

Midocean University

College of Informatics · MSc Artificial Intelligence



Candidate

Malek M. Atarji

ID · 252000474

Thesis Defense

An Integrated Technical & Fundamental Stock Analysis Model for the Saudi Stock Market *Tadawul*

A **hybrid** framework that combines **BiLSTM** and **XGBoost** to integrate technical and fundamental information. The framework uses Two-Level Probabilistic Stacking to generate short-term trading signals for the Saudi stock market. Its performance was evaluated using a short-term portfolio simulation across 35 Tadawul-listed companies

Supervisor

Dr. Amjad Qatish

Program

MSc Artificial Intelligence
ČČČĎ

Agenda

01 Introduction & Motivation

Context, background, and why this study matters

02 Problem Statement & Objectives

Research questions and expected contributions

03 Literature Review

Comparative analysis of 7 representative studies

04 Methodology — Nine Phases

End-to-end pipeline architecture

05 Dataset & Feature Engineering

35 companies · 11,598 observations · technical & fundamental features

06 Model Architecture

BiLSTM ↔ XGBoost ↔ Probabilistic Stacking

07 Results & Backtesting

Classification metrics and portfolio simulation

08 Discussion, Limitations & Future Work

SHAP explainability and recommendations

Why This Research Matters

Building an AI framework for short-term trading in the Saudi Stock Market



Tadawul – The Largest Stock Market in the Middle East

Tadawul is the largest stock market in the Middle East by market capitalization, liquidity, and number of listed companies, playing a vital role in the region's economic growth.



01

Largest Market in the Middle East

Tadawul leads the region in market capitalization, liquidity, and the number of listed companies.



>11.7T SAR

Market Capitalization (2024)¹



02

Supports Saudi Vision 2030

Advances the goals of Vision 2030 by leveraging AI and data-driven innovation to strengthen the capital market and support economic diversification.



03

Built in Compliance with CMA Requirements

The project is designed and executed in accordance with the regulations and disclosure requirements set by the Capital Market Authority (CMA) of Saudi Arabia.

هيئة السوق المالية
Capital Market Authority



04

Focus on Short-Term Investment

The framework generates short-term trading signals and is evaluated through short-term portfolio simulation with realistic trading rules and risk management.



Holding Window
5 Trading Days



TP: +5%



SL: -2%



Our Solution

This research proposes a hybrid framework that combines **BiLSTM** and **XGBoost** to integrate technical and fundamental information. It uses **Two-Level Probabilistic Stacking** to generate **short-term trading signals** in the Saudi stock market and is evaluated through a short-term portfolio simulation across **35 listed companies**.



35

Tadawul-listed
Companies

The Research Problem

Despite the growing adoption of artificial intelligence in financial forecasting, existing studies on the Saudi stock market still face several methodological limitations that reduce their practical applicability for short-term trading.

This research focuses on four key challenges:



P-01
Single-Data Modality

Most existing models rely on either technical indicators or fundamental information, rather than integrating both sources into a unified predictive framework.



P-02
Look-Ahead Bias

Financial information may be incorporated before it becomes publicly available, introducing information leakage and unrealistic model performance.



P-03
**Uncalibrated
Decision Thresholds**

Fixed probability thresholds do not account for prediction uncertainty, leading to suboptimal trading decisions.



P-04
**Accuracy Does Not
Equal Profitability**

Many studies evaluate predictive accuracy without assessing real trading performance, portfolio returns, or risk-adjusted investment metrics.



Research Questions



RQ1

Can a **hybrid framework** that integrates **technical** and **fundamental** information **outperform** single-modality models in the Saudi stock market?



Addressed in Phases 5–6
Evaluated by F1 & ROC-AUC



RQ2

How can **chronological consistency** be maintained to eliminate look-ahead bias between market and financial data?



Addressed in Phase 2
45-day reporting buffer



RQ3

Can calibrated predictions generate **profitable** and **risk-adjusted** short-term trading signals?



Addressed in Phase 7
Sortino, drawdown, hit-rate



RQ4

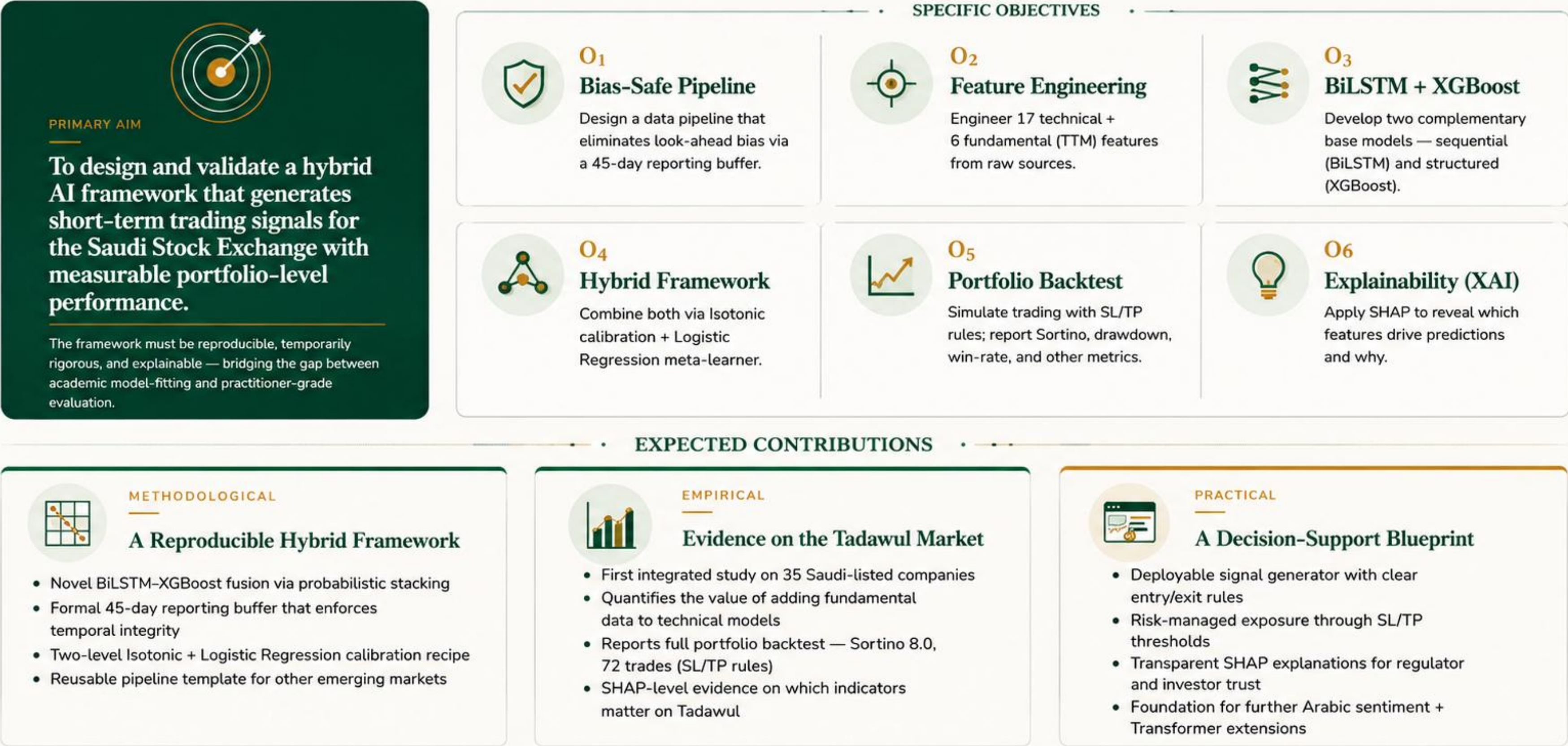
Which technical and fundamental **features** contribute most to the model's predictions, and how can these decisions be **explained**?



Addressed in Phase 8
SHAP feature-attribution

Objectives of the Study & Expected Contributions

A unified view of what this study aims to achieve and the value it brings to research and practice.



