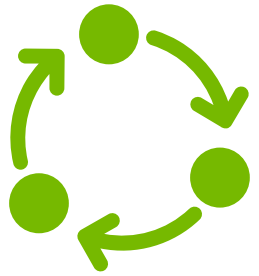
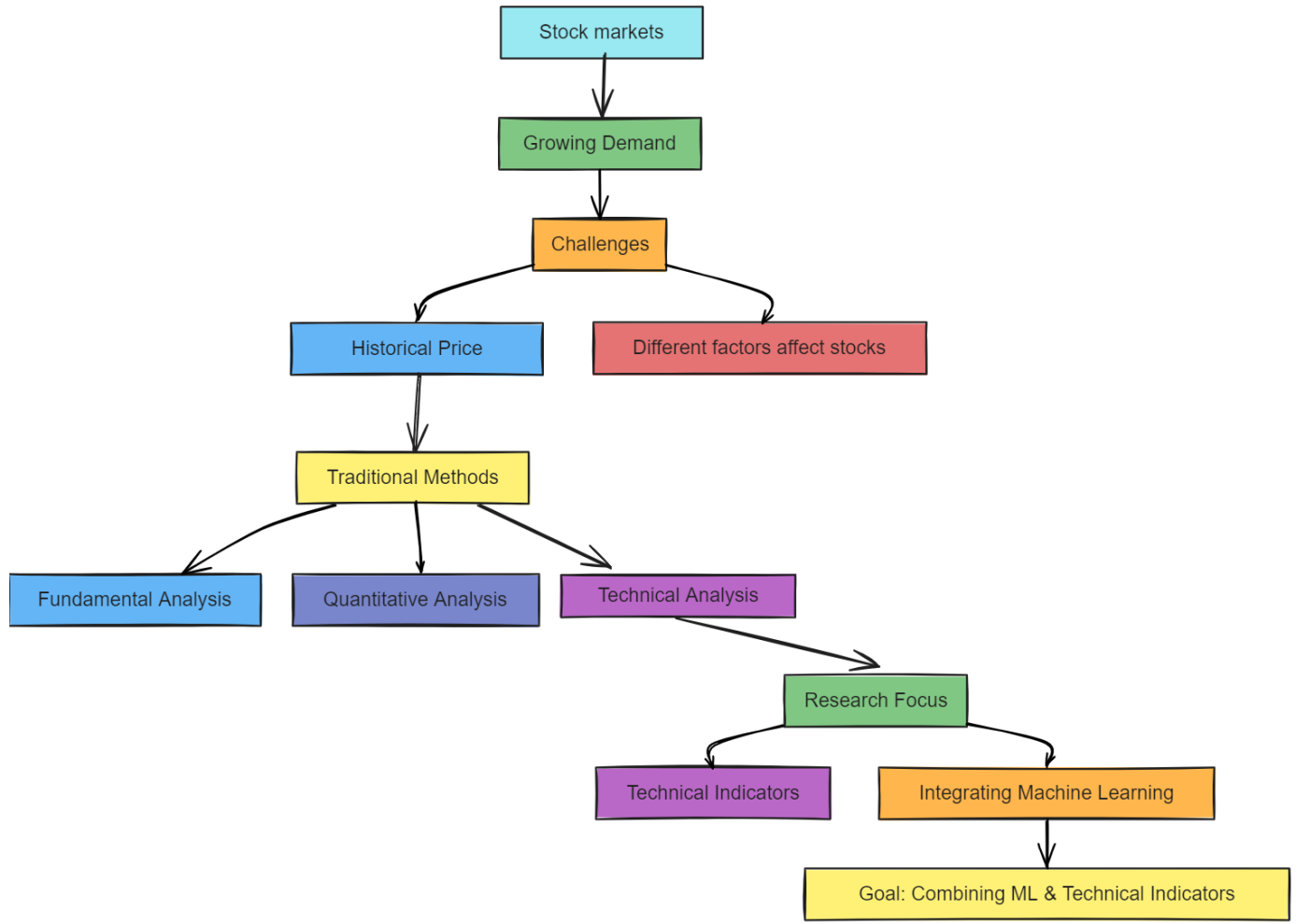
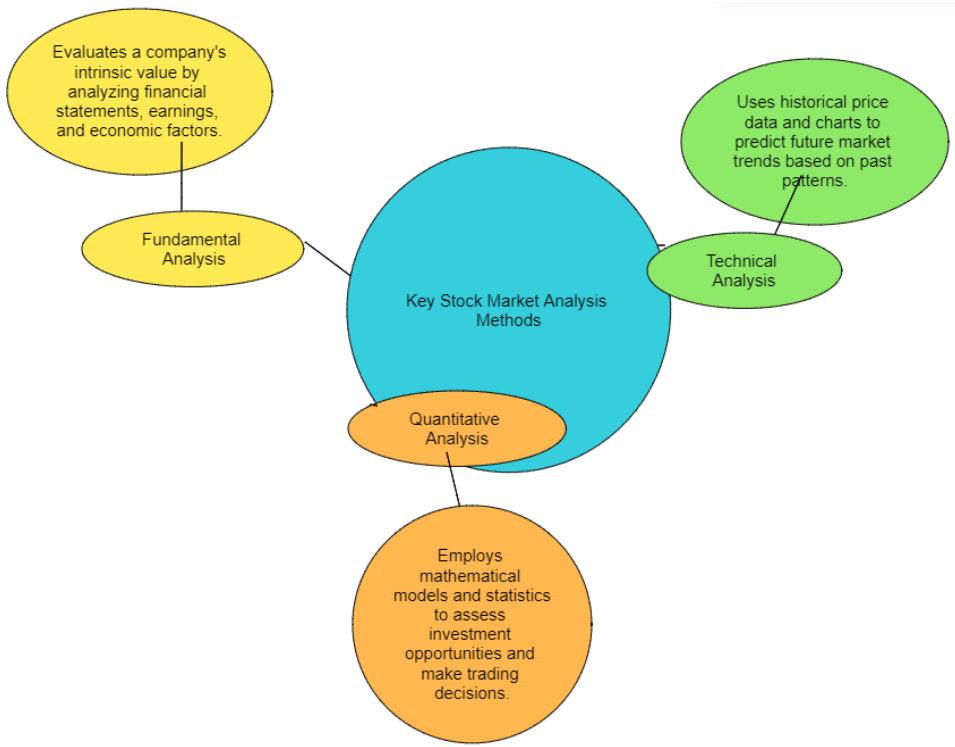
Financial Trading Application

