A GUIDE TO IMPROVING THE PERFORMANCE OF TRADING STRATEGIES

By Md. Nafis Rabbi
ID: 22366046

Binary VS Continuous

Binary

Crossovers: A crossover occurs when one indicator crosses above or below another indicator. For example, a bullish crossover would occur when a short-term moving average crosses above a long-term moving average.

Breakouts: A breakout occurs when a stock price moves outside of a trading range. For example, a bullish breakout would occur when a stock price moves above a resistance level.

Divergences: A divergence occurs when the price of a stock and a stock indicator are moving in opposite directions. For example, a bullish divergence would occur when a stock price is making lower lows, but a stock indicator is making higher lows.

Continuous

Indicators: Stock indicators can also generate numerical values, which can be used to quantify the strength of a signal. For example, the RSI can be used to generate a numerical value between 0 and 100. A higher RSI value indicates that the stock is more overbought, while a lower RSI value indicates that the stock is more oversold.

Example: Another example is the Bollinger Bands. The Bollinger Bands %B is a measure of how far the stock price is from the upper Bollinger Band. A higher Bollinger Bands %B value indicates that the stock price is closer to the upper Bollinger Band, which is a sign of overbought conditions. A lower Bollinger Bands %B value indicates that the stock price is closer to the lower Bollinger Band, which is a sign of oversold conditions. Numerical values from stock indicators can be used to create more sophisticated trading strategies. For example, a trader might use the RSI to filter out trades when the RSI is above a certain level.

KNN Implementation With Binary Values

Bollinger Bands are a volatility indicator that consists of a moving average (typically a 20-period simple moving average) and two standard deviation bands above and below the moving average. The bands widen and contract as volatility increases and decreases, respectively. Bollinger Bands can be used to identify overbought and oversold conditions, as well as potential trend reversals.

Keltner Channels are another volatility indicator that consists of a moving average (typically a 20-period exponential moving average) and two bands that are a multiple of the average true range (ATR) above and below the moving average. The ATR is a measure of volatility that takes into account the highs, lows, and closing prices of a security over a given period of time. Keltner Channels are similar to Bollinger Bands, but they are more responsive to changes in volatility.

The Relative Strength Index (RSI) is a momentum indicator that measures the magnitude of recent price changes to evaluate overbought or oversold conditions in the price of a stock or other asset. The RSI is displayed as an oscillator (a line graph that moves between two extremes) and can range from 0 to 100. Traditionally, and most commonly, the RSI is interpreted as being overbought when above 70 and oversold when below 30. However, these values can vary depending on the asset being traded and the time frame being used.

Accuracy	Precision	Recall	F1 Score	Features	Total Features	Threshold
0.53000000	1	0.53000000	0.53000000	[sma, ema, chaikin_oscillators, macd_histogram...	8	0.30000000
0.53000000	1	0.53000000	0.50000000	[sma, ema, chaikin_oscillators, williams_perce...	6	0.40000000
0.52000000	1	0.52000000	0.52000000	[sma, ema, chaikin_oscillators, macd_histogram...	9	0.10000000
0.52000000	1	0.52000000	0.52000000	[sma, ema, chaikin_oscillators, macd_histogram...	9	0.20000000
0.49000000	1	0.49000000	0.37000000	[williams_percent_r]	1	0.50000000

Training With Selected Binary Features

Model Training

Model	Accuracy	Precision	Recall	F1
random_forest	0.924731	0.933636	0.924731	0.925487
svm	0.922939	0.931739	0.922939	0.923682
decision_tree	0.922939	0.932357	0.922939	0.923732
logistic_regression	0.879928	0.896386	0.879928	0.880183
naive_bayes_gaussian	0.818996	0.828011	0.818996	0.818925

KNN Implementation With Continuous Values

FSO stands for **Fast Stochastic Oscillator**. It is a momentum indicator that compares the closing price of a security to its price range over a given period of time. The FSO is displayed as two lines, %K and %D. The %K line is the faster of the two lines and is typically calculated using a 5-period lookback period. The %D line is the slower of the two lines and is typically calculated using a 3-period moving average of the %K line.

Stochastic Oscillator is a momentum indicator that compares the closing price of a security to its price range over a given period of time. The Stochastic Oscillator is displayed as two lines, %K and %D. The %K line is the faster of the two lines and is typically calculated using a 14-period lookback period. The %D line is the slower of the two lines and is typically calculated using a 3-period moving average of the %K line.