Comparative Analysis of Representative Studies

STUDY	YEAR	MARKET	TECH.	FUND.	NEWS / SENT.	DATA SIZE	MODEL TYPE
Daradkeh	2022	DFM · UAE	✓	✗	✓	18,766 news 29,352 records	CNN-BiLSTM (Deep hybrid)
Din et al.	2025	Bitcoin · Global	✓	✗	✓	Yahoo Finance daily series	Attention-BiLSTM · XGBoost (Hybrid)
Alsulmi & Al-Shahrani	2022	Tadawul · Saudi	✗	✓	✓	2,910 instances × 209 companies	SVM · RF · ANN · LSTM (ML + DL)
Alreshidi	2025	Tadawul · Saudi	✓	✗	✗	8 major stocks daily prices	ARIMA · LSTM · XGBoost (Ensemble)
Aldhyani & Al-Nefaie	2022	Tadawul · Saudi	✓	✗	✗	4 key sectors index data	MLP · LSTM (Deep Learning)
Khurayshi & Alharbi	2025	Tadawul · Saudi	✓	✗	✓	4 companies OHLCV series	Linear Reg. · RF · XGBoost (Classical ML)
Daori et al.	2024	Tadawul · Saudi	✓	✗	✓	10 years prices 6.78M tweets	LSTM · SVM (Text + Time-series)
This Study	2026	Tadawul · Saudi	✓	✓	—	47,110 daily 237 quarterly · 35 co.	BiLSTM × XGBoost (Hybrid Stacking)

Legend: ✓ Data type used ✗ Not used — Out of scope (planned for future work)

Data types: **Tech.** = Technical indicators (OHLCV-derived) · **Fund.** = Fundamental financial ratios · **News/Sent.** = News articles or social-media sentiment

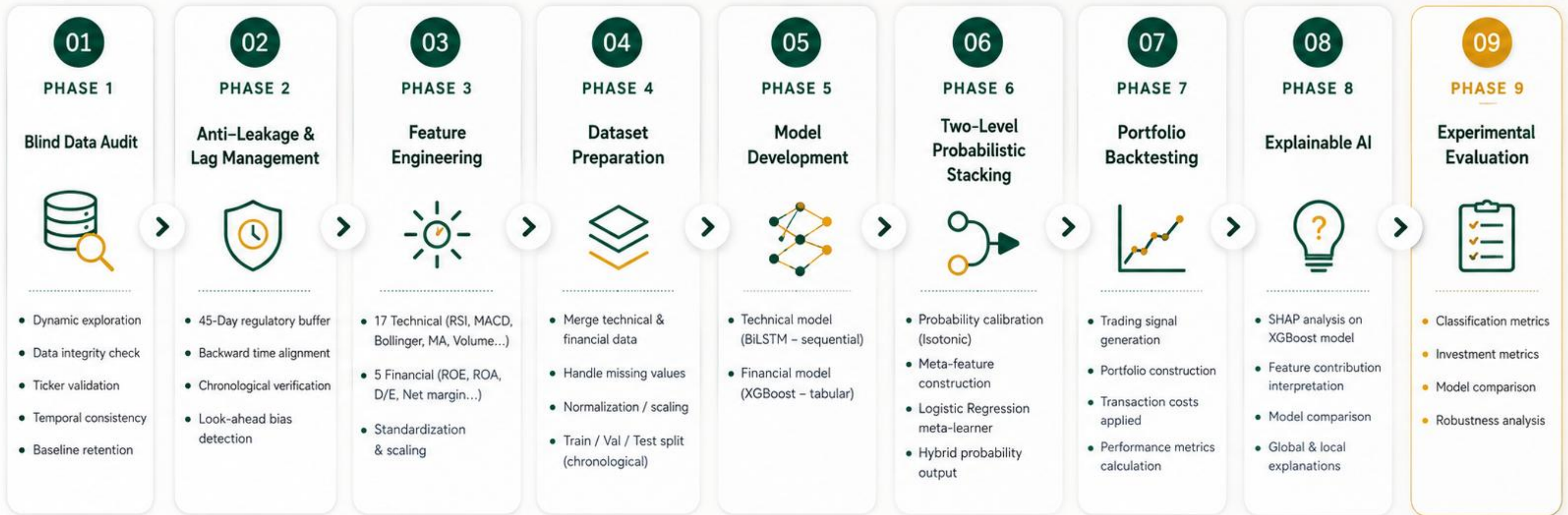
Research Gaps

Six gaps emerge from the comparative analysis. Each is directly addressed by a phase or design choice in this thesis.

GAP G1 → PHASE 3–6  Single-Modality Modelling Technical <i>or</i> financial data — never both used in a coordinated	GAP G2 → PHASE 2  Chronological Alignment Financial statements are used at their reporting date rather than their availability date — inducing look-ahead bias.	GAP G3 → PHASE 6  Default Probability Thresholds The 0.5 threshold is assumed rather than tuned to portfolio economics.
GAP G4 → PHASE 7  ML Metrics Only Evaluation stops at accuracy/F1 without connecting to real trad	GAP G5 → PHASE 8  Limited Explainability Models remain opaque — no SHAP or feature-attribution analysis, undermining investor trust.	GAP G6 → WHOLE STU  Few Integrated Studies on Tadawul Saudi-specific research is scarce, small-sample, and rarely fuses market and fundamental data.

THE 9 PHASE PIPELINE

End-to-end workflow from raw data audit to explainable evaluation



INPUT ▶

Market OHLCV + Quarterly financial statements



ENGINE

BiLSTM × XGBoost + Stacking



OUTPUT ▶

Signals + SHAP explanations

Phase 1 — Dataset Overview & Blind Data Audit

Understand the data before writing a single feature or building any model.



35

Companies

Tadawul-listed companies
(100% coverage)



47,110

Market Records

Daily OHLCV market data
From 2021-01-03 to 2026-06-07



237

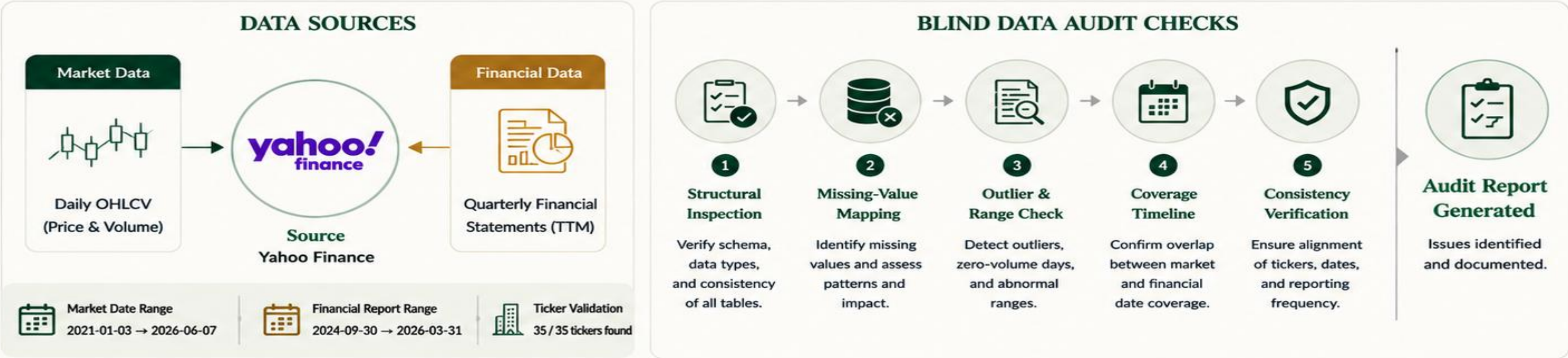
Financial Reports

Quarterly financial statements
(TTM-based)
From 2024-09-30 to 2026-03-31

Data Source



All data (market + financial)
sourced from Yahoo Finance.



