Research Objective: To evaluate the effectiveness of rule-based and machine learning signals in predicting sector trends

This study evaluates whether technical indicators (e.g., RSI, MACD, Coppock Curve) and machine learning models can effectively predict relative sector performance over 5-period horizons. The goal is to determine if a systematic, rules-based strategy using these signals can generate consistent alpha relative to the US market benchmark (US_REBALANCE).

Scope:

Analyzed 25 GICS sectors using daily data from 2004–2025
Built a comprehensive feature set of 30+ technical indicators
Trained 9 machine learning models per sector (Random Forest, XGBoost, SVM, etc.)
Backtested a tactical trading strategy based on model predictions
Measured performance via accuracy, hit rate (win rate), and total return

Definitions

- Rule-based signals: Predefined technical rules (e.g., RSI > 80% for overbought) achieved accuracy between 40%–60%, depending on the sector.
- Model-identified signals: Machine learning (ML) models using technical features, combined with predefined rules, performed better than rules alone.
 Reported accuracy generally ranged from 40%–60%. In some cases, win rates (i.e., proportion of profitable trades) reached up to ~75%.
 - Important clarification: When I previously referenced 75% accuracy, I should have stated this referred to win rate ranges, not overall predictive accuracy.

Classification-Based Technical Trading Strategies Using ML Indicators

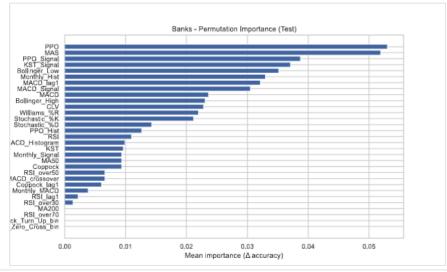
Key Findings

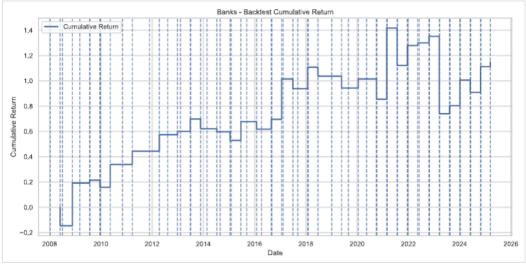
- Rule-based signals alone achieved 40%–60% accuracy.
- ML-enhanced models improved performance, with **win rates up to ~75%** (clarified: refers to trading win rate, not model accuracy).

Performance Report: Sep 2025

Sector	ML_Accuracy	Cumulative_Return	Win_Rate	Num_Trades
Automobiles & Components	0.412	0.078	0.514	37
Banks	0.56	1.15	0.606	33
Capital Goods	0.462	0.624	0.733	30
Commercial & Professional Services	0.478	0.725	0.618	34
Consumer Discretionary Distribution & Retail	0.44	1.009	0.656	32
Consumer Durables & Apparel	0.412	0.61	0.686	35
Consumer Services	0.522	0.672	0.645	31
Consumer Staples Distribution & Retail	0.516	2.198	0.63	27
Energy	0.429	1.412	0.621	29
Equity Real Estate Investment Trusts (REITs)	0.44	-0.018	0.486	37
Financial Services	0.478	0.297	0.667	33
Food, Beverage & Tobacco	0.434	0.554	0.611	36
Health Care Equipment & Services	0.467	0.518	0.541	37
Household & Personal Products	0.555	0.843	0.56	25
Insurance	0.434	1.3	0.724	29
Materials	0.489	0.452	0.697	33
Media & Entertainment	0.566	0.948	0.727	33
Pharmaceuticals, Biotechnology & Life Sciences	0.456	1.84	0.649	37
Real Estate Management & Development (New)	0.516	8.472	0.677	31
Semiconductors & Semiconductor Equipment	0.505	1.809	0.629	35
Software & Services	0.434	1.498	0.649	37
Technology Hardware & Equipment	0.527	0.071	0.529	34
Telecommunication Services	0.599	0.636	0.625	32
Transportation	0.56	1.729	0.643	28
Utilities	0.445	0.839	0.556	36

Pdf report with Feature Importance across Sectors





This research integrates technical analysis with machine learning (ML) to evaluate sector performance relative to a benchmark (US_REBALANCE) and to test systematic trading strategies. The approach combines traditional rule-based signals with predictive models, and validates results through structured backtesting.