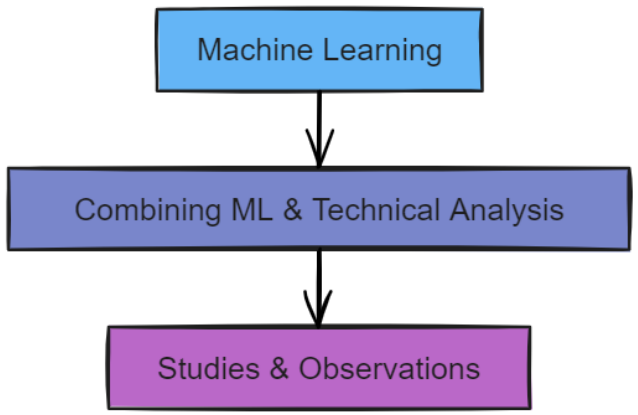
Financial Technical indicatrors + Machine Learning Classifiers for Financial Trading

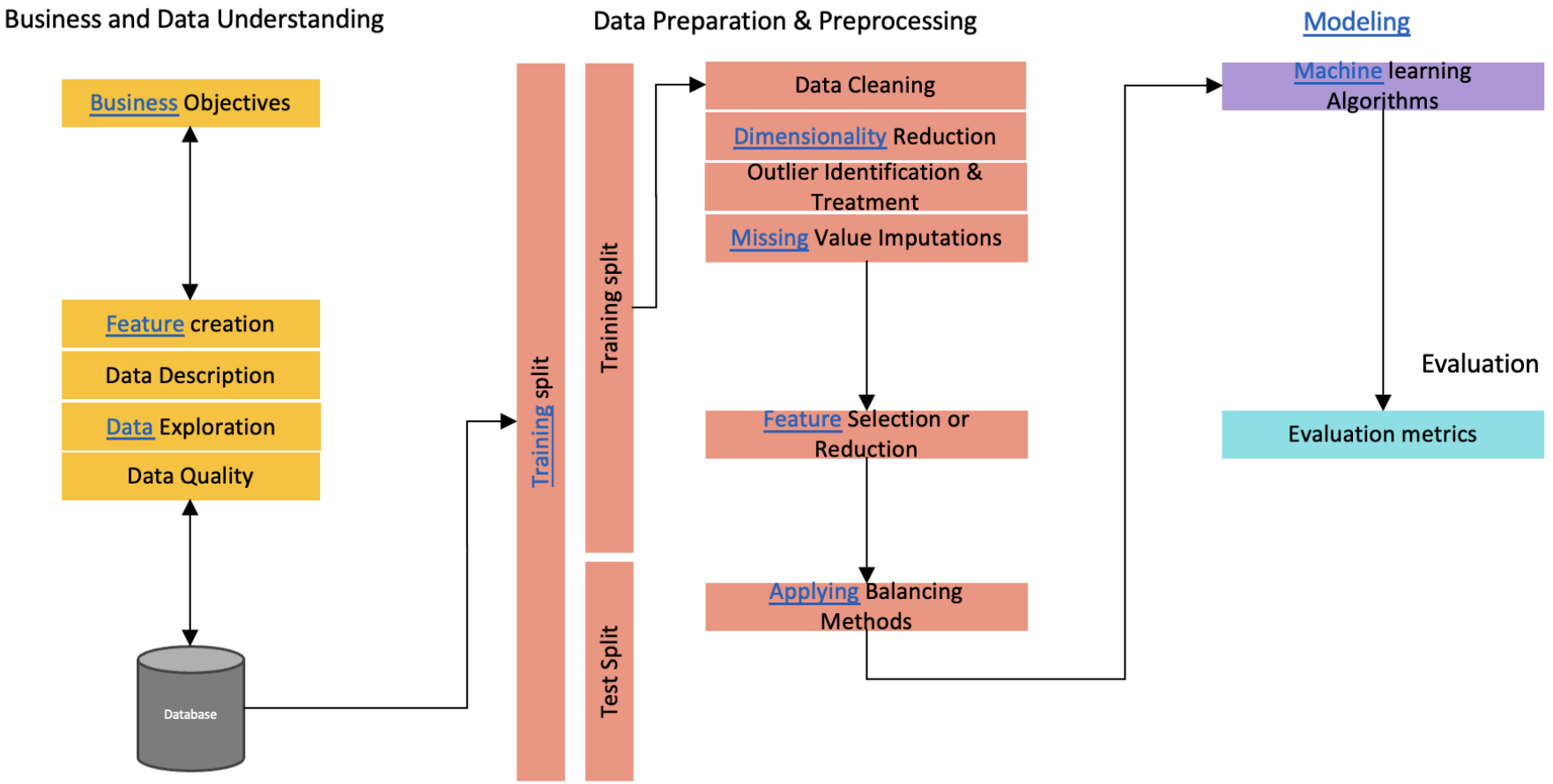
- Table of Content




Stock markets
Growing Demand
Challenges
Historical Price
Different factors affect stocks
Traditional Methods
Fundamental Analysis
Quantitative Analysis
Technical Analysis
Research Focus
Technical Indicators
Integrating Machine Learning
Goal: Combining ML & Technical Indicators


Evaluates a company's intrinsic value by analyzing financial statements, earnings, and economic factors.
Fundamental Analysis
Key Stock Market Analysis Methods
Uses historical price data and charts to predict future market trends based on past patterns.
Technical Analysis
Quantitative Analysis
Employs mathematical models and statistics to assess investment opportunities and make trading decisions.