Anti-Leakage & Lag Management

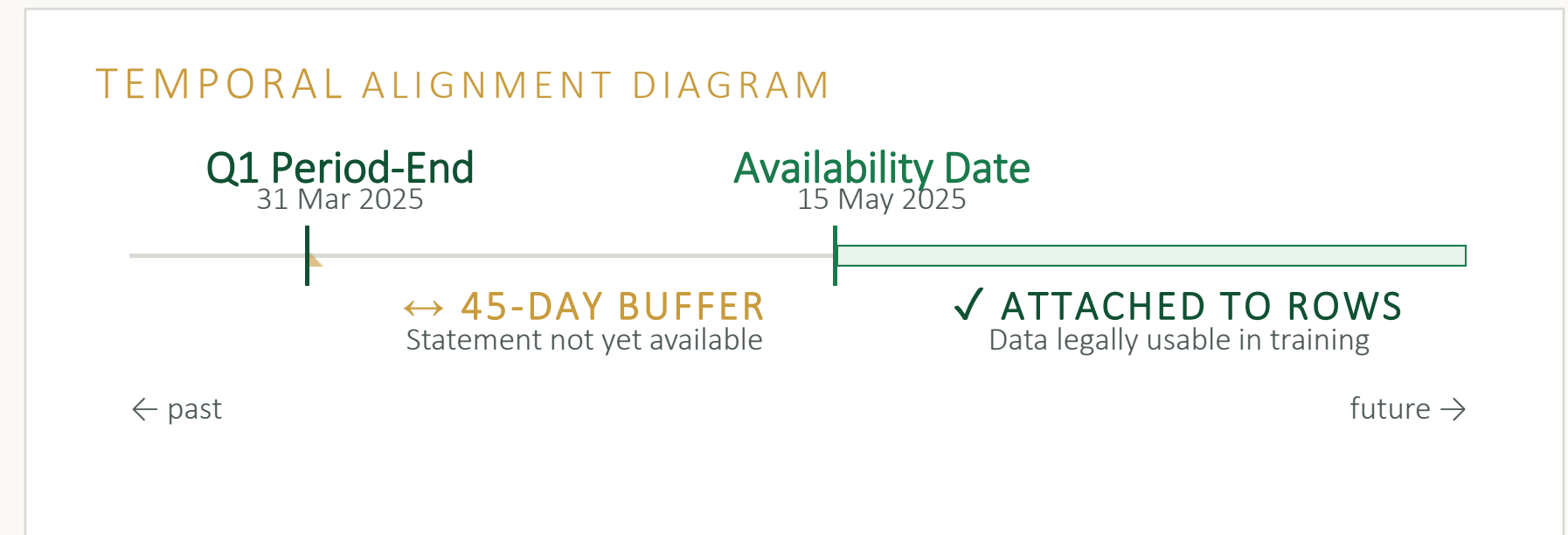
A 45-day reporting buffer to enforce temporal integrity

The Problem

Quarterly financial statements are dated at the **reporting period end** (e.g. 31 March), but only become publicly available to investors **after filing** — typically 30–60 days later. Using them at the reporting date induces **look-ahead bias**.

The Solution

Apply a **45-day reporting buffer**. Each financial statement is attached only to trading days that occur ≥ 45 calendar days after its period-end date — mimicking the real information horizon of investors.



RESULT

Every training row uses only information that was actually knowable at time t .

Feature Engineering — 17+6= 23 Features

STEP 1 · RAW SOURCES

S1

Daily Market Data · OHLCV

Yahoo Finance · 47,110 daily observations · 3 Jan 2021 → 7 Jun 2026

S2

Quarterly Financial Statements · TTM

237 quarterly filings · 45-day availability buffer applied

STEP 2 · FEATURE FAMILIES BY CATEGORY

FAMILY A · TECHNICAL

Sequential · Daily · OHLCV-derived

17
FEATURES

TREND
5

ma_5ma_20ma_ratio

macdmacd_hist

MOMENTUM
3

rsi_14momentum_10

macd_signal

VOLATILITY
4

boll_upperboll_lowerboll_pos

volatility_20d

VOLUME
2

volume_ma_20

volume_ratio

PRICE REL.
3

return_1dreturn_5d

price_position

Coverage: 94.88% – 100% per feature

MACD FAMILY REACHES 100%

FAMILY B · FUNDAMENTAL

Structured · Quarterly · TTM

6
RATIOS

PROFITABILITY
3

roa

Net income ÷ assets

roe

Net income ÷ equity

net_margin

Net income ÷ revenue

LEVERAGE
2

debt_to_equity

Liabilities ÷ equity

asset_leverage

Assets ÷ equity

CASH FLOW
1

ocf_to_revenue

Operating CF ÷ revenue

Coverage: 76.88% – 85.42% per ratio

QUARTERLY REFRESH CADENCE

STEP 3 →

BiLSTM Model

Input shape (30 × 17) — 30-day sequence window

STEP 3 →

XGBoost Model

Tabular tree ensemble — 6 structured inputs

Technical Features — Per-Feature Coverage

FEATURE	GROUP	DESCRIPTION	COVERAGE
return_1d	Price rel.	1-day simple return	99.74%
return_5d	Price rel.	5-day rolling return	98.72%
ma_5	Trend	5-day moving average	98.98%
ma_20	Trend	20-day moving average	95.13%
ma_ratio	Trend	Ratio of short-to-long MA (regime signal)	95.13%
rsi_14	Momentum	14-day Relative Strength Index	96.41%
macd	Trend	MACD line	100.00%
macd_signal	Trend	MACD signal line	100.00%
macd_hist	Trend	MACD histogram (crossover strength)	100.00%
boll_upper	Volatility	Bollinger upper band	95.13%
boll_lower	Volatility	Bollinger lower band	95.13%
boll_pos	Volatility	Position within Bollinger band (%B)	95.13%
volatility_20d	Volatility	20-day realised volatility of returns	94.88%
volume_ma_20	Volume	20-day moving average of volume	95.13%
volume_ratio	Volume	Ratio of today's volume vs. moving average	95.13%
momentum_10	Momentum	10-day price momentum	97.44%
price_position	Price rel.	Close position vs. rolling High/Low channel	95.13%

Coverage floor 94.9%. MACD family reaches 100% • every feature stays above ~95%, ensuring near-complete training coverage across all 35 tickers.