Williams Percent R is a momentum indicator that compares the closing price of a security to its highest high and lowest low over a given period of time. The Williams Percent R is displayed as a single line that can range from -100 to 0. Traditionally, and most commonly, the Williams Percent R is interpreted as being overbought when above -20 and oversold when below -80. However, these values can vary depending on the asset being traded and the time frame being used.

Accuracy	Precision	Recall	F1 Score	Features	Total Features	Threshold
0.61000000	1	0.61000000	0.61000000	[williams_percent_r, stochastic_oscillator, rs...	4	0.40000000
0.54000000	1	0.54000000	0.54000000	[williams_percent_r]	1	0.50000000
0.54000000	1	0.54000000	0.54000000	[williams_percent_r]	1	0.60000000
0.48000000	1	0.48000000	0.48000000	[chaikin_oscillator, macd_histogram, williams_...	6	0.10000000
0.48000000	1	0.48000000	0.48000000	[chaikin_oscillator, macd_histogram, williams_...	6	0.20000000
0.48000000	1	0.48000000	0.48000000	[chaikin_oscillator, macd_histogram, williams_...	6	0.30000000

Training With Selected Continuous Features

Model Training

Model	Accuracy	Precision	Recall	F1
svm	0.928322	0.937590	0.928322	0.929100
random_forest	0.923077	0.933238	0.923077	0.923918
decision_tree	0.919580	0.930127	0.919580	0.920447
logistic_regression	0.877622	0.895279	0.877622	0.877900
naive_bayes_gaussian	0.805944	0.819044	0.805944	0.805940

Comparison of Binary VS Continuous

Model	Accuracy	Precision	Recall	F1
random_forest	0.924731	0.933636	0.924731	0.925487
svm	0.922939	0.931739	0.922939	0.923682
decision_tree	0.922939	0.932357	0.922939	0.923732
logistic_regression	0.879928	0.896386	0.879928	0.880183
naive_bayes_gaussian	0.818996	0.828011	0.818996	0.818925

Binary

Model	Accuracy	Precision	Recall	F1
svm	0.928322	0.937590	0.928322	0.929100
random_forest	0.923077	0.933238	0.923077	0.923918
decision_tree	0.919580	0.930127	0.919580	0.920447
logistic_regression	0.877622	0.895279	0.877622	0.877900
naive_bayes_gaussian	0.805944	0.819044	0.805944	0.805940

Continuous

Backtesting Binary VS Continuous

Start date	2021-09-29
End date	2023-04-04
Total months	26
	Backtest
Annual return	14.363%
Cumulative returns	34.248%
Annual volatility	25.761%
Sharpe ratio	0.65
Calmar ratio	0.56
Stability	0.31
Max drawdown	-25.48%
Omega ratio	1.14
Sortino ratio	0.95
Skew	-0.03
Kurtosis	3.98
Tail ratio	1.02
Daily value at risk	-3.179%

Start date	2021-09-15
End date	2023-04-04
Total months	27
	Backtest
Annual return	20.302%
Cumulative returns	51.569%
Annual volatility	26.262%
Sharpe ratio	0.83
Calmar ratio	0.82
Stability	0.38
Max drawdown	-24.651%
Omega ratio	1.18
Sortino ratio	1.26
Skew	0.17
Kurtosis	4.19
Tail ratio	1.09
Daily value at risk	-3.222%

Combining Binary & Continuous

Model	Accuracy	Precision	Recall	F1
decision_tree	1.000000	1.000000	1.000000	1.000000
random_forest	0.985294	0.985523	0.985294	0.985296
svm	0.972689	0.972943	0.972689	0.972705
logistic_regression	0.962185	0.962293	0.962185	0.962161
naive_bayes_gaussian	0.710084	0.754302	0.710084	0.709467

Backtesting Combined Features Prediction

Annual return	7.594%
Cumulative returns	13.27%
Annual volatility	24.7%
Sharpe ratio	0.42
Calmar ratio	0.24
Stability	0.53
Max drawdown	-31.948%
Omega ratio	1.08
Sortino ratio	0.62
Skew	0.18
Kurtosis	3.60
Tail ratio	1.15
Daily value at risk	-3.071%

Conclusion

➔ Numerical features produce excellent backtesting results.
➔ Binary features produce poor backtesting results.
➔ Combining numerical and binary features also produces poor backtesting performance, but it improves model training.
➔ When choosing features for backtesting, it is important to prioritize numerical features. Binary features should be used with caution, as they can produce poor results. Combining numerical and binary features can also produce poor backtesting performance, but it can improve model training.