Machine Learning
Combining ML & Technical Analysis
Studies & Observations


Business and Data Understanding
Data Preparation & Preprocessing
Modeling
Business Objectives
Feature creation
Data Description
Data Exploration
Data Quality
Database
Training split
Training split
Test Split
Data Cleaning
Dimensionality Reduction
Outlier Identification & Treatment
Missing Value Imputations
Feature Selection or Reduction
Applying Balancing Methods
Machine learning Algorithms
Evaluation
Evaluation metrics

Research question:

How to create a robust and replicable machine learning model to different stock markets using technical indicators?

Objective of this project

• Create a machine learning model capable of predicting the trends and classify in the 3 classes called "buy," "hold," or "sell" for a given stock. This model will have as input data technical indicators calculated from the stock opening price to identify market movements.

• Select the model that better predicts the opening prices according to the evaluation metrics.

• Compare the winning model to a base model.

West Texas Intermediate (WTI oil)

This stock, also known as Texas light sweet, is a benchmark for crude oil globally, it is used as a reference for oil prices and trading activities around the world for its quality.

ETF iShares MSCI World Index

This stock incorporates an exchange-traded fund (ETF). This ETF seeks to track the performance of the MSCI World Index, which measures the performance of different companies in developed countries.

Meta

Also known as Facebook, it is a recognized company in the technology sector, it serves as a reference to review the behavior of other companies in the technology sector, its focus is technological, which is why they invest in innovation, virtual and augmented reality.

Business and data understanding

Yahoo Finance

Dataset:

- Date
- Open Price
- High
- Low
- Close
- Volume
- Adj Close

Indicator

- RSI (Relative Strength Index)
- MACD (Moving Average Convergence Divergence)
- Moving Averages
- Stochastic Oscillator
- EMA (Exponential Moving Average)
- Bollinger Bands
- Ichimoku Cloud
- Standard Deviation

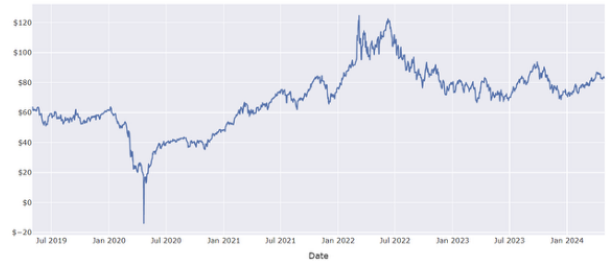
$$Target_t = \begin{cases} 1 \; if \; Daily\_change_t > Sell\_threshold \\ 2 \; if \; Daily\_change_t < Buy\_threshold \\ 0 \; otherwise \end{cases}$$

$$Daily\_Change_t = \left(\frac{\text{Open}_t - \text{Open}_{t-1}}{\text{Open}_{t-1}} \right) \times 100$$

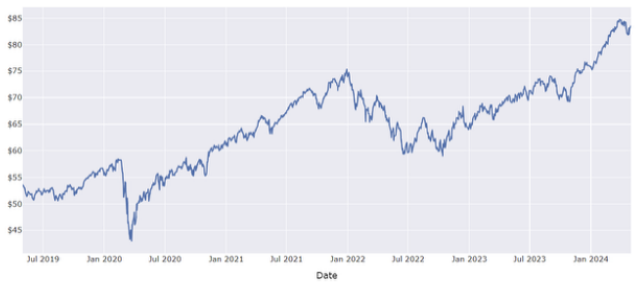
sell_threshold=1.0
buy_threshold=-1.0

Business and data understanding

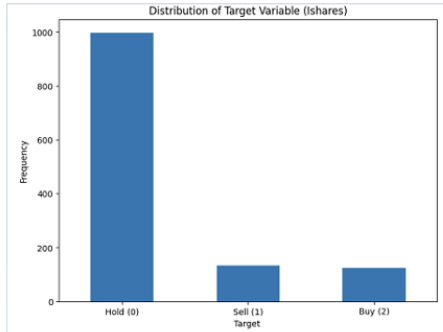
West Texas Intermediate (WTI oil)