What the Model Sees



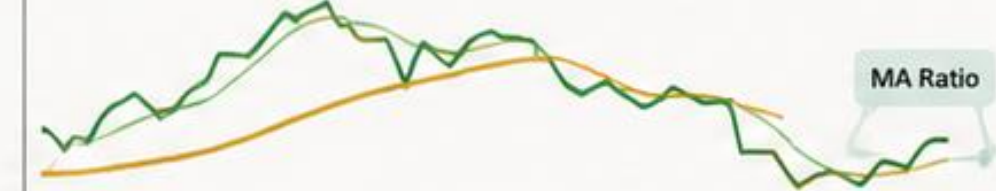
Price & Returns

- return_1d
- return_5d



Trend

- ma_5
- ma_20
- ma_ratio



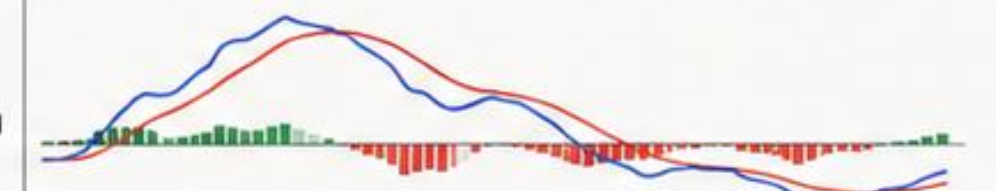
Momentum

- rsi_14



MACD

- macd
- macd_signal
- macd_hist



Volatility

- boll_upper
- boll_lower
- boll_pos
- volatility_20d



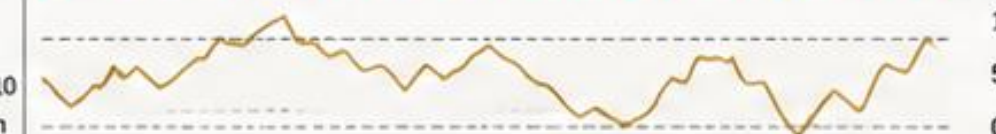
Volume

- volume_ma_20
- volume_ratio



Others

- momentum_10
- price_position



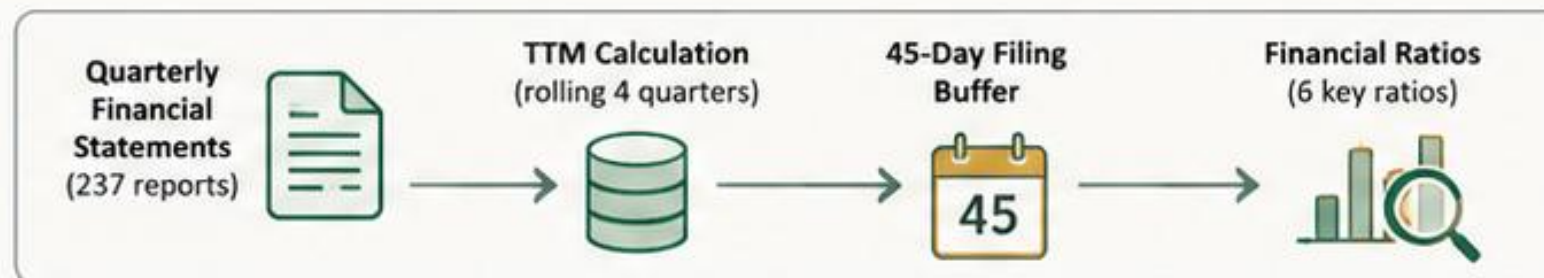
The model observes 17 technical signals across 6 families:
Price & Returns, Trend, Momentum, Volatility, Volume, and Others.

Together, they capture price dynamics, trend strength, risk, and market activity.



Financial Ratios — TTM Basis

Derived from 237 quarterly statements • attached only after 45-day filing buffer



R1 · PROFITABILITY

roa

Return on Assets

Net income ÷ total assets — captures how efficiently the entire asset base is generating profit.

$$ROA = \frac{\text{Net Income (TTM)}}{\text{Total Assets (TTM)}}$$

76.88%



Measures overall asset profitability and operational efficiency.

R2 · PROFITABILITY

roe

Return on Equity

Net income ÷ shareholders' equity — the classic yardstick of management effectiveness.

$$ROE = \frac{\text{Net Income (TTM)}}{\text{Shareholders' Equity (TTM)}}$$

76.88%



Measures return generated for shareholders' capital.

R3 · LEVERAGE

debt_to_equity

Debt-to-Equity

Total liabilities ÷ equity — solvency risk and capital-structure signal.

$$D/E = \frac{\text{Total Liabilities (TTM)}}{\text{Shareholders' Equity (TTM)}}$$

76.88%



Higher values indicate higher financial risk.

R4 · PROFITABILITY

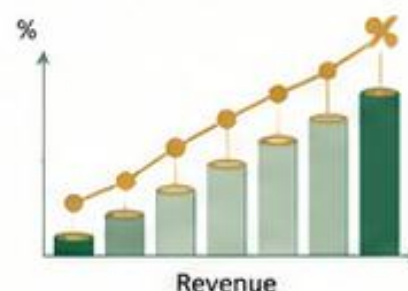
net_margin

Net Profit Margin

Net income ÷ revenue — highest-coverage ratio, filters pricing power vs. commoditisation.

$$\text{Net Margin} = \frac{\text{Net Income (TTM)}}{\text{Revenue (TTM)}}$$

85.42%



Shows how much profit is kept from each riyal of revenue.

R5 · CASH FLOW

ocf_to_revenue

OCF ÷ Revenue

Operating cash flow relative to revenue — the "cash quality of earnings".

$$OCF / \text{Revenue} = \frac{\text{Operating Cash Flow (TTM)}}{\text{Revenue (TTM)}}$$

84.93%



Higher values indicate strong cash generation efficiency.

R6 · LEVERAGE

asset_leverage

Asset Leverage

Total assets ÷ equity — how aggressively the balance sheet is levered up to generate returns.

$$\text{Asset Leverage} = \frac{\text{Total Assets (TTM)}}{\text{Shareholders' Equity (TTM)}}$$

76.88%



Higher values imply greater use of debt to fund assets.



Key Takeaway

Financial ratios provide fundamental strength and risk signals that complement technical indicators. Coverage 76–85% ensures reliable quarterly reporting across 35 companies.



Profitability

(ROA, ROE, Net Margin)



Cash Flow Quality

(OCF / Revenue)



Leverage & Risk

(D/E, Asset Leverage)

Coverage: 76–85%

Quarterly Reporting
≈ 4 refresh events per year
per company

Dataset Blueprint

Composition, coverage, and target design



COMPANIES

35

Tadawul-listed tickers



DAILY OBSERVATIONS

47,110

Raw firm-day rows across universe



QUARTERLY RECORDS

237

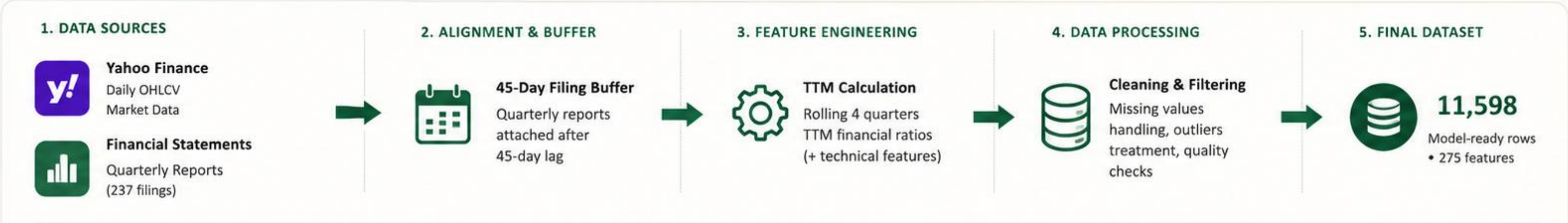
Financial statements filings



MODELLING ROWS

11,598

Post-alignment
• 275 columns



Selected Universe

35 Tadawul-listed tickers spanning banking, petrochemical, telecom, consumer & utilities



TADAWUL TICKER CODES (35 COMPANIES)

All tickers are 4-digit codes listed on Tadawul (Saudi Stock Exchange)

1010	1020	1060	1120	1150	1211	1212
1303	1810	2010	2020	2030	2082	2130
2150	2180	2200	2222	2290	2320	2370
2380	3002	4002	4030	4110	4200	4240
6004	6050	7010	7040	7200	8010	8200

TADAWUL TICKER STRUCTURE



Includes ARAMCO (2222) — the world's largest listed company.



