PROJECT 2

UNC AI Bootcamp

Predicting Covered Call Options
on SPY using Machine Learning

GROUP 3:

Yeongjin Nam
Mat Nicholas
Nee Buntoum
Jana Avery
Jay Requarth
Dr. Dick Davis
John Andrews

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

The Problem Statement

Investors like retirees face optimizing covered call strategies to protect their investments and create ongoing revenue.

Some investors use covered calls to generate income from their portfolios while preserving capital and limiting downside risk.

The goal of this project is to use machine learning to answer the question on whether or not this is a viable strategy and if so, help investors maximize their returns.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Let’s use a lemonade stand analogy to understand covered calls.

Lemonade Stand Story:

Imagine you own a small lemonade stand that you believe is worth $100. One sunny day, a neighbor comes to you and says,

“I see you’re doing well with your lemonade stand. I’d like to have the option to buy it from you for $100 next month because I think it will be worth more by then. I’ll pay you $10 right now for that option.”

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

You get to keep the $10 no matter what…

If your neighbor decides to buy the lemonade stand for $100, you’ve made a bit of profit (the $10 option fee).

If your neighbor decides not to buy the stand because it's not worth more than $100, you still keep your stand and the $10.

Why Covered Calls?

Covered calls are a smart way for stock owners to make extra money without having to sell their stocks. Imagine you own a piece of a company and think its value won't change much soon.

This strategy is perfect for people who want to keep their stocks for a long time and are okay with making a bit of money even if the stock doesn't go up much. It's a cautious way to earn from your investments without expecting big surprises in the stock's price.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

This situation is similar to selling a covered call in the stock market.

The lemonade stand is like your stock.

The $10 payment is the option premium that an option buyer pays you, the stock owner, for the right (but not the obligation) to buy your stock at a predetermined price (the strike price, in our analogy, $100) within a certain period.

AI to the Rescue

Our AI project tackles the complexities of the covered call investment strategy, offering a smarter, AI-driven solution for investors, especially retirees, looking to generate income without selling their stocks. **Here's how:**