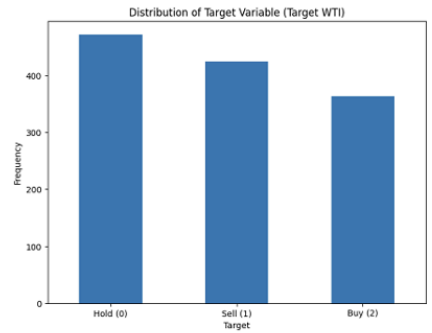


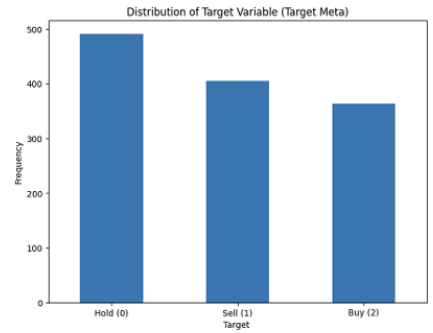
ETF iShares MSCI World Index



Meta







Business and data understanding

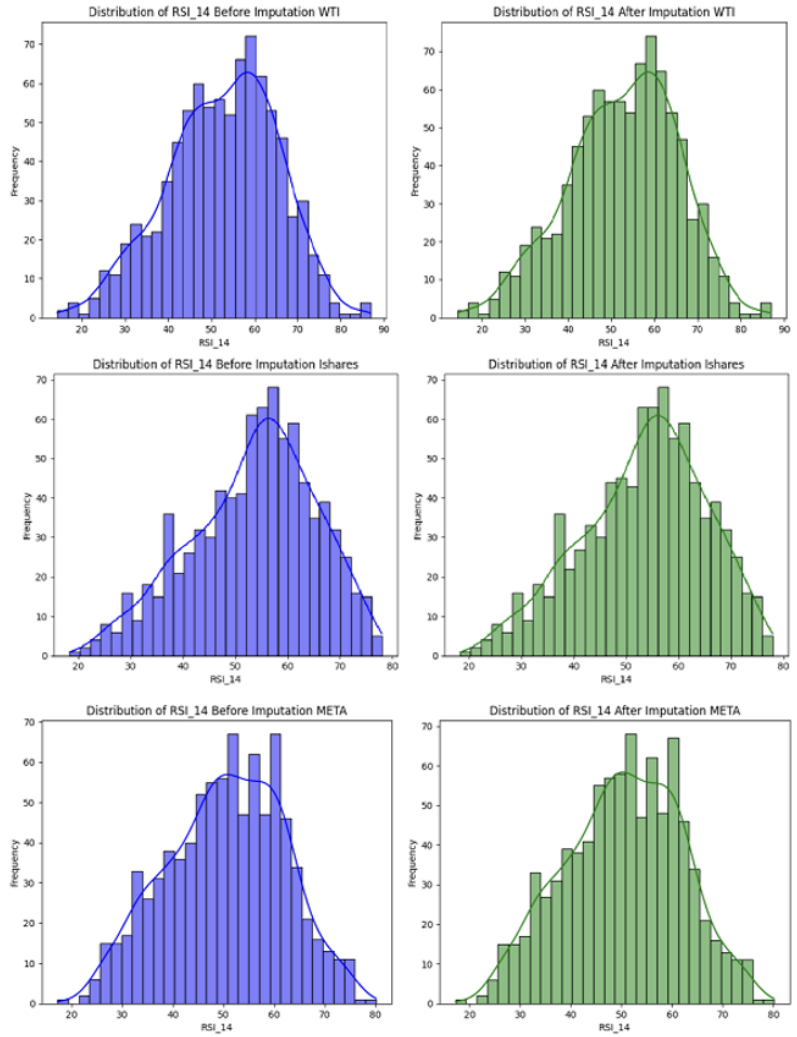
Dataset	X_train_shape	y_train_shape	X_test_shape	y_test_shape
meta	(881, 19)	(881,1)	(378, 19)	(378,1)
WTI	(882, 19)	(882,1)	(378, 19)	(378,1)
iShares	(878, 19)	(878,1)	(377, 19)	(377,1)

Variance calculation

Feature	META	ISHARES	WTI
RSI	154.6	138.7	145.6
MACD	59.0	0.4	4.1
MACD-Histogram	3.9	0.0	0.3
Signal	54.3	0.3	3.6
Moving Averages	6236.6	62.3	405.2
Stochastic Oscillator %K	639.2	776.6	592.4
Stochastic Oscillator %D	598.1	729.4	555.8
Exponential Moving Average	7374.3	70.8	415.9
Lower Band	6719.6	71.8	397.4
Middle Band	7130.3	68.9	415.0
Upper Band	7816.6	68.1	458.8
Bandwidth	104.0	16.2	638.2
Percent B	0.1	0.1	0.1
Senkou Span A	5962.9	62.4	469.4
Senkou Span B	5148.5	58.8	565.4
Tenkan-sen	7372.5	71.4	442.0
Kijun-sen	6897.7	68.5	491.9
Chikou Span	7616.3	70.9	429.1
Standard deviation	56.7	0.4	4.4

After calculating the variance, only 1 feature was removed for iShares, leaving it with 18 and meta and WTI with 19.