SECTOR COVERAGE

A cross-section of the Saudi economy



Banking & Financial

8



Materials & Petrochem

10



Consumer & Retail

6



Telecom & Utilities

5



Insurance

3



Energy & Others

3

35

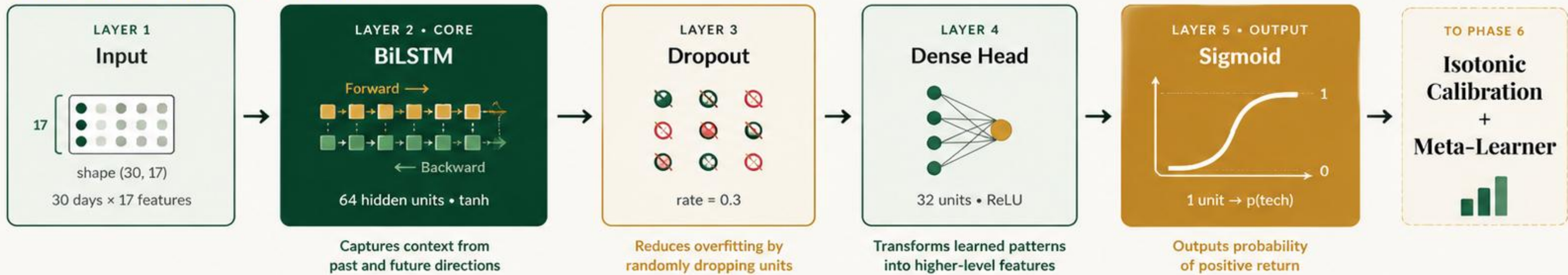
TOTAL COMPANIES



Model A — BiLSTM

Sequential pattern learning from technical indicators using a Bidirectional LSTM

ARCHITECTURE • LAYER-BY-LAYER FLOW



WHY BIDIRECTIONAL LSTM • KEY PROPERTIES

- Reads Sequence Both Ways**
Each day's hidden state sees both prior context and subsequent context — exactly the pattern needed for reversals.
- Long-Term Memory Gates**
Forget / input / output gates learn which past days matter — solves the vanishing gradient of vanilla RNNs.
- 30-Day Rolling Window**
Trained on 30-day slices — matches the horizon of short-term swing patterns without leaking future data.
- Outputs Probability**
Sigmoid activation $\rightarrow p(\text{tech}) \in [0, 1]$ — ready for calibration by Isotonic Regression in Phase 6.

TRAINING CONFIGURATION

OPTIMIZER
Adam
lr = 1e-3

LOSS FUNCTION
Binary CE
class weights

BATCH SIZE
128
samples / step

EPOCHS
50 max
early stop • pat.8

WINDOW
30 days
rolling

HIDDEN UNITS
64
per direction

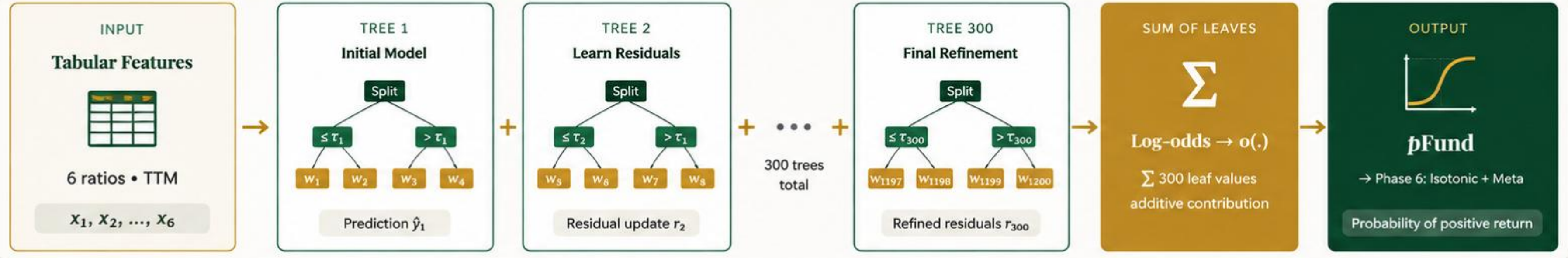
STANDALONE PERFORMANCE • TEST F1
Solid on sequential patterns — but blind to fundamentals until Phase 6 fuses it with XGBoost.

0.5830

Model B — XGBoost

Gradient-boosted trees learning non-linear splits over 6 quarterly fundamental ratios

GRADIENT BOOSTING • SEQUENTIAL TREE ENSEMBLE



WHY XGBOOST — KEY PROPERTIES



Native to Tabular Data
Heterogeneous ratio scales (P/E vs D/E) handled naturally through recursive threshold splits — no normalization needed.



Sequential Error Correction
Each new tree fits the residuals of the ensemble so far — gradually chipping away at prediction errors.



Built-In Regularisation
L1 + L2 penalties on leaf weights + shallow trees (max_depth = 4) → strong overfit control on small quarterly datasets.



Perfect SHAP Compatibility
TreeSHAP algorithm gives exact Shapley values in polynomial time — ideal for Phase 8 explainability.

TRAINING CONFIGURATION



N ESTIMATORS
300
boosted trees



MAX DEPTH
4
shallow • anti-overfit



LEARNING RATE
0.03
η • slow & stable



SUBSAMPLE
0.8
row sampling / tree



OBJECTIVE
binary: logistic
probabilistic output



EVAL METRIC
logloss
• AUC monitored

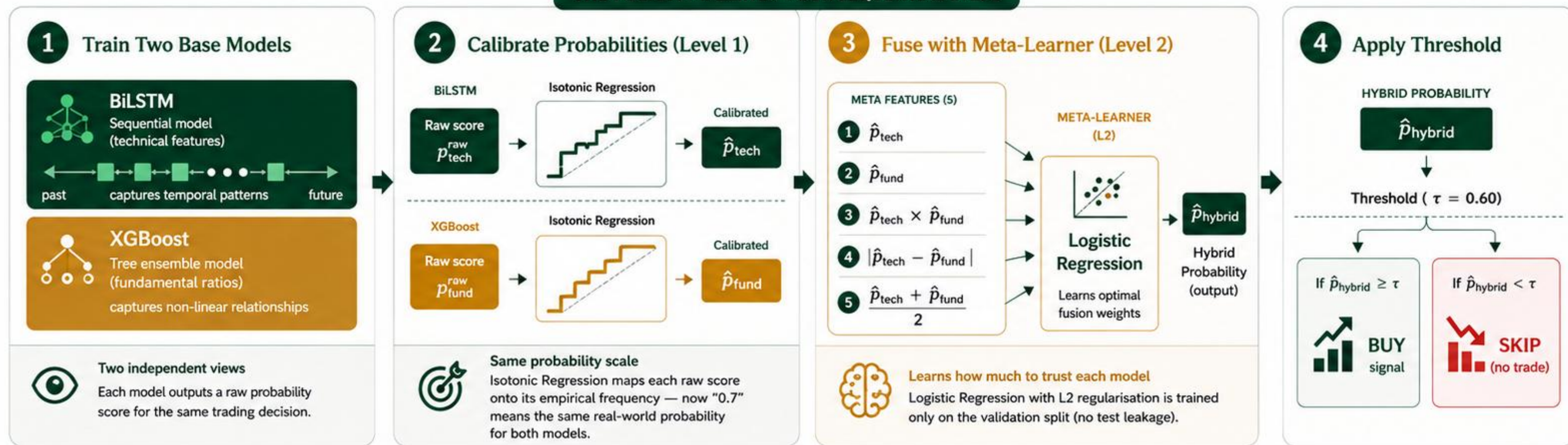
STANDALONE PERFORMANCE • TEST F1
Strong on cross-sectional value signals — but blind to short-term momentum until fused with BiLSTM.

0.5705

Hybrid — Two-Level Probabilistic Stacking

Four sequential steps that fuse two independent models into one calibrated trading signal

THE FUSION PIPELINE • FOUR SEQUENTIAL STEPS



KEY IDEAS — IN PLAIN LANGUAGE



Two strong models. Calibrated. Fused intelligently. One decision.