TAG: US_REBALANCE_ANALYSIS_CLASSIFIER

Approach

Approaching as a classification problem. Technical features and a Buy/Hold/Sell Signal. Predict t+5 periods(experiments with weekly sector returns)

Data: Sector time series standardized against US_REBALANCE.
 Indicators: RSI, MACD, Bollinger Bands, Coppock Curve, and others.
 Modeling: Random Forest and XGB classifiers trained on lagged/binary features.
 Target: Predict 5-period forward returns (binary classification).

• **Backtesting:** Strategy based on consecutive positive predictions with fixed holding periods.

Definitions

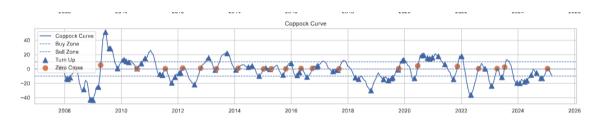
Rule-Based vs. ML-Based vs. Combined

- Rule-Based = fixed cutoffs (transparent but rigid)
 Example: RSI > 70 → Sell, RSI < 30 → Buy, Coppock > 0 → Bullish.
- ML-Based = feature-driven learning (adaptive, finds hidden patterns)
 Example: Learns that "Coppock turning up + RSI rising from oversold + price near lower Bollinger Band → high chance of positive return."
 Capture non-linear patterns.

Combined rule-based + ML approaches delivered **better classification and stronger backtest results**.

Link

Features explored



Feature Type	Example Feature(s)	Typical Signal Output	Interpretation / Use	Binary Signal Example
Moving Averages (MA)	SMA, EMA (various windows: 10, 21, 50, 100, 200)	Continuous	Trend direction, smoothing	-

MA Crossovers	Golden/Death Cross (50 > 200)	Binary/Flag	Major trend regime shifts	1 if short MA > long MA (Golden Cross)
Rate of Change (ROC)	n-day ROC	Continuous	Momentum strength	-
Relative Strength Index (RSI)	RSI value (0–100), overbought/oversold bands	Continuous & Binary	Momentum, reversal risk	1 if RSI > 70 (overbought), <30 (oversold)
MACD	MACD diff, Signal, Histogram	Continuous & Binary	Trend shifts, cross confirmations	1 if MACD > Signal (bullish cross)
Know Sure Thing (KST)	KST, Signal, Cross	Continuous & Binary	Long-term trend strength	1 if KST > Signal
Bollinger Bands	Band upper/lower, %b, band width	Continuous & Binary	Volatility, breakout risk	1 if Close > Upper Band
Rolling Std Dev	Std dev (5, 21, 63 days)	Continuous	Short/long-term volatility	-
Average True Range (ATR)	ATR value	Continuous	Volatility, position sizing	-
High-Low Range	High minus Low	Continuous	Daily/intra-bar volatility	-

Stochastic Oscillator	%K, %D, Overbought/Oversold cross	Continuous & Binary	Momentum extremes, cycles	1 if %K crosses > %D (bull/bear signal)
Williams %R	%R value	Continuous & Binary	Fast momentum/reversal	1 if %R < -80 (oversold), > -20 (overbought)
Commodity Channel Index (CCI)	CCI value	Continuous & Binary	Cyclical trend/mean-reversion	1 if CCI > 100 or < -100
Momentum Oscillator	Price[today] vs. Price[n-bars ago]	Continuous	Raw momentum, crossover signals	-
New High/New Low	1 if new N-bar high/low	Binary	Breakout or mean reversion alert	1 if Close = rolling max/min
Drawdown Flags	(Close/rolling max)-1	Continuous & Binary	Mean reversion/extension	1 if drawdown > threshold
Percent Distance from MA	(Close/MA_n)-1	Continuous & Binary	Overextension from trend	1 if distance exceeds X%

None

US_REBALANCE_ANALYSIS_CLASSIFIER

Available sectors:

- Automobiles & Components
- Banks
- Capital Goods
- Commercial & Professional Services
- Consumer Discretionary Distribution & Retail (New Name)
- Consumer Durables & Apparel
- Consumer Services
- Consumer Staples Distribution & Retail (New Name)
- Energy
- Equity Real Estate Investment Trusts (REITs) (New Name)
- Financial Services (New Name)
- Food, Beverage & Tobacco
- Health Care Equipment & Services
- Household & Personal Products
- Insurance
- Materials
- Media & Entertainment
- Pharmaceuticals, Biotechnology & Life Sciences
- Real Estate Management & Development (New)
- Semiconductors & Semiconductor Equipment
- Software & Services
- Technology Hardware & Equipment
- Telecommunication Services
- Transportation
- Utilities

Compare multiple Machine Learning Models performance on Technical Signals

Notable Findings

Hit Rate ≠ Model Accuracy

Financial Services: Model Accuracy = 42.3%, Hit Rate = 66.7%

Energy: Model Accuracy = 34.5%, Hit Rate = 27.3%

Software & Services: Model Accuracy = 40.5%, Hit Rate = 60.0%

This is critical: A low-accuracy model can still produce profitable trades if losses are small and wins are large.

Top Features consistently included:

RSI_lag1, MACD_crossover, Coppock_lag1, Momentum_5, Volatility_10

Some Sectors Are Unpredictable

Consumer Discretionary Retail: Hit Rate = 0.0% despite 42.3% accuracy

Energy: Low accuracy and hit rate → poor signal quality Utilities: Low volatility → few profitable opportunities

The goal of this research was to:

Predict relative outperformance of equity sectors vs. the broad market (US_REBALANCE) Use technical indicators (RSI, MACD, Coppock, Bollinger Bands, etc.) as input features Train multiple ML models to classify future 5-day return direction Backtest a tactical trading strategy and evaluate its realized performance using hit rate (win rate) and total return

Methodology

Data Source: TechnicalData.xlsx (daily pricing for 25 GICS sectors)
Benchmark: US REBALANCE (equal-weighted market index)

Relative Pricing: All sector prices normalized by US REBALANCE → relative price = sector /

index

Sectors Analyzed: 25 sectors including Technology, Financials, Healthcare, Energy, etc.

Feature Engineering

We constructed a rich set of technical features:

Trend: MA50, MA200, MACD, PPO

Momentum: RSI, Stochastic, Williams %R, Rate of Change

Volatility: Bollinger Bands, Volatility (5/10/20-day)

Reversal: Coppock Curve (with turn-up and zero-cross signals)

Market Regime: Monthly MACD, KST, MAS, CLV

Binary Signals: RSI > 30/50/70, MACD crossover, Coppock_Turn_Up

2.3 Target Definition

Classification Task: Predict 5-day forward return direction

0 = Negative return

 $1 = Neutral return (\pm 1\%)$

2 = Positive return

Time-Series Split: 3-fold TimeSeriesSplit to avoid lookahead bias

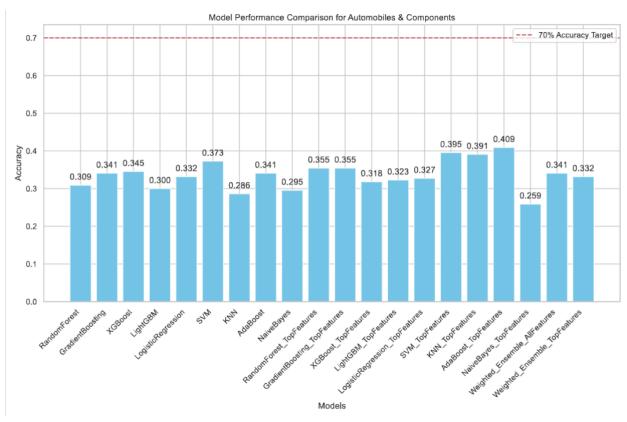
**	_	_	_
Sector	Best_Model	Accuracy	Hit_Rate
Automobiles & Components	AdaBoost_TopFeatures	0.409	0.407
Banks	AdaBoost	0.418	0.486
Capital Goods	NaiveBayes_TopFeatures	0.486	0.471
Commercial & Professional Services	LogisticRegression	0.505	0.538
Consumer Discretionary Distribution & Retail	Weighted_Ensemble_TopFeatures	0.423	0
Consumer Durables & Apparel	SVM	0.505	0.472
Consumer Services	AdaBoost	0.414	0.52
Consumer Staples Distribution & Retail	SVM	0.527	0.536
Energy	NaiveBayes_TopFeatures	0.345	0.273
Equity Real Estate Investment Trusts (REITs)	KNN_TopFeatures	0.505	0.403
Financial Services	NaiveBayes_TopFeatures	0.423	0.667
Food, Beverage & Tobacco	SVM	0.464	0.537
Health Care Equipment & Services	GradientBoosting_TopFeatures	0.405	0.561
Household & Personal Products	LogisticRegression	0.459	0.469
Insurance	AdaBoost_TopFeatures	0.414	0.52
Materials	LightGBM_TopFeatures	0.436	0.465
Media & Entertainment	GradientBoosting_TopFeatures	0.423	0.57
Pharmaceuticals, Biotechnology & Life Sciences	XGBoost	0.541	0.518
Real Estate Management & Development (New)	NaiveBayes_TopFeatures	0.486	0.556
Semiconductors & Semiconductor Equipment	NaiveBayes	0.445	0.556
Software & Services	SVM	0.405	0.6
Technology Hardware & Equipment	XGBoost_TopFeatures	0.382	0.609
Telecommunication Services	NaiveBayes	0.468	0.406
Transportation	KNN_TopFeatures	0.468	0.5
Utilities	AdaBoost_TopFeatures	0.418	0.5