Missing Values and imputation

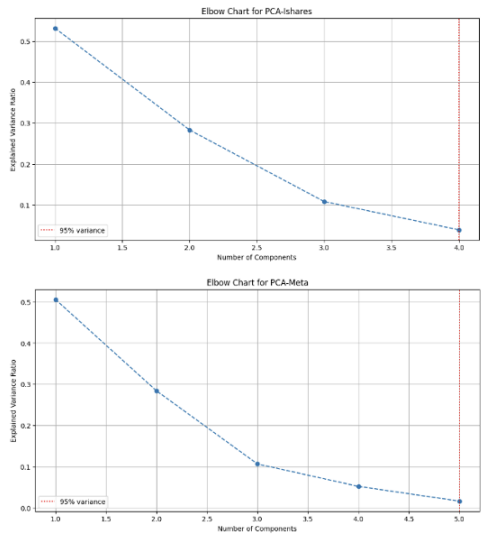


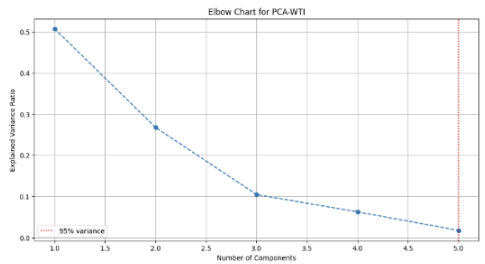
Feature Reduction

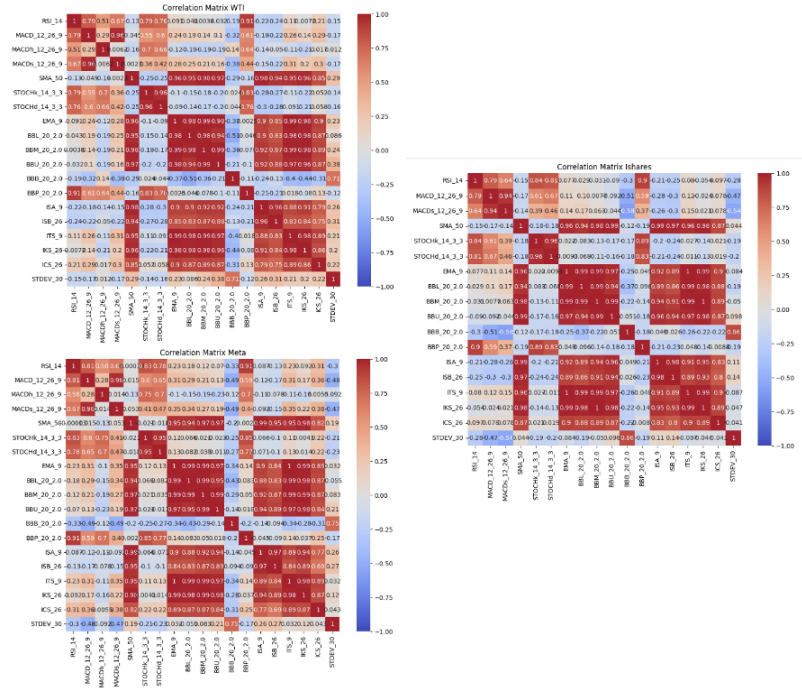
KMO Value

Range	Description
0.00 to 0.49	Unacceptable
0.50 to 0.59	Miserable
0.60 to 0.69	Mediocre
0.70 to 0.79	Middling
0.80 to 0.89	Meritorious
0.90 to 1.00	Marvelous

Dataset	KMO Model	Result
Meta	0.806	Meritorious
WTI	0.816	Meritorious
iShares	0.823	Meritorious

