

# 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

## System Overview

**Trend-following system with:**

- Trades only in trend direction
- Enters on pullbacks, not reversals
- ML ensemble as validation filter
- ATR-based dynamic risk management

## Execution Flow

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. Exit	ATR-based stops

## ML Component

**Ensemble:** XGBoost (60%) + LightGBM (40%)  
**Input** RSI, MACD, ADX, KER features  
**Output** P(Up) vs P(Down)  
**Role** Reject misaligned trades  
**Training** Walk-forward validation

## Parameter Robustness

Parameter	Range	Selected	Sharpe
Pullback	4–10%	6%	5.27–6.61
Stop ATR	1.2–2.0x	1.5x	6.61
Target ATR	2.5–5.0x	3.0x	6.61
ADX Filter	12–20	15	6.61

Only pullback shows sensitivity. Others invariant.

## Risk Management

Control	Rule
Position Size	risk = base × (conf/0.70)
Max Risk	5% per trade
Stop Loss	1.5× ATR
Take Profit	3× ATR
Trailing Stop	2.5× ATR
Regime Filter	ADX < 15 → skip

## Backtest Results

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

## Forward Predictions (Jan 1–8, 2026)

Date	Signal	Dir	Conf	Regime
Jan 1	HOLD	DOWN	68%	Uptrend
Jan 2	HOLD	DOWN	68%	Uptrend
Jan 5	HOLD	DOWN	68%	Uptrend
Jan 6	HOLD	DOWN	68%	Uptrend
Jan 7	HOLD	DOWN	68%	Uptrend
Jan 8	HOLD	DOWN	68%	Uptrend

Model predicts DOWN but regime is Uptrend. Conservative approach: HOLD.

## Data Integrity

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

FYERS API → Features → XGBoost+LightGBM → Signal Gate → Risk Mgmt → Execute

**Key Insight:** High win rate (88.9%) + tight stops (1.5x ATR) + minimal drawdown (0.05%) = sustainable edge. System compounds small, consistent gains while preserving capital.

## 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 and is never used for model training or trade decisions. Model training and backtesting are restricted to **November 1 – December 31, 2025**. Forward signals for **January 1–8, 2026** are generated as true out-of-sample forecasts. The strategy is rule-based and structure-driven. Trend direction is identified using the EMA(10)–SMA(20) relationship and slope, with entries taken only on controlled pullbacks within an established trend. The Kaufman Efficiency Ratio (KER) provides additional trend quality assessment.

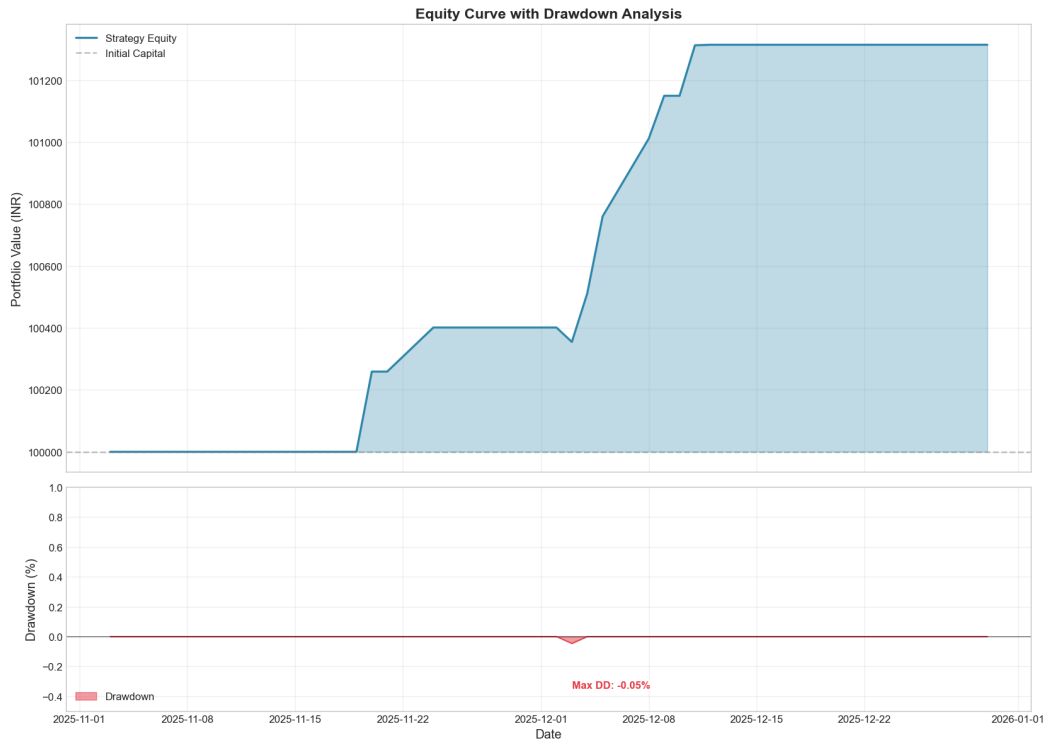


Figure 1: Equity curve and drawdown during the Nov–Dec 2025 backtest period

Machine learning serves as a validation layer, not a predictor. The XGBoost + LightGBM ensemble estimates directional confidence from trend, momentum, volatility, and volume features. Trades execute 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/low prices for realistic stop-loss and take-profit triggers.

A dedicated audit process verifies 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. Triple-barrier labeling uses high/low prices for barrier detection.



Figure 2: Price action with executed trades

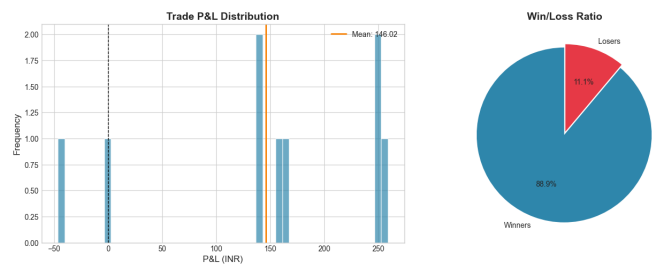


Figure 3: Distribution of trade outcomes

**Summary:** A rule-based trend-following system with probabilistic validation, evaluated under strict data constraints with realistic execution and verified absence of look-ahead bias.