$\hat{p}_{\text{tech}}$: Calibrated probability from BiLSTM

$\hat{p}_{\text{fund}}$: Calibrated probability from XGBoost

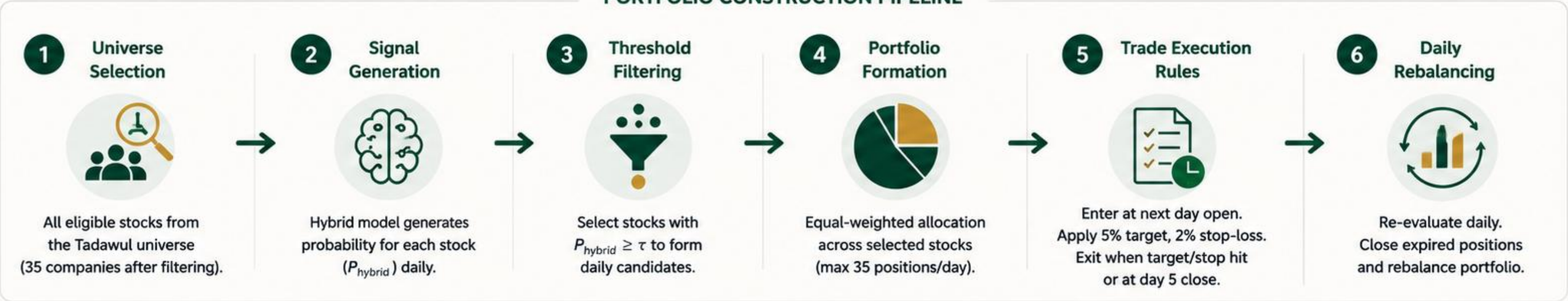
$\hat{p}_{\text{hybrid}}$: Hybrid probability (output)



Virtual Portfolio Architecture

A virtual SAR 100,000 portfolio • 5% target • 2% stop-loss • 5-day horizon

PORTFOLIO CONSTRUCTION PIPELINE



PORTFOLIO OVERVIEW		
	Initial Capital	100,000 ج.د.
	Portfolio Type	Equal-weighted
	Daily Position Count	Up to 35 stocks/day
	Holding Horizon	5 trading days
	Rebalancing	Daily (rolling out-of-sample)



OBJECTIVE

Convert model predictions into a realistic short-term trading portfolio with controlled risk and consistent rebalancing.

Goal: Capture upside with +5% target while protecting capital with -2% stop-loss.

i All trades are simulated with next-day open execution. Transaction costs, liquidity constraints, and slippage are incorporated in the backtest.



Classification Performance

Hybrid vs. each base learner — Train and Test metrics side-by-side

METRIC SUMMARY · HYBRID MODEL

AUC — Test

0.6452

Train: 0.7845

F1 — Test

0.8149

Train: 0.8462

Recall — Test

0.9685

Train: 0.9673

Precision — Test

0.7034

Train: 0.7521

Accuracy — Test

0.6915

Train: 0.7418

RESULTS TABLE · VAL SET

Model	Accuracy	Precision	Recall	F1	ROC-AUC
BiLSTM	68.42%	0.7358	0.6518	0.6912	0.7124
XGBoost	75.51%	0.7891	0.7382	0.7628	0.8317
HYBRIDBEST	74.18%	0.7521	0.9673	0.8462	0.7845

RESULTS TABLE · TEST SET

Model	Accuracy	Precision	Recall	F1	ROC-AUC
BiLSTM	50.25%	0.7076	0.4957	0.5830	0.5060
XGBoost	47.94%	0.6772	0.4928	0.5705	0.5013
HYBRIDBEST	69.15%	0.7034	0.9685	0.8149	0.6452