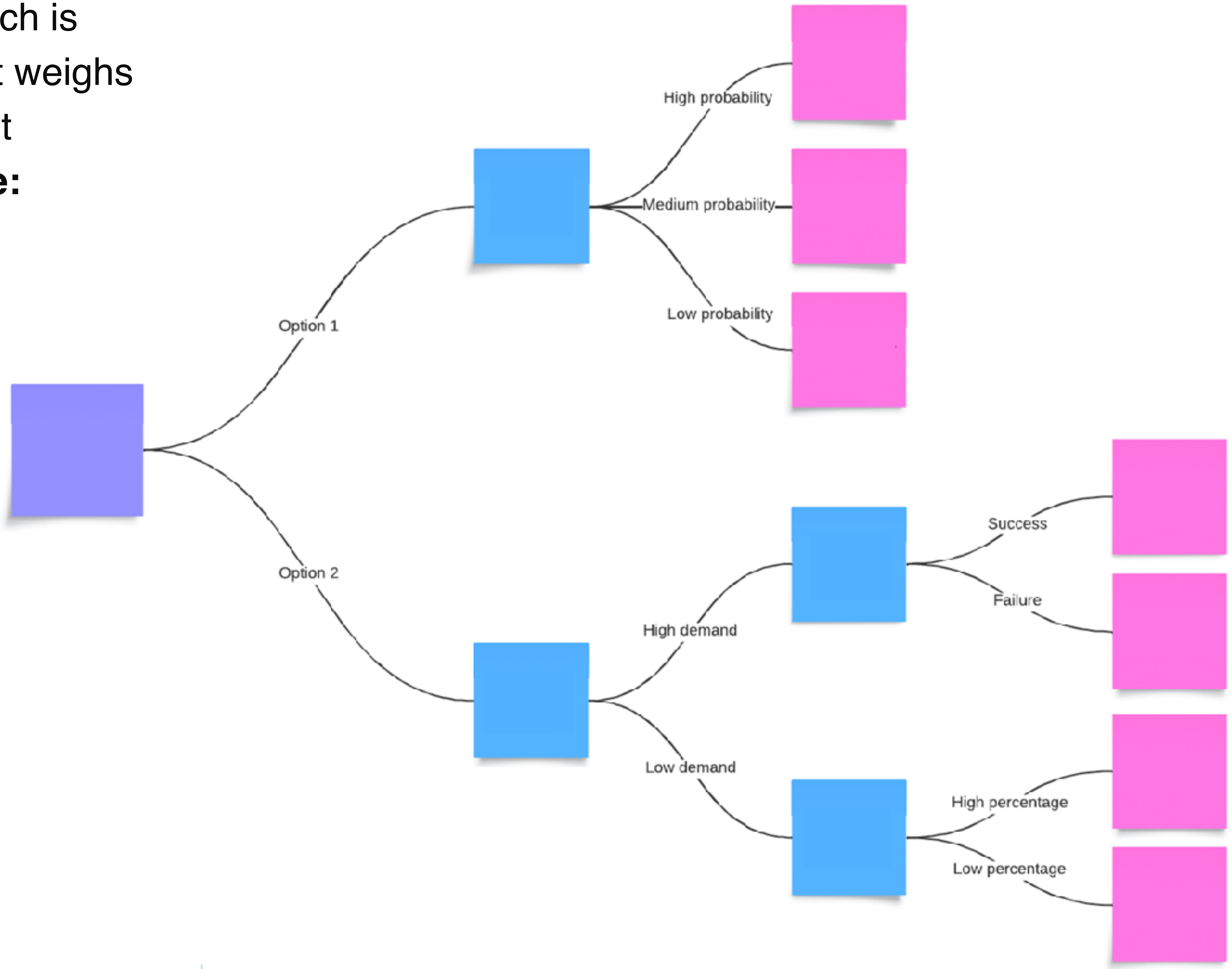
- **AI simplifies choices** by quickly analyzing data to identify the best stock options and call strategies, removing guesswork.

- **It enhances risk management** by evaluating potential risks and returns from historical trends, helping investors make informed decisions.

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

Choosing the right AI model is akin to selecting the best path in a complex maze. Our approach is methodical, guided by a decision tree that weighs various probabilities and demands against expected outcomes. **Here's our rationale:**

- **Assessing Probabilities**
- **Demand Analysis**
- **Success Metrics**
- **Quantitative Rigor**
- **Iterative Refinement**
- **Outcome-Focused**



OUR SOLUTION:

AI-Powered Covered Calls Service

We propose to develop an AI-driven service that simplifies and optimizes the covered call strategy for individual investors, mainly targeting retirees seeking to enhance their income. Our AI system will analyze vast amounts of stock and options data in real time, recommending the most promising covered call opportunities based on the individual's portfolio, risk tolerance, and income goals.

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Optimizing Covered Call Strategies with AI

Data Integration

Beginning our journey with robust data integration, we harness real-time market insights from reputable sources. Our dedicated API key from Polygon.io serves as a gateway to the latest financial data, laying the groundwork for informed AI analysis.

Key Insights:

- At-the-Money CC strategy mitigates downturns but may underperform in a strong bull market.
- Adjusting strike prices could significantly influence these results.

Table for Project-2:

Overall: April 6, 2022, to March 20, 2024.
437 rows x 36 columns = 15,732 data points.
Bear market: April 6, 2022, to Oct 7, 2022
Bull market: Oct 8, 2022, to March 20, 2024

	SPY open	SPY close	SPY gain	SPY CC
Overall	$446.12	$519.17	+16.4%	+24.1%
Bear market	$446.12	$361.99	-18.9%	+5.7%
Bull market	$361.99	$519.17	+41.1%	+26.4%

CONCLUSIONS: The At-the Money CC strategy decreases loss during a bear market; however, it performs poorly during a strong bull market.

