ML-Driven Trading System for Sonata Software

KSHITIJ 2026 · FinStreet Hackathon · IIT Kharagpur

Aman Behera · IIT Roorkee · Team RobinHoodHashing

Win Rate	Sharpe Ratio	Total Return	Max Drawdown	Profit Factor
88.9%	**6.62**	**1.31%**	**0.05%**	**29.27**

What We Built

Trend-following system that:

- Trades **only** in trend direction
- Enters on **pullbacks**, not reversals
- Uses ML as a **filter**, not predictor
- Prioritizes **capital preservation**

How It Works

Step	Action
1. Trend ID	EMA(10)/SMA(20) + ADX
2. Entry	Wait for 6% pullback
3. Validate	ML confirms direction
4. Size	Scale by confidence
5. Manage	ATR-based stops

Why This Approach

- Sonata showed **continuation patterns**
- Mean-reversion **failed** in backtest
- Pullbacks = **low-risk entries**
- ML filters **bad setups**

ML Component

Ensemble: XGBoost (60%) + LightGBM (40%)

Input	RSI, MACD, ADX features
Output	P(Up) vs P(Down)
Role	Reject Tech $\neq$ ML trades
Training	Walk-forward, no leakage

Risk Management

Control	Rule
Position Size	risk = base $\times$ (0.5 + conf)
Max Risk	5% per trade
Stop Loss	1.5$\times$ ATR
Take Profit	3$\times$ ATR
Trailing	2.5$\times$ ATR
Regime Filter	ADX $<$ 15 $\rightarrow$ no trade

Parameter Optimization

Method: Randomized search (150 configs)

Param	Range	Best
Pullback	6–12%	**6%**
Risk/Trade	2–5%	**3–5%**
Stop ATR	1.2–1.8$\times$	**1.5**$\times$
Target ATR	3–6$\times$	**3**$\times$
ADX Min	12–18	**15**
Trailing	1.5–2.5$\times$	**2.5**$\times$

Objective: Max(Return $\times$ Sharpe) $-$ DD penalty

Backtest Results

Metric	Value	Metric	Value
Period	Nov–Dec 2025	Trades	9
Symbol	SONATSOFTW	Winners	8
Capital	100,000	Win Rate	88.9%
Final	101,314	Sharpe	6.62
Return	1.31%	Max DD	0.05%

Forward Predictions (Jan 1-8, 2026)

XGBoost output on Dec 31, 2025 state:

Date	Signal	Conf	Regime
Jan 1	HOLD	69%	Uptrend
Jan 2	HOLD	69%	Uptrend
Jan 5	HOLD	69%	Uptrend
Jan 6	HOLD	69%	Uptrend
Jan 7	HOLD	69%	Uptrend
Jan 8	HOLD	69%	Uptrend

HOLD = ML direction conflicts with technical regime

Data Integrity & Compliance

Check	Status
No future data in features	
Labels from T+2 to T+5 returns	
Walk-forward validation	
FYERS API only (data + exec)	
Audit script passed	

`FYERS API → Features (RSI, MACD, ADX) → XGBoost+LightGBM → Signal Gate → Risk Mgmt → Execute`

Key Insight: High win rate (88.9%) + tight stops (1.5$\times$ ATR) + minimal drawdown (0.05%) = **sustainable edge**
System compounds small, consistent gains while protecting capital — built for real-world deployment.

Methodology and Execution Details

All market data is sourced exclusively from the **FYERS API** using daily OHLCV candles. Data from **Jan–Oct 2025** is used only for indicator warmup (e.g., moving averages, momentum) and is never used for model training or trade decisions.
Model training and backtesting are restricted to the permitted window from **November 1 to December 31, 2025**. Forward signals for **January 1–8, 2026** are generated as a true out-of-sample forecast.
The strategy is rule-based and structure-driven. Trend direction is identified using the EMA(10)–SMA(20) relationship and its slope, with entries taken only on controlled pullbacks within an established trend to improve risk–reward.