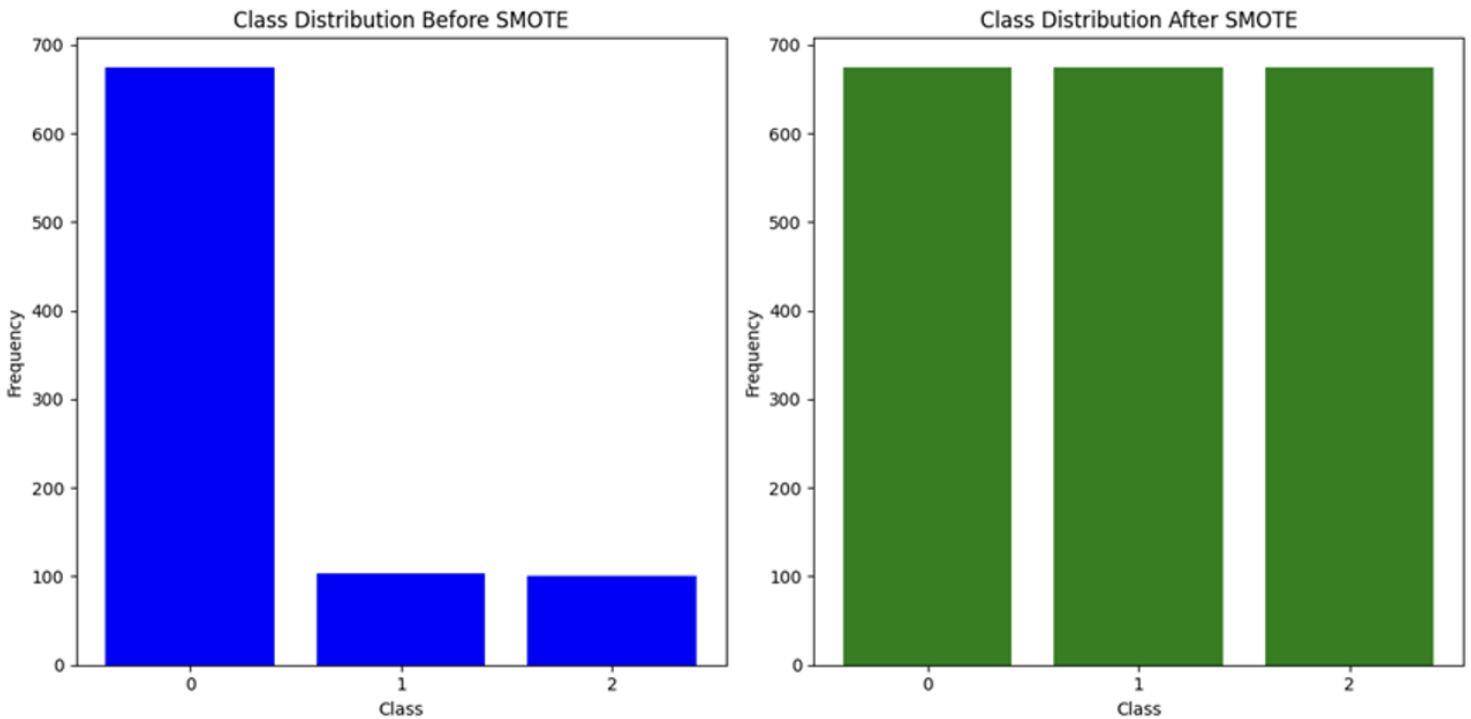




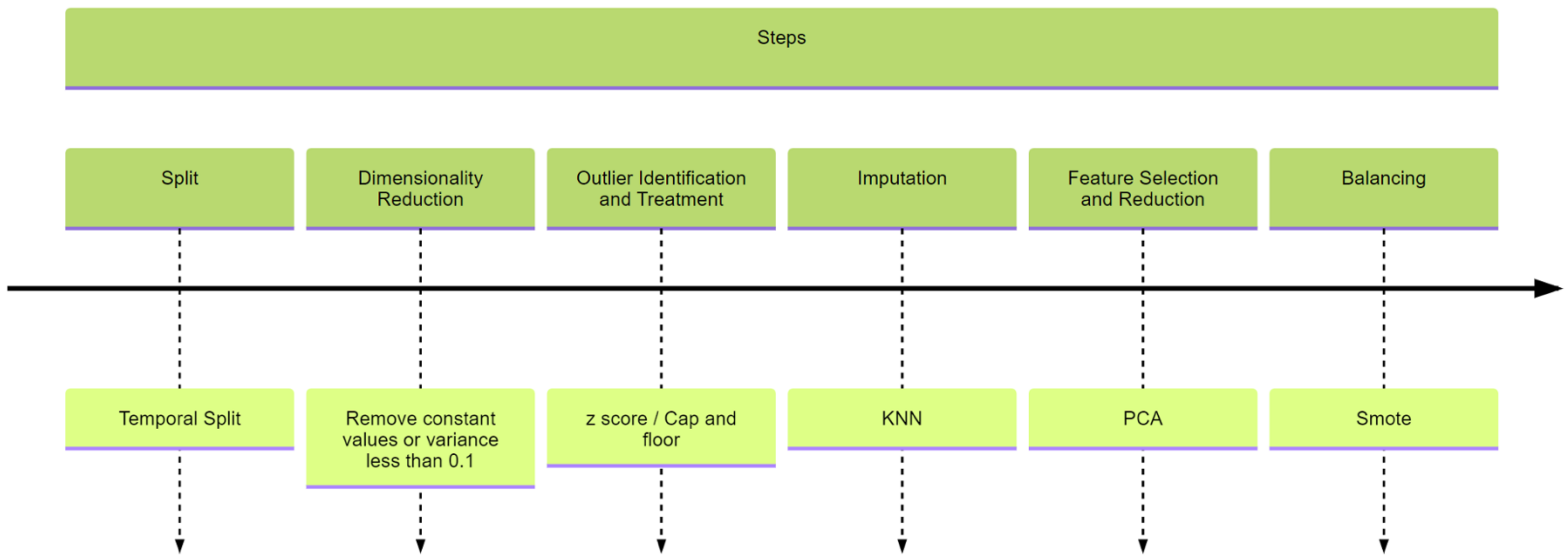


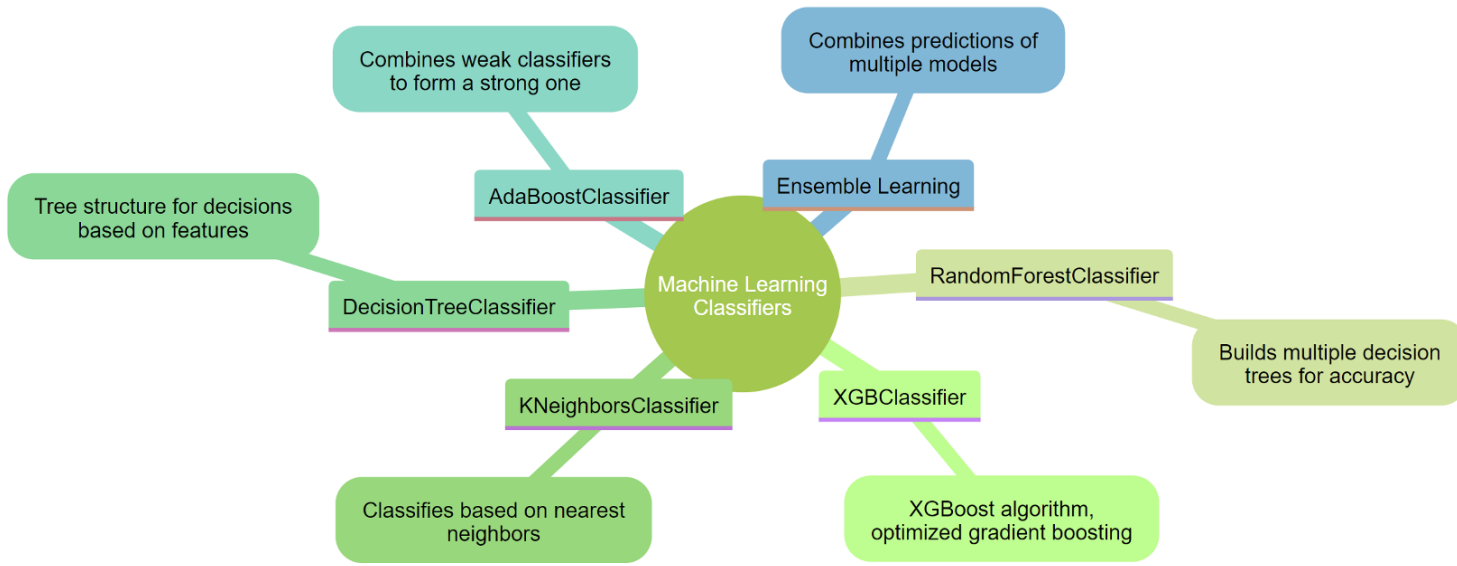
PCA effectively reduces data dimensionality for all the stocks, with KMO test indicating data suitability for PCA

