ML Algorithmic Trading

FinTech Butlers

Michelle Silver
Malika Ajmera
Matt Glasgow
Joseph Knight
Mike Nguyen
Luis Torres

University of California Berkeley FinTech Bootcamp : Project 2

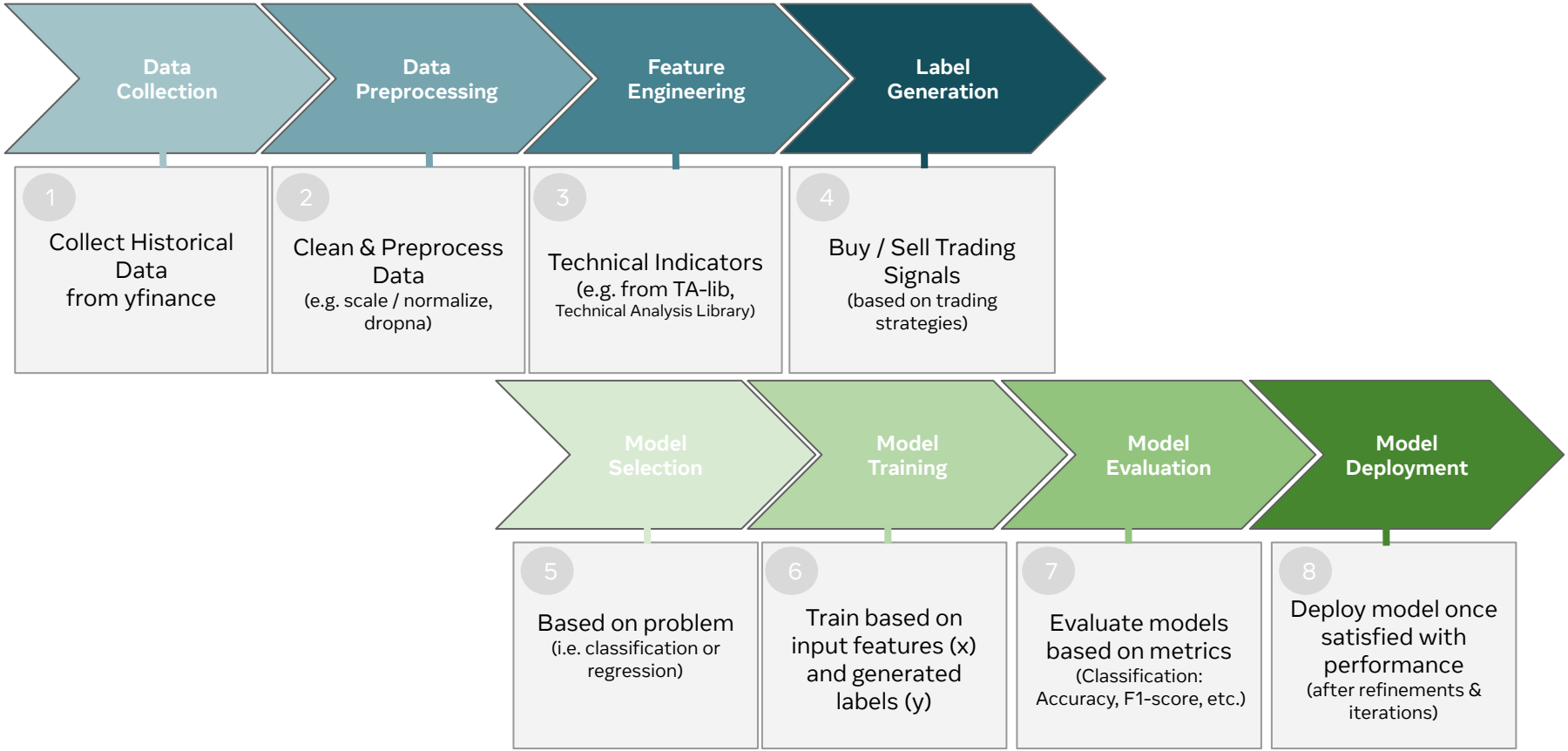
July 2023

ML Algorithmic Trading Project

Objective: Build predictive models using machine learning algorithms for buy-sell decision making

- **Why?**
  - Help develop profitable trading strategies based on machine learning model results
- **How?**
  - Leveraging historical stock data from yfinance, TA-lib technical analysis library, & various machine learning models