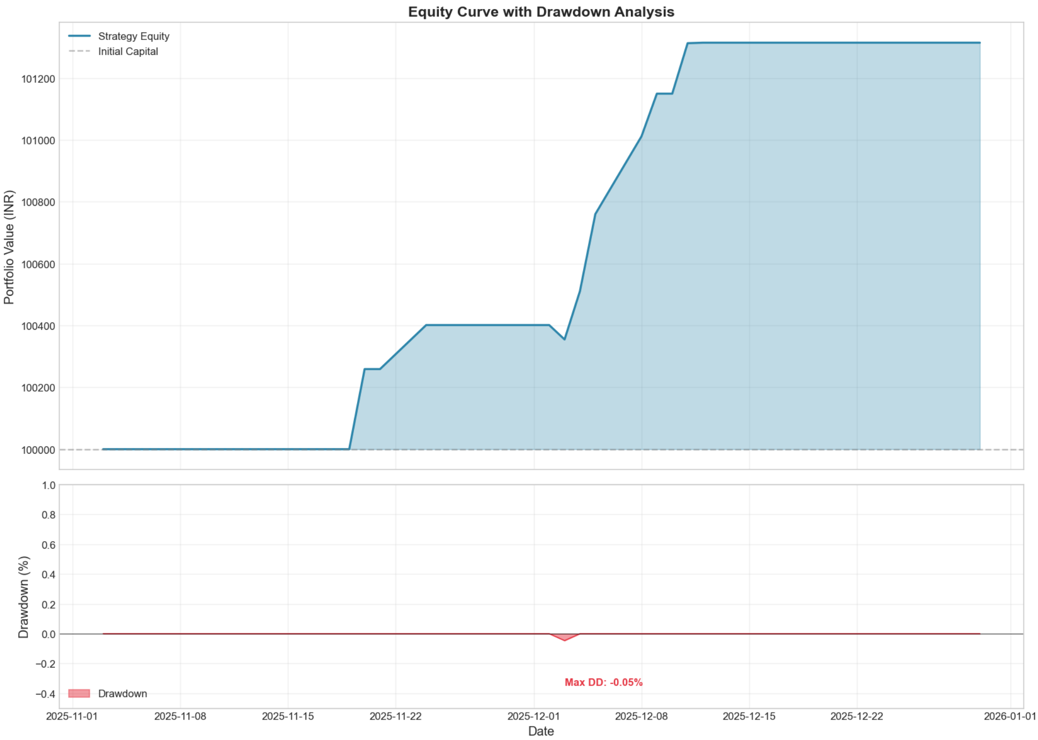

Figure 1: Equity curve and drawdown during the Nov–Dec 2025 backtest period

Machine learning is used only as a validation layer. It estimates directional confidence from recent trend, momentum, volatility, and volume features, and does not predict prices or returns. Trades are taken only when ML and technical signals align; otherwise, they are skipped.
Backtesting follows a strict chronological walk-forward process. Signals are generated after market close and executed at the next session open. Exits are evaluated using intraday high and low prices, with positions closed at end-of-day if neither stop-loss nor target is reached.
A dedicated audit process verifies the absence of look-ahead bias. Features use only past data, rolling indicators respect their lookback windows, and labels are derived from forward returns by design.

Figure 2: Price action with executed trades

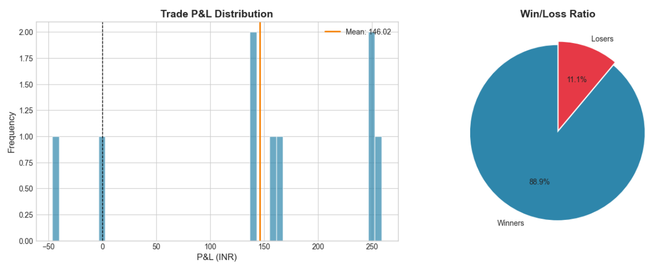

Figure 3: Distribution of trade outcomes

Summary: A rule-based trend-following system enhanced by probabilistic validation, evaluated under strict data constraints with realistic execution and no look-ahead bias.