Smote before and after for iShares

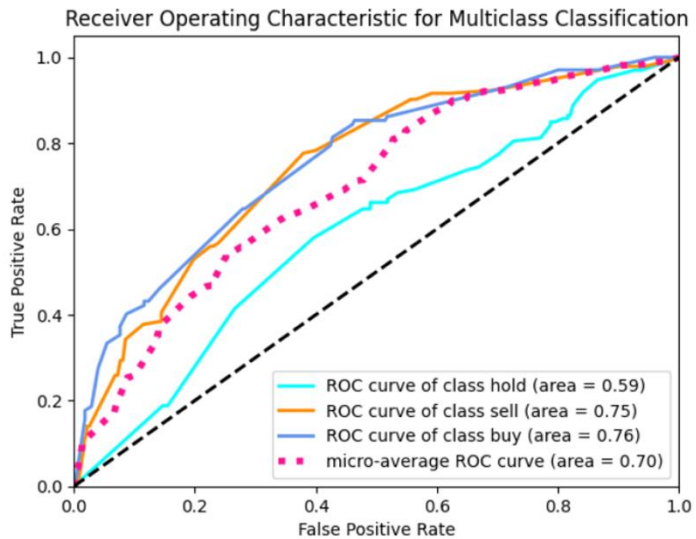


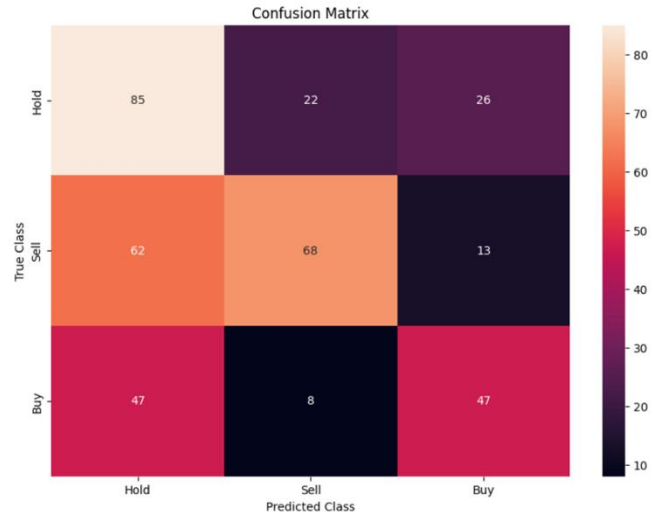
Summary steps




Machine Learning Classifiers
Ensemble Learning
Combines predictions of multiple models
AdaBoostClassifier
Combines weak classifiers to form a strong one
DecisionTreeClassifier
Tree structure for decisions based on features
RandomForestClassifier
Builds multiple decision trees for accuracy
KNeighborsClassifier
Classifies based on nearest neighbors
XGBClassifier
XGBoost algorithm, optimized gradient boosting

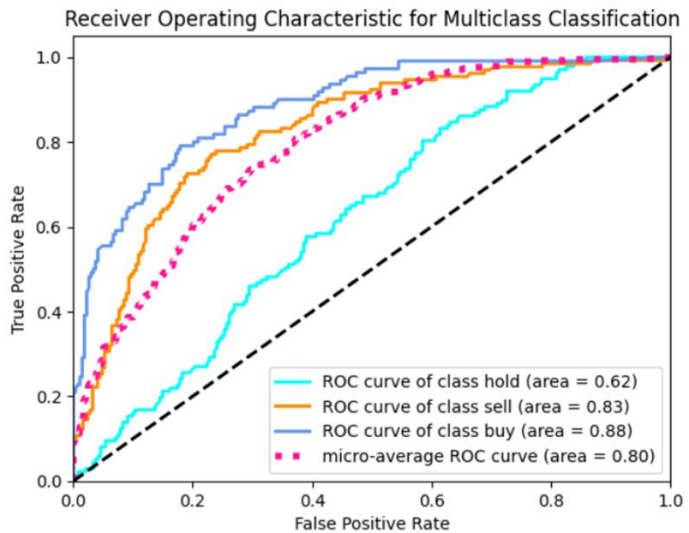
Meta Winning Model

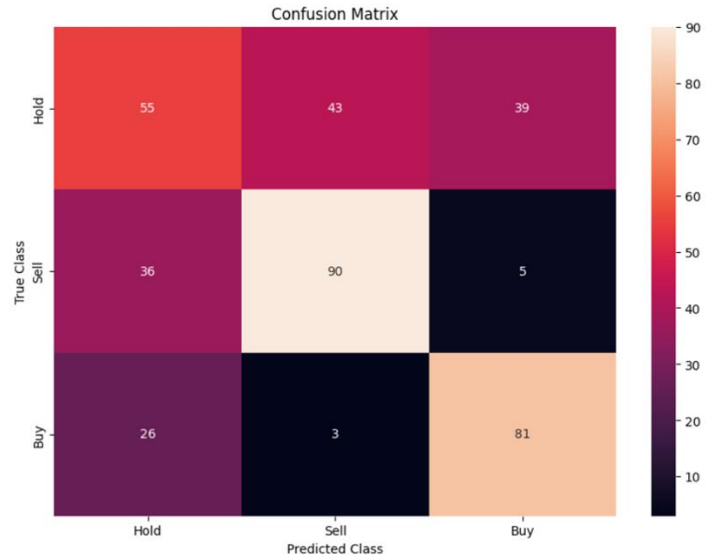




Decision trees emerged as the winning model for Meta, outperforming iShares in terms of accuracy, with higher rates of true positives and false predictions. it is possible to see that in all classes they are better than a random classification.