Moving the strike price out a few dollars will change these findings considerably.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Scraping Data from polygon.io

```
def cc_close_spy(row):
  #
  string = str(row['Timestamp']) # Grab Date "2024-01-01"
  odate = string[2:4] + "" + string[5:7] + "" + string[8:] # 6 digit code for date (240401)
  strike = str(row['SPY Strike']) # Strike Price
  if len(strike) == 2:
    strike = '0' + strike
  ocode = f'O:SPY{odate}C00{strike}000'
  # ---- #O:SPY240401C00511000
  query_url = f'{base_url}/v2/aggs/ticker/{ocode}/range/5/minute/{string}/{string}?adjusted=true&sort=asc&limit=50000&apiKey={api_key}'
  response = requests.get(query_url) # Using Response API
  json_data = response.json()
  try: #Error - Data doesn't exist
    if json_data['resultsCount'] > 0:
        # Convert JSON to Pandas Dataframe
        tdf = pd.json_normalize(json_data['results'])
        tdf['t'] = pd.to_datetime(tdf['t'], unit='ms')
        cprice = tdf.loc[(tdf['t'].dt.hour == 19) & (tdf['t'].dt.minute == 45)]['h'].mean()
        return cprice
    else:
      return 0
  except:
    return -1

testtest = df.copy()
testtest['SPY Open CC'] = testtest.apply(cc_open_spy, axis = 1)
testtest['SPY Close CC'] = testtest.apply(cc_close_spy, axis = 1)
testtest.tail(30)
```

