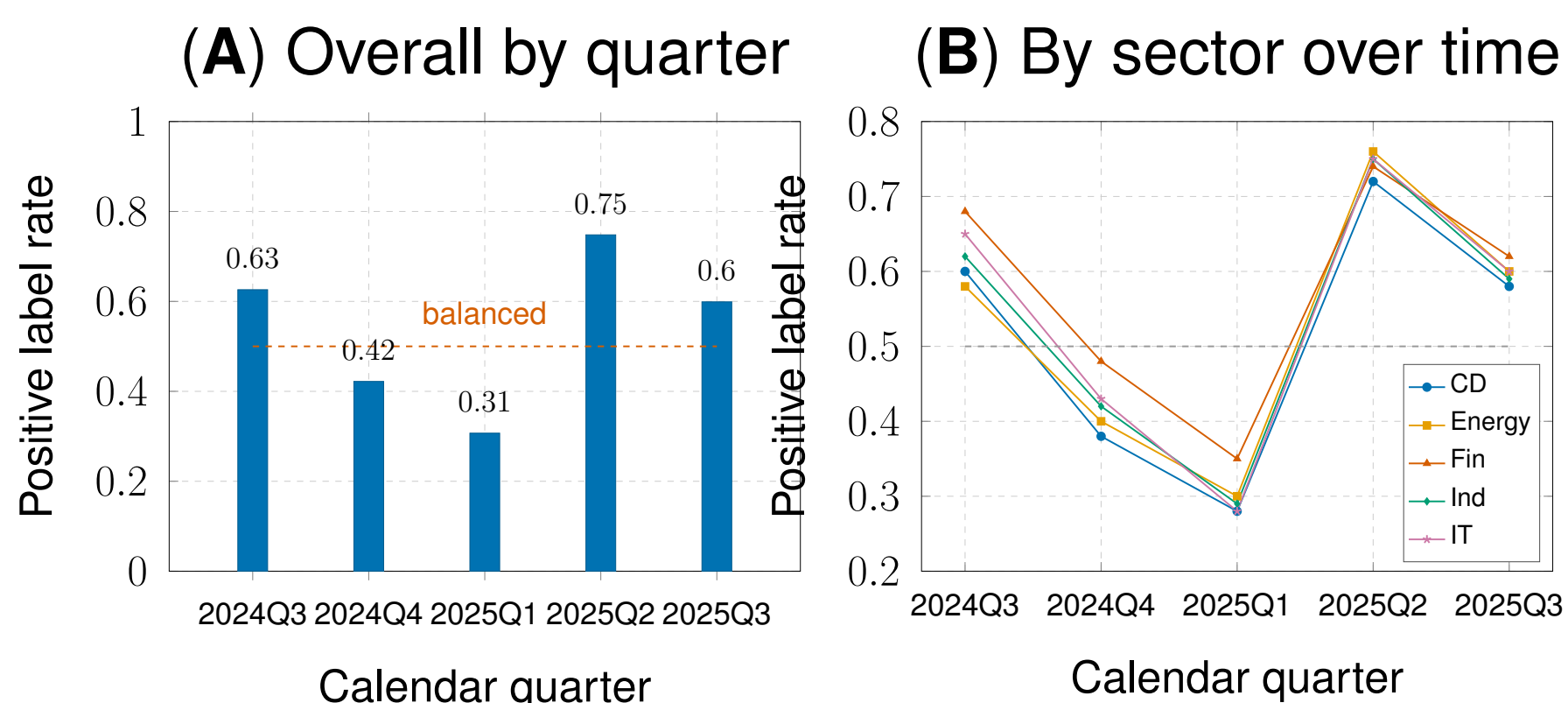
## Tree models dominate; signal is regime-dependent



Balanced accuracy by model; 2024Q4–2025Q2 validation, 2025Q3 out-of-sample test.

- Inverted-U pattern:** signal peaks in 2025Q1, weakens as market structure shifts.
- Tree vs. LSTM:** GBT/XGBoost dominate in favorable regimes; LSTM is more stable when the signal fades.
- Sector/size heterogeneity:** small caps and certain sectors become harder to predict in weaker regimes.

## Label balance shifts across quarters



Positive-label rate by quarter and over time.

## Signal is real but regime-dependent

- Fed, macro, and FinBERT features contain **modest but meaningful** signal for 60-day relative stock prediction.
- The signal is **regime-dependent**, not stable across adjacent quarters.
- Walk-forward validation** and **diagnostic transparency** matter more than one headline score.

**Acknowledgments** Prof. Shixin Xu, Prof. David Ye, DKU, and my family.