MACHINE LEARNING: APPROACH



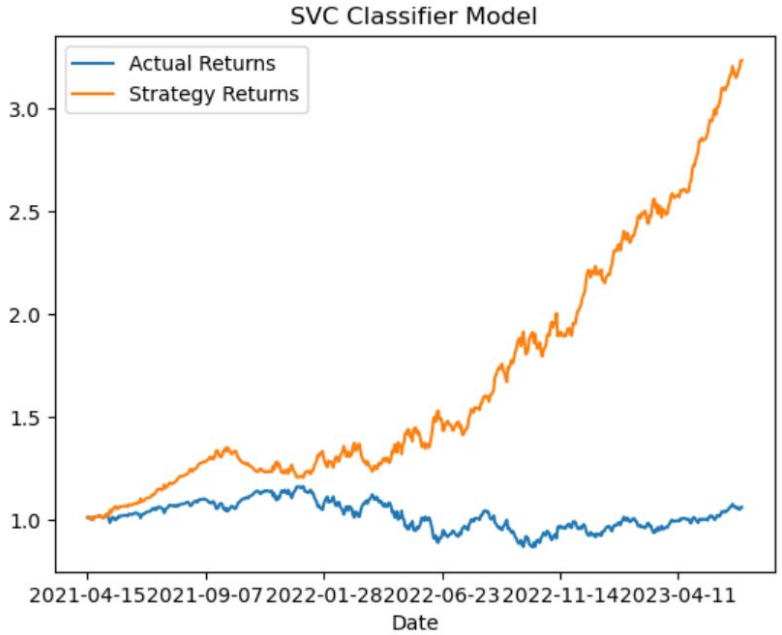
MACHINE LEARNING MODELS

Model	Description	Variables	Challenges	Results	Conclusion
SVC Classifier	Good for high dimensional spaces	Signal - 10d stock price change, SMAs, RSI, MACD, P/E	Deciding variables and indicator	Accuracy Score: .72	Promising Evaluation Metrics
Decision Tree Classifier	Model that builds flowchart-like tree structure	Signal - 10d stock price change, SMAs, RSI, MACD, P/E	Deciding variables and indicator	Accuracy Score: .85	Promising Evaluation Metrics
Random Forest Classifier	Predict Buy / Sell Trading Signals (Classification Model)	• Relative Strength Index (RSI) • Close Price	Multiple iterations required to improve performance (e.g. adjust data time frame, overbought / oversold thresholds)	Accuracy Score: .98	Promising Evaluation Metrics
AdaBoostClassifier (AbC-3Y-7030)	Fits a base or weak learner to the training data. Assigns equal weights to each training example iteratively training multiple weak learners	• 3Y SPY closing prices • Short window = 4 days • Long Window = 100 days • 70/30 training split	• Training/Testing split using trading days	Accuracy: • Training : 70% • Test: 50%	• Useful when features don't provide enough info to the model • Increase training split (e.g. 80/20) to improve performance

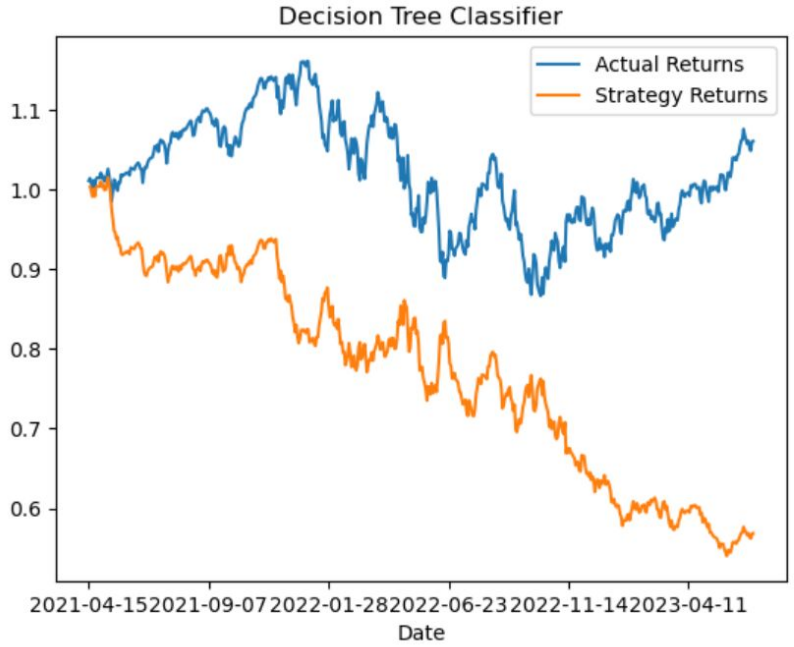
MACHINE LEARNING MODELS

Model	Description	Variables	Challenges	Results	Conclusion
Long Short-Term Memory (LTSM) Neural Network	Predict Price of Stock & Returns (Regression Model)	• SPY data (3Y) • Closing price • 80/20 Split	Understanding the LSTM model itself and what parameters to choose	Train Loss: 0.0014 Test Loss: 0.00061 MSE - loss: 0.0019	Works best for time series analysis
Logistic Regression (LogReg-3Y7030)	Models the relationship between the features and the probability of the target variable belonging to a specific class. It is a classification, not a regression algorithm	• 3Y SPY closing prices • Short window = 4 days • Long Window = 100 days • 70/30 training split	• Training/Testing split using trading days	• Training : 56% • Test: 24% • Poor recall on sell signal	• Do not use • Try a different Model (see AdaBoostClassifier)
SVC Classifier	Predict Buy / Sell Trading Signals based (Classification Model)	• Force Index • Close.diff * • Volume	The sudden spikes in the direction of the price moment can help confirm the breakout. Inverse the signal	Accuracy Score: 0.49	Confirmation indicator

RESULTS & CONCLUSION



	precision	recall	f1-score	support
-1	0.60	0.96	0.74	231
1	0.95	0.55	0.70	324



	precision	recall	f1-score	support
-1	0.82	0.82	0.82	231
1	0.87	0.87	0.87	324

NEXT STEPS

Questions

- Are there other pairs of indicators that will produce better results?
- What accuracy score is sufficient?
- What model loss is acceptable?

Future Research

- Explore other machine learning models
- Runtime/ Resource efficiency
- Test on other asset classes, stocks and sectors
- Test on ‘paper trading’ mode

Future Releases

- Add UI/UX
- API integration
- Deploy to the cloud
- Blockchain record keeping

LINKS

- **GitHub repo: *P2_ML_AlgoTrade***
  - Presented Models in <main branch>
  - All other models in dev branches (5)