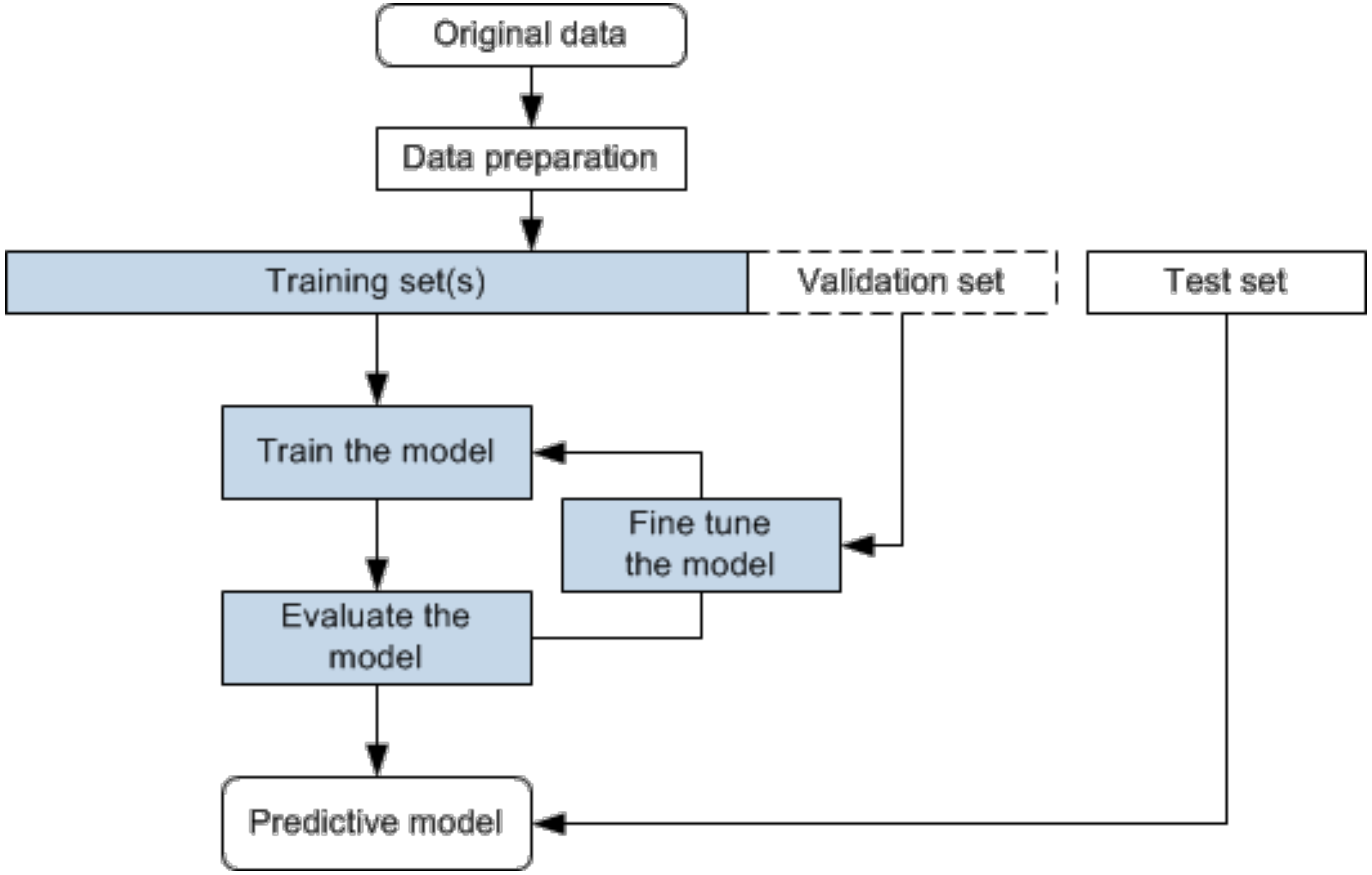
	Timestamp	SPY Opening Price	QQQ Opening Price	VXX Opening Price	DIA Opening Price	SPY Closing Price	QQQ Closing Price	VXX Closing Price	DIA Closing Price	SPY 5D Avg Volume	...	QQQ 1D Stock Change	VXX 5D Stock Change	VXX 3D Stock Change	VXX 1D Stock Change	DIA 5D Stock Change	DIA 3D Stock Change	DIA 1D Stock Change	SPY Strike	SPY Open CC	SPY Close CC
466	2024-01-18	474.47	410.790	15.4000	372.320	474.9400	412.1100	15.3850	373.700	72398793.50	...	0	0.09	0.00	0	-0.01	0.00	0	475	0.48	0.81
467	2024-01-19	479.02	416.300	15.0700	376.730	481.6200	420.2300	15.0400	378.480	75885352.25	...	0	0.06	0.00	0	-0.01	-0.00	0	480	0.20	2.22
468	2024-01-25	486.79	427.420	14.4000	378.600	485.8000	424.5785	14.5600	378.550	81971519.60	...	0	-0.07	-0.03	0	0.02	0.00	0	487	0.87	0.15
469	2024-01-26	486.85	424.000	14.6400	379.490	487.7050	424.3304	14.3900	380.910	78101267.80	...	0	-0.03	0.02	0	0.00	0.00	0	487	1.43	1.07
470	2024-01-31	488.38	420.890	14.3200	384.930	487.7400	421.5100	14.5550	384.800	70048347.40	...	0	0.00	-0.02	0	0.01	0.01	0	489	1.01	1.02
471	2024-02-01	484.79	419.370	14.8300	381.690	487.8200	421.1600	14.7800	384.190	78905551.80	...	0	0.01	-0.00	0	0.02	0.00	0	485	1.49	3.60
472	2024-02-02	489.83	423.900	14.8000	384.010	495.3700	429.7900	15.0400	387.380	82782881.40	...	0	0.01	0.00	0	0.01	0.00	0	490	1.12	5.70

Finding the correct data for our model proved challenging, given the limits of our access to our dataset.