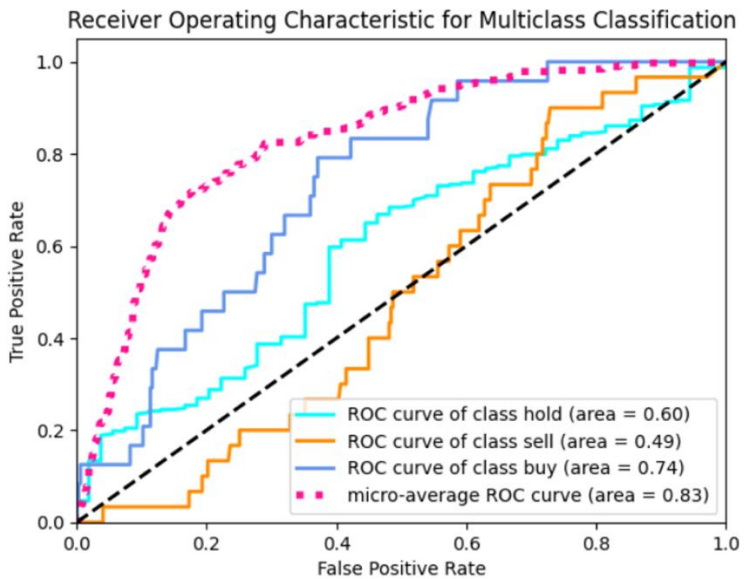
WTI Winning Model

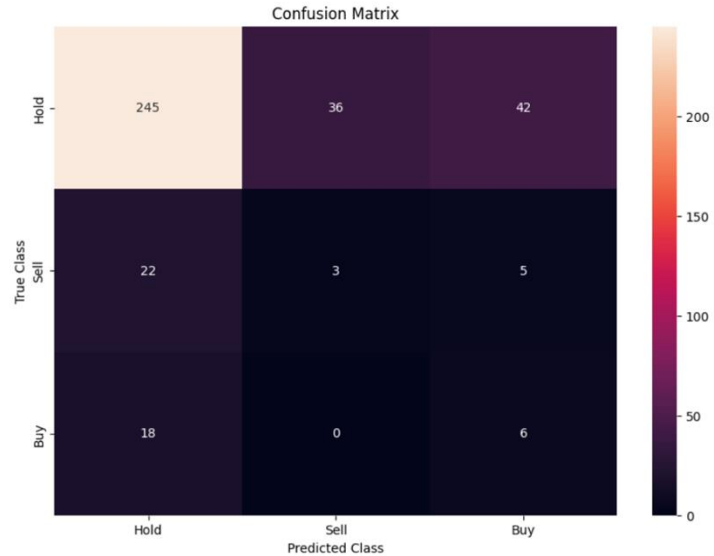




The predictive model for WTI(Random forest) is more accurate for buy and sell classes, but has a high false positive rate. It generally yields positive results, but struggles with hold indicator classification.

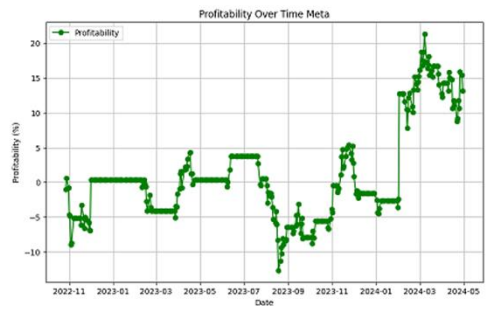
iShares Winning Model

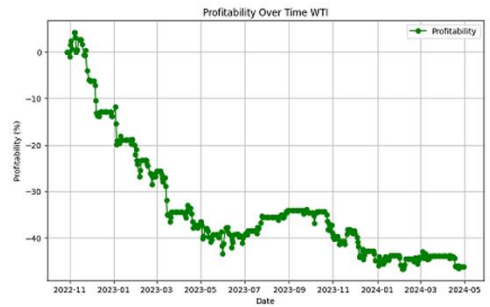




The random forest model for iShares effectively differentiates the Hold class but struggles with the sell and buy classes, as shown in the average Micro ROC curve

Backtesting results





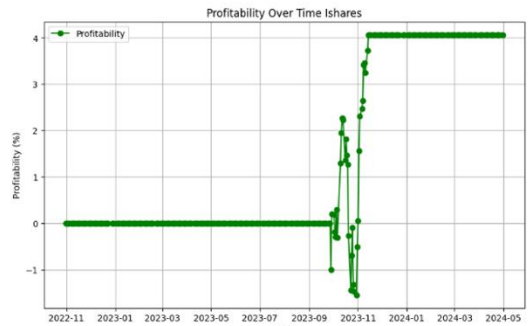
The machine learning model's profitability analysis showed a positive trend for iShares, with Meta offering the highest margin but high risk, while WTI experienced constant falls.

Backtesting results

Date	Open	Target	Predicted Signal	Signal	Portfolio Value	Action	Profitability %	Holdings	Cash
4/22/2024	489.72	2	2	buy	108.80	hold	8.80	0.22	0.00
4/23/2024	491.25	0	2	buy	109.14	hold	9.14	0.22	0.00
4/24/2024	508.06	1	1	sell	111.75	sell	11.75	0.00	111.75
4/25/2024	421.4	2	2	buy	110.63	buy	10.63	0.26	0.00

Target: the real market signal.
Predicted Signal: the market signal predicted by the model.
Signal: the actual name representing that class, i.e. 0=Hold, 1=sell and 2=buy.
Portfolio Value: the total value of the portfolio that day (cash + holdings * price).
Action: the action taken that day according to the available balance or available stocks.
Profitability: the change in the portfolio due to the action taken.
Holdings: the number of shares held that day.

Backtesting results – Base strategy



The base strategy, which uses RSI for buy and sale decisions, demonstrated the highest performance across all studied stocks, demonstrating its precision across various sectors.

Conclusion

- The Roc curve and confusion matrix show consistent models, but base strategy wins in all cases, suggesting ML models require improvement.
- This evaluation shows that in general the model can capture well the trends of certain sectors such as technology and the diversified market.
- The difference in performance (Roc and confusion matrix) for Meta and WTI models regarding iShares may be related to class balancing.
- iShares exhibits a conservative growth pattern due to its constant market growth, resulting in stable profits but not high compared to other stocks.
- Machine learning models perform better with stock with rapid changes and fluctuations, but not with stable models or constant upward trends.