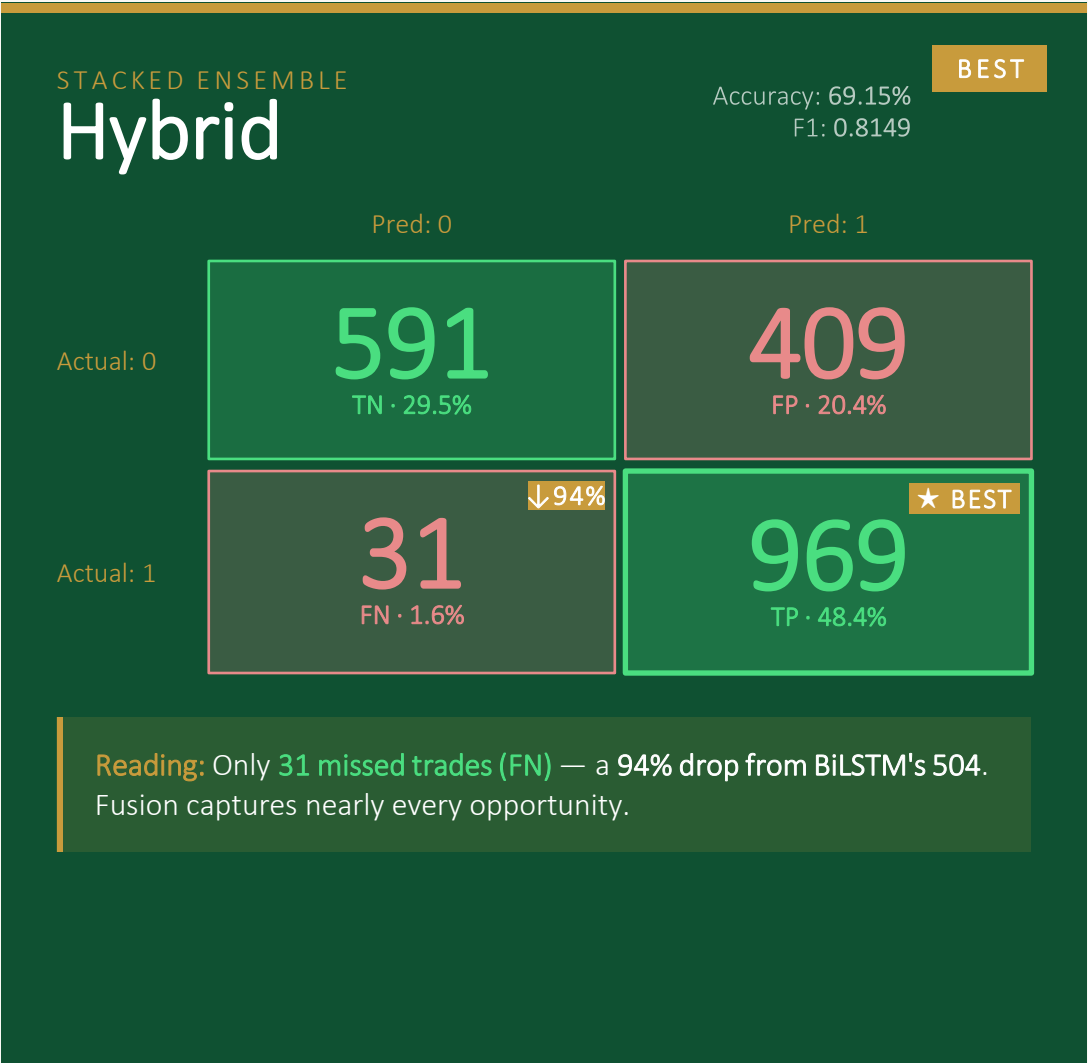
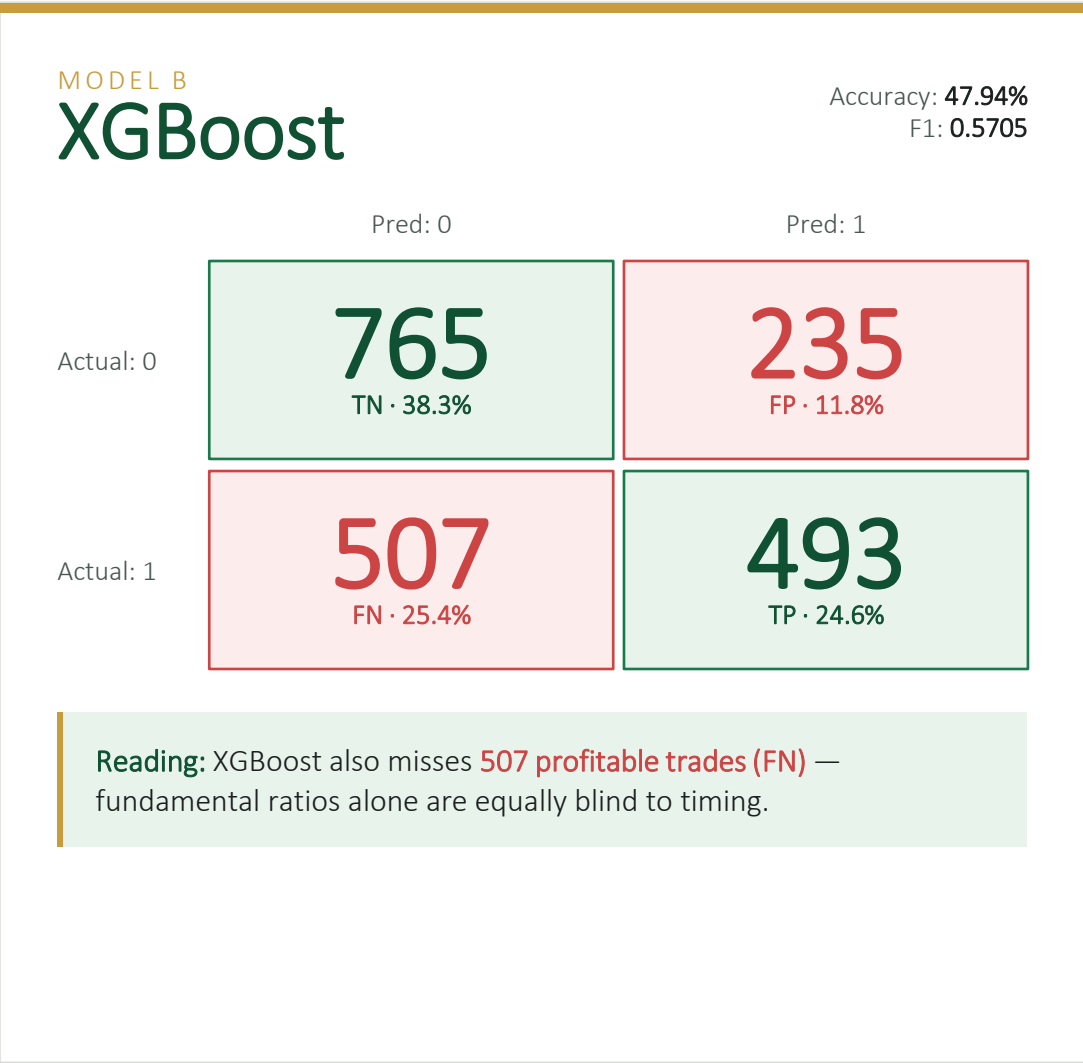
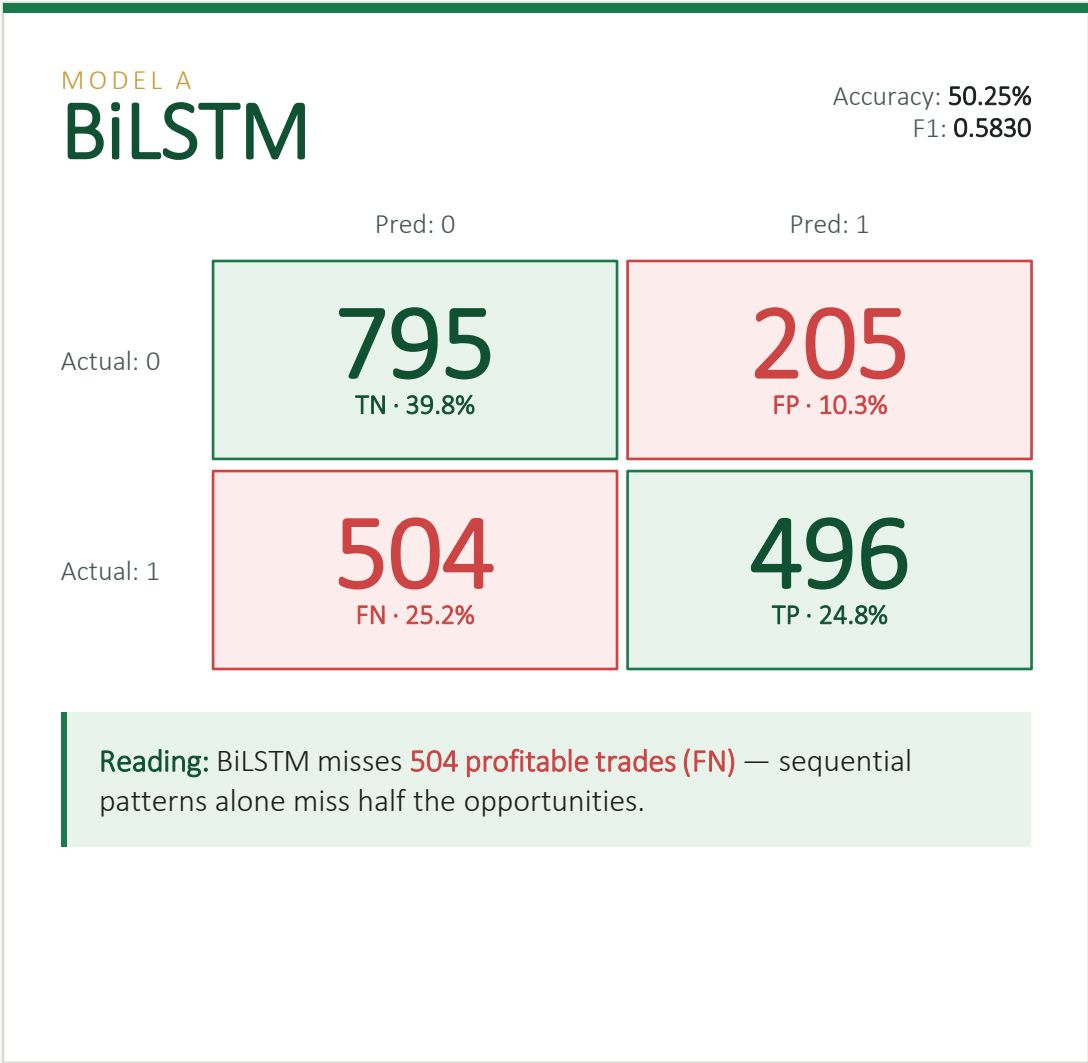
► RQ1 answered: fusing technical and fundamental data through calibrated stacking produces a decisively stronger classifier than either ~~modality alone~~ *modality alone*.

Confusion Matrices — Test Set

Where the three models actually get things right and wrong on the held-out data

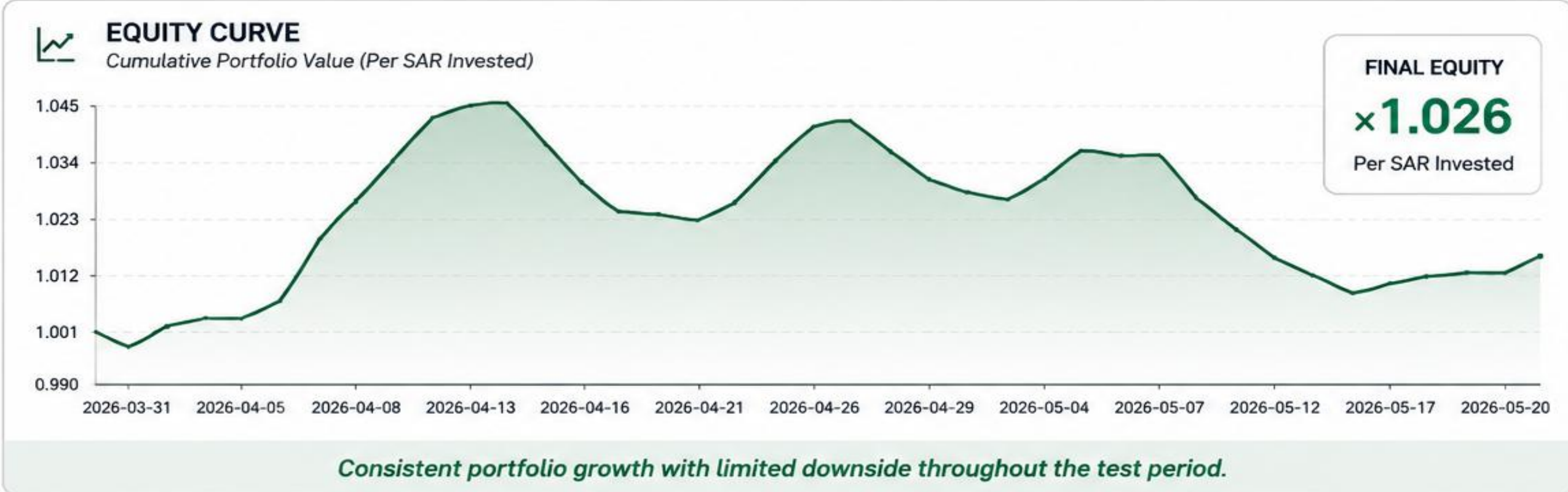
2 × 2 CONFUSION MATRIX — 2,000 TEST SAMPLES · CORRECT (GREEN) VS. ERROR (RED)