Link to summary report

Machine Learning Models

We trained 9 classifiers per sector:

Random Forest, XGBoost, LightGBM, Gradient Boosting Logistic Regression, SVM, KNN, AdaBoost, Naive Bayes We also created:



Sector wise feature importance

Top-Features Models: Trained on top 5 features via Random Forest importance Weighted Ensemble: Model combination using accuracy-based weights

Backtesting & Performance Metrics

For each sector:

Strategy: Enter long if model predicts "positive return", exit when signal flips

Position: Position = Signal.shift(1) → no lookahead

Key Metrics:

Model Accuracy: % correct predictions

Hit Rate (Win Rate): % of trades with positive P&L

Total Return: Cumulative strategy return

Feature Importance: Top drivers of predictions

Non Machine Learning: Vote-Based Trade Signal Using Coppock Scores

TAG: COPPOCK_VOTE_STRATEGY

1. Introduction

This study explores a non–machine learning trading strategy driven by Coppock scores derived from daily price data. Coppock scores, which combine rate-of-change calculations with a weighted moving average, are used here as a proxy for long-term momentum. In our approach, trading signals are generated when the Coppock score exceeds a defined positive threshold (buy signal) or falls below a negative threshold (sell signal). The strategy is then backtested, and key performance metrics (total return, CAGR, Sharpe ratio, and maximum drawdown) are computed. In addition, the Coppock curve is plotted with overlaid buy and sell markers for visual interpretation.

2. Data and Methodology

Data Loading and Preparation

Data Source:

The strategy uses daily security-level price data, with each asset identified by its SEDOL.

• Pivoting:

The dataset is pivoted so that each SEDOL becomes a separate column representing the relative price

Coppock Score Calculation

Weighted Moving Average (WMA):

A custom WMA function is defined to calculate weights for a rolling window.

Coppock Calculation:

For a given price series, the 11-day and 14-day rates of change (ROC) are computed. Their sum is then smoothed with a 10-day WMA to generate the Coppock score. The final "total_copp_score" is set equal to this Coppock score.

Trading Signal Generation

Value-Based Signals:

A trading signal is generated according to the following rule:

- Buy Signal: When total_copp_score > 5Sell Signal: When total_copp_score < -5
- Neutral: Otherwise

The signal is shifted by one day (to avoid lookahead bias) and applied as the trading position for the next day.

Backtesting

Daily Returns:

The strategy computes daily percentage returns using the asset's price.

Strategy Returns:

The daily strategy return is the product of the shifted signal (position) and the daily return.