Data Preparation

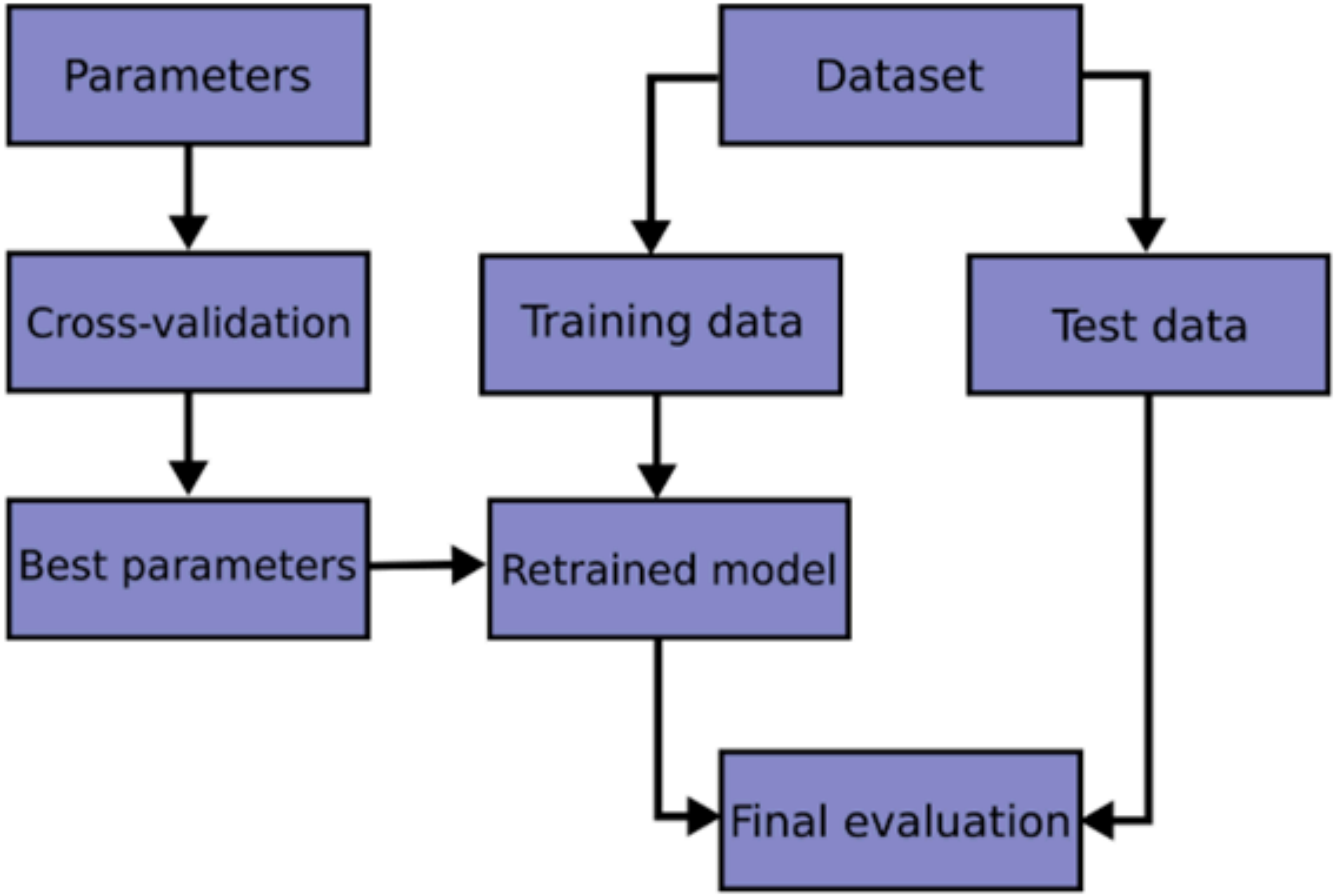
Feature Engineering

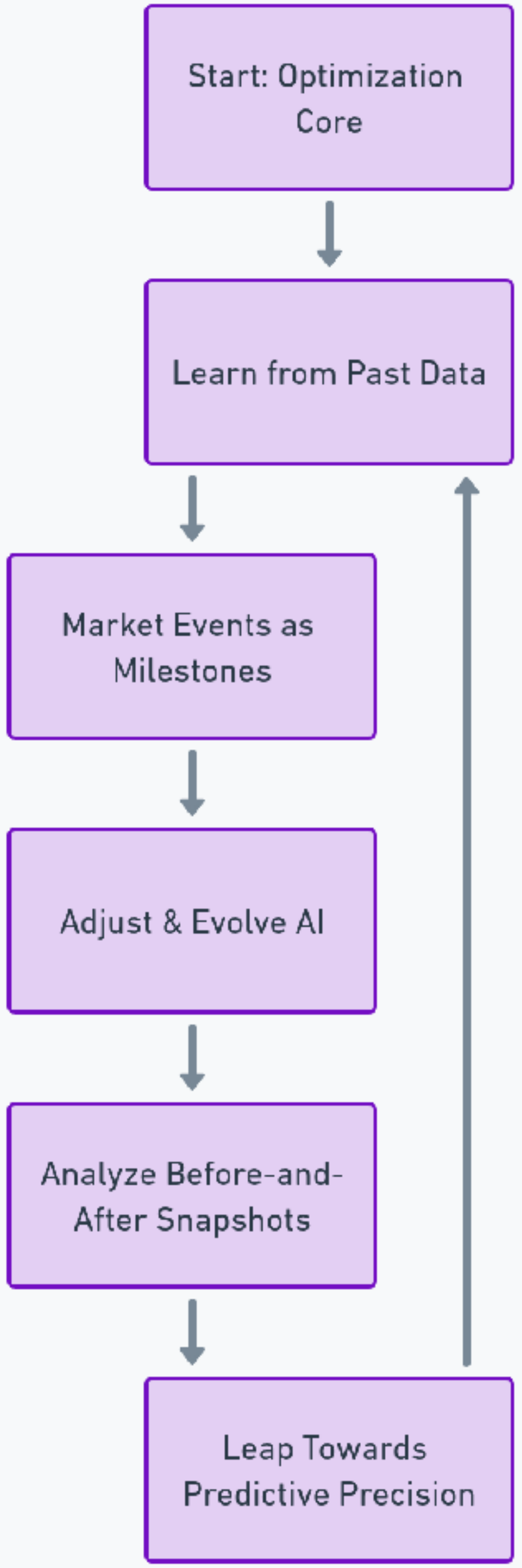
- Developed a target variable, 'trade,' to identify optimal trading days based on market movement predictions. This was complemented by rolling window calculations for short-term market trend analysis.



Validation Approach

Selectivity is key in feature selection. We sieve through our data, weighing the relevance of each factor against the scale of covered call strategy success. It's a blend of correlation and machine intelligence that determines the features with foresight.





Optimization Process

Considering the iterative nature of AI learning and the significance of market events in guiding AI evolution, a flowchart would be the most suitable visualization. This flowchart will visualize the continuous cycle of learning from past data, using market events as milestones for adjustment, and aiming towards predictive precision with the aid of before-and-after snapshots.

Data Methodology Overview

Sourcing Excellence

Outline the high-quality sources of your data, emphasizing the reliability and real-time nature of your data streams. Mention the use of Polygon.io API for the latest market data.

Data Transformation

Briefly describe the steps taken to clean and prepare the data for analysis, such as handling missing values, adjusting closing prices, and feature engineering.

Analytical Rigor

Summarize the analytical methods used to process the data. Highlight the use of time-series analysis to avoid look-ahead bias and the application of machine learning algorithms to identify trading signals.