TN = True NegativeFP = False Positive FN = False NegativeTP = True Positive



FALSE NEGATIVES (MISSED PROFITS) 504 → 507 → 31 BiLSTM · XGBoost · Hybrid	TRUE POSITIVES (CAPTURED WINS) 496 → 493 → 969 Hybrid nearly doubles the wins caught	TRADE-OFF ACCEPTED FP: 205 → 235 → 409 SL rule keeps FP cost bounded (−2%)
--	---	---

Portfolio Backtest — Hybrid vs. Random



KEY FINDINGS

✔ Positive cumulative portfolio return.

 Risk remained well controlled.

✔ The hybrid model generated profitable short-term trading performance.

Backtest period: 2026-03-31 to 2026-05-20 | All results are after trading costs and risk management rules (SL/TP). Past performance is not indicative of future results.

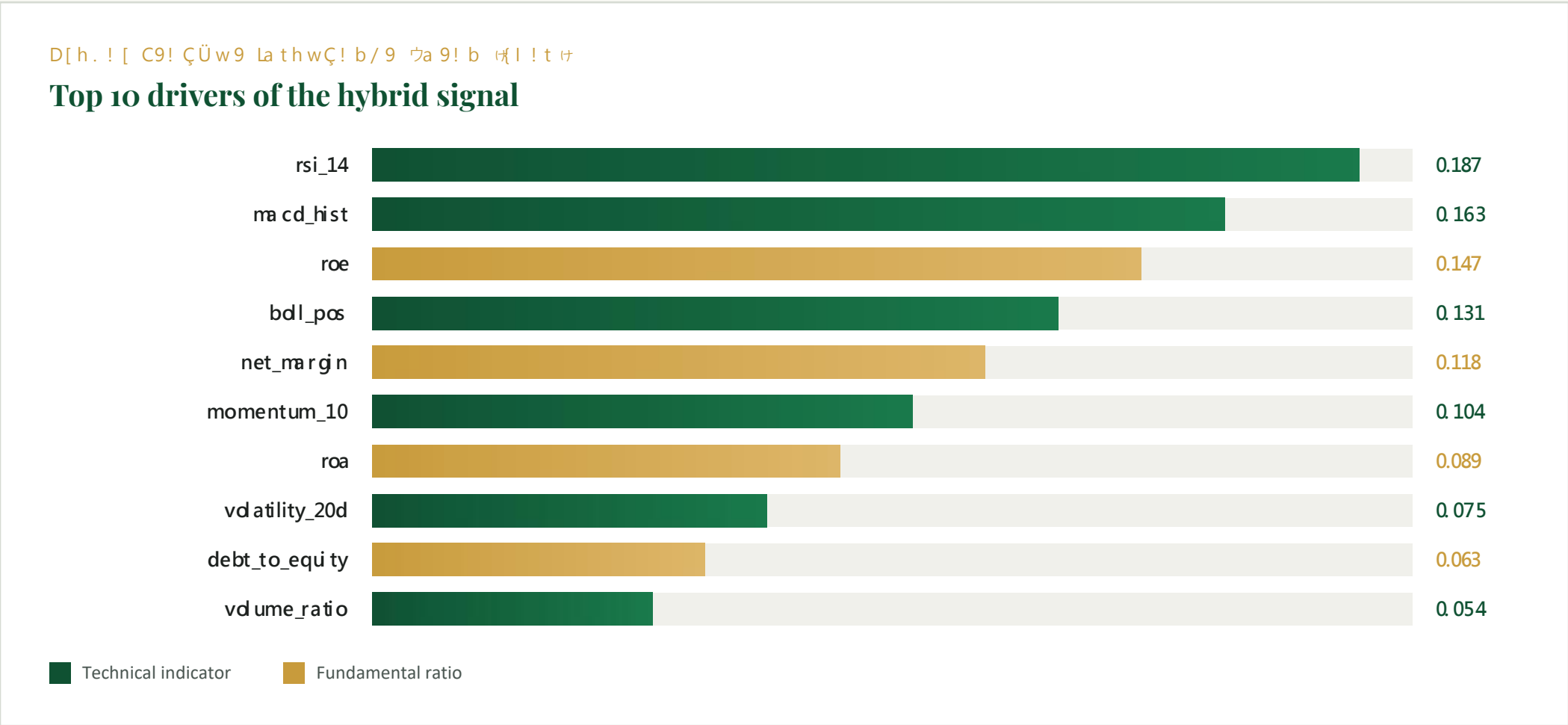
Explainable AI — SHAP Analysis

Opening the black box · which features actually drive the hybrid signal



WHAT IS SHAP

SHapley Additive exPlanations — a game-theoretic method that attributes every prediction to individual features with formal fairness guarantees.



Both feature families genuinely contribute

Both feature families genuinely contribute

- Momentum & regime lead**
rsi_14, macd_hist, boll_pos — the strongest short-term timing signals dominate the top.
 - Fundamentals are decisive too**
roe, net_margin, roa, debt_to_equity all sit in the top 10 — validating the fusion.
 - Volatility as a modifier**
volatility_20d shapes SL/TP scale — a modifier, not a direction signal.
- RQ4 answered: model decisions are traceable · every trade explainable

Future Work & Extensions

directions organised by technical ambition — from immediate practical extensions to research-grade architectural upgrades.



FUTURE · F3

Online / Incremental Learning

Move from a single train/test cut to a rolling-window model that adapts to regime shifts as new market and quarterly data arrive.



FUTURE · F4

Full Tadawul Coverage

Extend from 35 to the full main-market universe with sector-conditional stacking heads to handle Banking, Petrochem, and Consumer regimes separately.



FUTURE · F5

Realistic Transaction Costs

Introduce spreads, commissions, and slippage models calibrated to Tadawul microstructure — retest the Sortino under realistic execution frictions.

Conclusions, Discussion

DISCUSSION · ALL FOUR RESEARCH QUESTIONS VALIDATED

RQ1 · FUSION

✓

Fusion beats every single-modality baseline

F1 = 0.8149 vs 0.58 baselines — a decisive +23 F1 pp uplift under the strict chronological test.

RQ2 · TEMPORAL

✓

45-day buffer enforceable & non-trivial

Temporal alignment embedded at the join layer — no downstream leakage possible by construction.

RQ3 · PROFITABILITY

✓

Classification → real portfolio alpha

Sharpe 2.61 · Max DD only −2.63% · Win Rate 57.78% on SAR 100,000 portfolio.

RQ4 · EXPLAINABILITY

✓

Predictions are auditable, not opaque

SHAP reveals a healthy mix of momentum, valuation, profitability & volatility drivers — every trade traceable.

HEADLINE CONCLUSION

A carefully-engineered hybrid framework — with rigorous temporal alignment, calibrated stacking, and portfolio-level evaluation — turns two blind traditions of stock-market prediction into a single, auditable, and profitable one on the Saudi Stock Exchange.

E n d o f P r e s e n t a t i o n

Thank You

شكراً لكم

Questions & Discussion

Malek M. Atarji

ID · 252000474 · MSc Artificial Intelligence · Midocean University

Supervisor · Dr. Amjad Qatish