Cumulative Return and Performance Metrics:

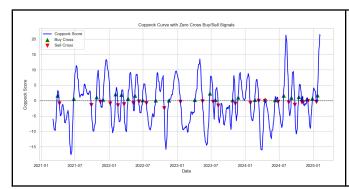
The cumulative strategy return is computed as the cumulative product of 1 plus daily strategy returns. Key performance metrics include:

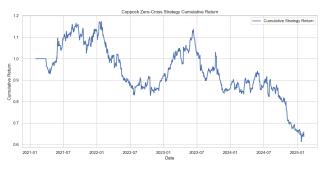
- o **Total Return:** Final portfolio value minus one.
- CAGR (Compound Annual Growth Rate): Annualized return computed from the cumulative return.
- Sharpe Ratio: Risk-adjusted performance based on the mean and standard deviation of strategy returns.
- o Maximum Drawdown: The worst observed decline in the portfolio value.

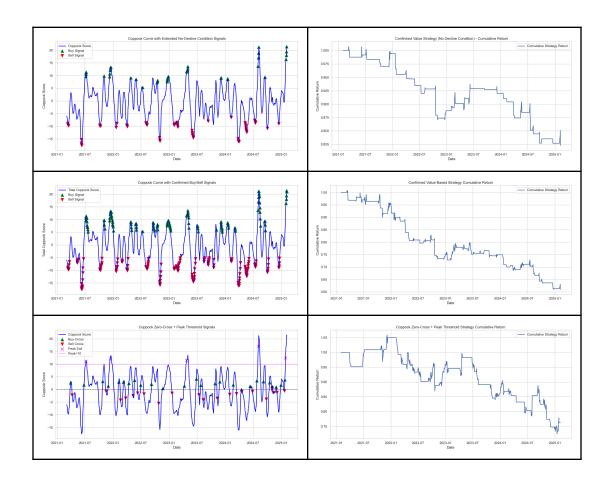
Signal Plotting

Overlaying Signals:

The Coppock curve is plotted over time, and buy signals (when the Coppock score exceeds 5) are marked with green up-arrows while sell signals (when the score is below –5) are marked with red down-arrows. This overlay facilitates a visual review of when the strategy would enter or exit a trade.





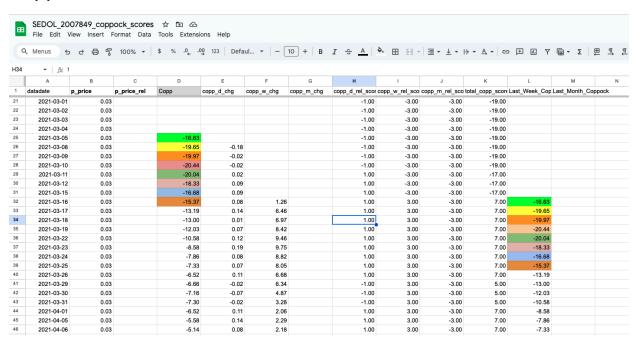


Strategy Entry Rule Exit Rule **Short Logic?** Pros Cons 1) Basic - Buy if Simple to May buy/sell at - Sell if Optional Value-Based Coppock > implement local extremes Coppock < Captures quick (Absolute Frequent trades **+**X -X Peaks) moves

2) Confirmation-B ased (Consecutive Days)	- Must hav Coppock (or < -X) for consecutiv	> +X	- Exit if cond fails (or the opposite consecutive triggers)		Optional	Reduces whipsaws Ensures momentu sustained	s ım is	Can miss sudden reversals More complex logic
No-Decli but a ne Copp	ilar to #2 Iso require lock not to on final day		utive rule or pative slope	Optiona		•	pull	miss partial packs that me upward
Zero-Cross fi	Buy on cross rom below 0 delbove 0	to fron	ell on cross n above 0 to ow 0	Option	portion trend	le, classic	ea of La	isses arliest part move ate to exit if oss is slow
Zero-Cross c	ross from ⇒ 0 to ≥0	- Exit if cro >0 to ≤0 c Coppock (peak)	or if	Optional	indicato overhea	r its id-based	early it	exit too f the market rising es threshold
Zero-Cross I + Peak (With s	ong on		ock > +X g, flip to shor ely	Ye t s	Always i market Profits fr downtre	or om si nds R	•	ling in s markets s robust

down

Coppock features used



Link

Inputs:

Data: A dataset containing historical prices (e.g., p price rel or close) and a datadate column.

Price Column: The column to use for calculations (e.g., p_price_rel or close).

Window Sizes:

Short-term ROC window: 11 periods.

Long-term ROC window: 14 periods.

Weighted Moving Average (WMA) window: 10 periods.

Weekly change window: 7 periods.

Monthly change window: 30 periods.

Outputs:

Coppock Scores: A DataFrame containing:

Copp: The Coppock Curve value.

copp_d_chg: Daily change in Coppock Curve.

copp w chg: Weekly change in Coppock Curve.

copp m chg: Monthly change in Coppock Curve.

copp_rel_score: Daily score based on Coppock change.

copp_w_rel_score: Weekly score based on Coppock change.

copp m rel score: Monthly score based on Coppock change.

total copp score: Total weighted Coppock score.

Last_Week_Coppock: Coppock value 7 periods ago.

Last_Month_Coppock: Coppock value 30 periods ago.

```
Python
ALGORITHM Coppock_Score_Calculation
INPUT: Dataset, Price_Column

    Load Dataset

   - Convert `datadate` to datetime
   - Set `datadate` as index
2. Define WMA Function
   - weights = [1, 2, \ldots, window]
   - WMA = (sum(x_i * w_i) / sum(w_i)
3. Calculate Coppock Curve
   - ROC11 = (Price_t / Price_t-11 - 1) * 100
   - ROC14 = (Price_t / Price_t-14 - 1) * 100
   - Sum ROC = ROC11 + ROC14
   - Copp = WMA_10(Sum ROC)
   - Copp_Change = (Copp_t - Copp_t-1) / |Copp_t-1|
4. Calculate Weekly and Monthly Changes
   - copp_w_chg = (Copp_t - Copp_t-7) / |Copp_t-7|
   - copp_m_chg = (Copp_t - Copp_t-30) / |Copp_t-30|
```

Appendix

Feature Importances across Sectors

Sector Key Indicators Rationale

Automobiles & Components	EMA_21, SMA_50, BB_Lower, BB_Upper	Sensitive to short-to-medium-term trends; volatility significant

Banks	EMA_21, BB_Lower	Sensitive to interest rates and economic cycles
Capital Goods	MAS, SMA_50	Tied to longer economic cycles & infrastructure spending
Commercial & Professional Services	SMA_50, MAS	Benefits from medium-term trend following and momentum tracking
Consumer Discretionary (Distribution & Retail)	BB_Upper, BB_Lower	Consumer spending is volatile
Consumer Durables & Apparel	BB_Lower, SMA_50	Sensitive to consumer demand; focus on downside protection
Consumer Services	SMA_50, EMA_21	Blend of medium and short-term trends is strongest predictor
Consumer Staples (Distribution & Retail)	MAS, SMA_50	Stable sector; moving averages more relevant than Bollinger Bands
Energy	MAS, SMA_50	Follows long-term supply-demand cycles

Equity Real Estate Investment Trusts (REITs)	MAS, MACD	Interest rate-sensitive and trend-driven
Financial Services	SMA_50, MAS	Similar to banks but with broader market exposure
Food, Beverage & Tobacco	MAS, RSI	Defensive sector; relies on overbought/oversold signals
Healthcare Equipment & Services	SMA_50, BB_Upper	Growth sector; breakouts to new highs are key signals
Household & Personal Products	EMA_21, SMA_50	Stable, defensive sector
Insurance	SMA_50, BB_Upper	Value-oriented, stable sector
Materials	BB_Lower, SMA_50	Cyclical sector tied to economic growth
Media & Entertainment	MAS, BB_Upper	Volatile, trend-driven sector
Pharmaceuticals, Biotechnology & Life Sciences	SMA_50, BB_Lower	Mix of defensive plays and high-growth opportunities

Real Estate Management & Development	MAS, SMA_50	Similar to REITs, focus on development projects
Semiconductors & Equipment	SMA_50, BB_Upper	Cyclical, high-growth sector
Software & Services	RSI, MAS	High-growth, high-momentum sector
Technology Hardware & Equipment	MAS, BB_Lower	Cyclical industry with strong downside risk
Telecommunication Services	BB_Upper, RSI	Defensive sector with potential growth spurts
Transportation	BB_Lower, MAS	Highly cyclical sector tied to economic trends
Utilities	SMA_50, EMA_21	Stable, defensive sector