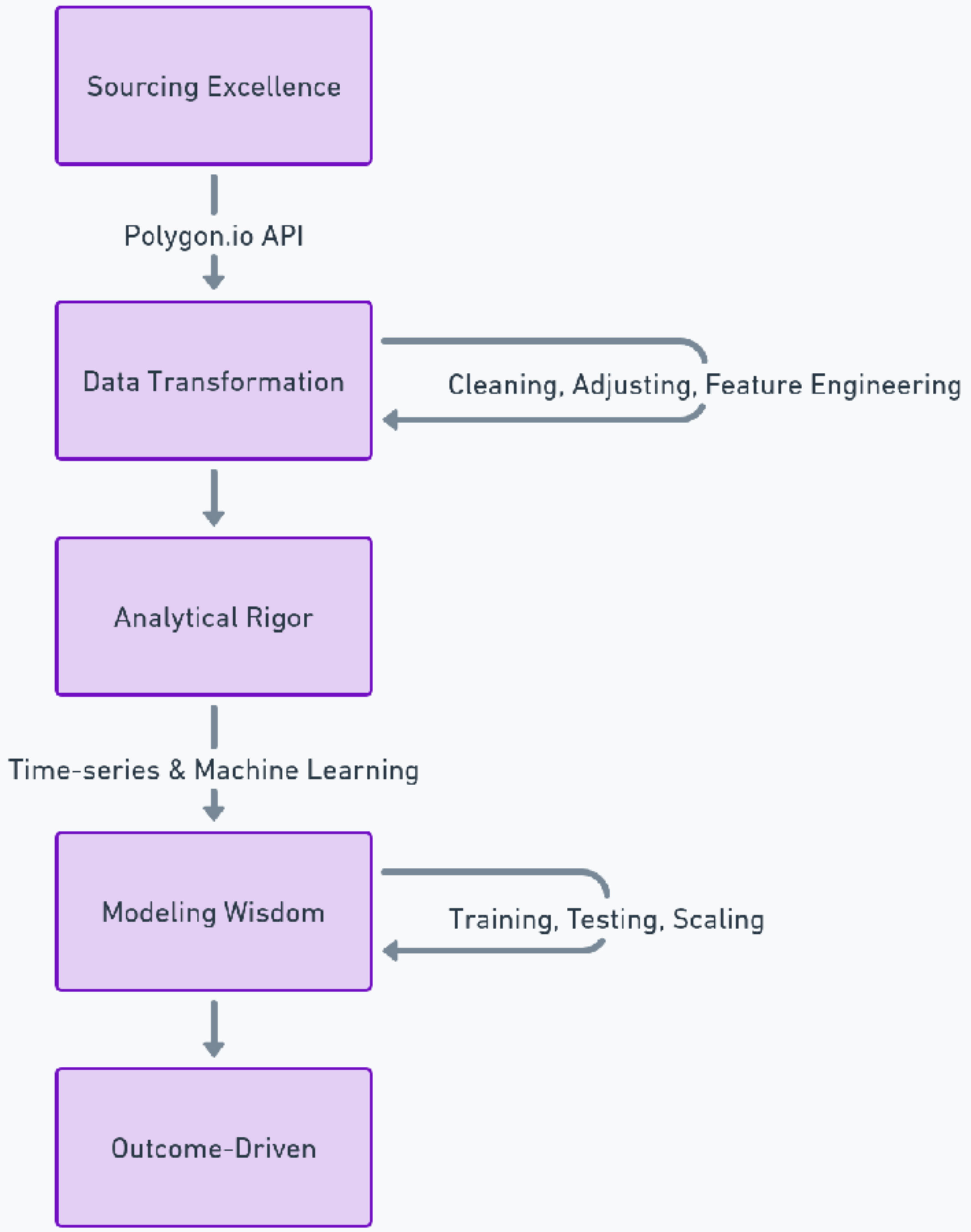
Modeling Wisdom

Explain how various models were trained, tested, and selected, and touch upon the importance of scaling data to ensure model accuracy.

Outcome-Driven

Conclude with how the data methodology leads to actionable insights and decisions, underpinning the model's predictions and the overall project's goal to assist in covered call trading decisions.

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL



Presentation Summary

In our quest to optimize trade recommendations for covered calls, we've harnessed two years of data, funneling it through a series of machine learning classifiers. Here's how they stacked up:

Results – Project-2

2 years of data:
432 rows, 25 columns

Target = 'trade' - was a trade recommended

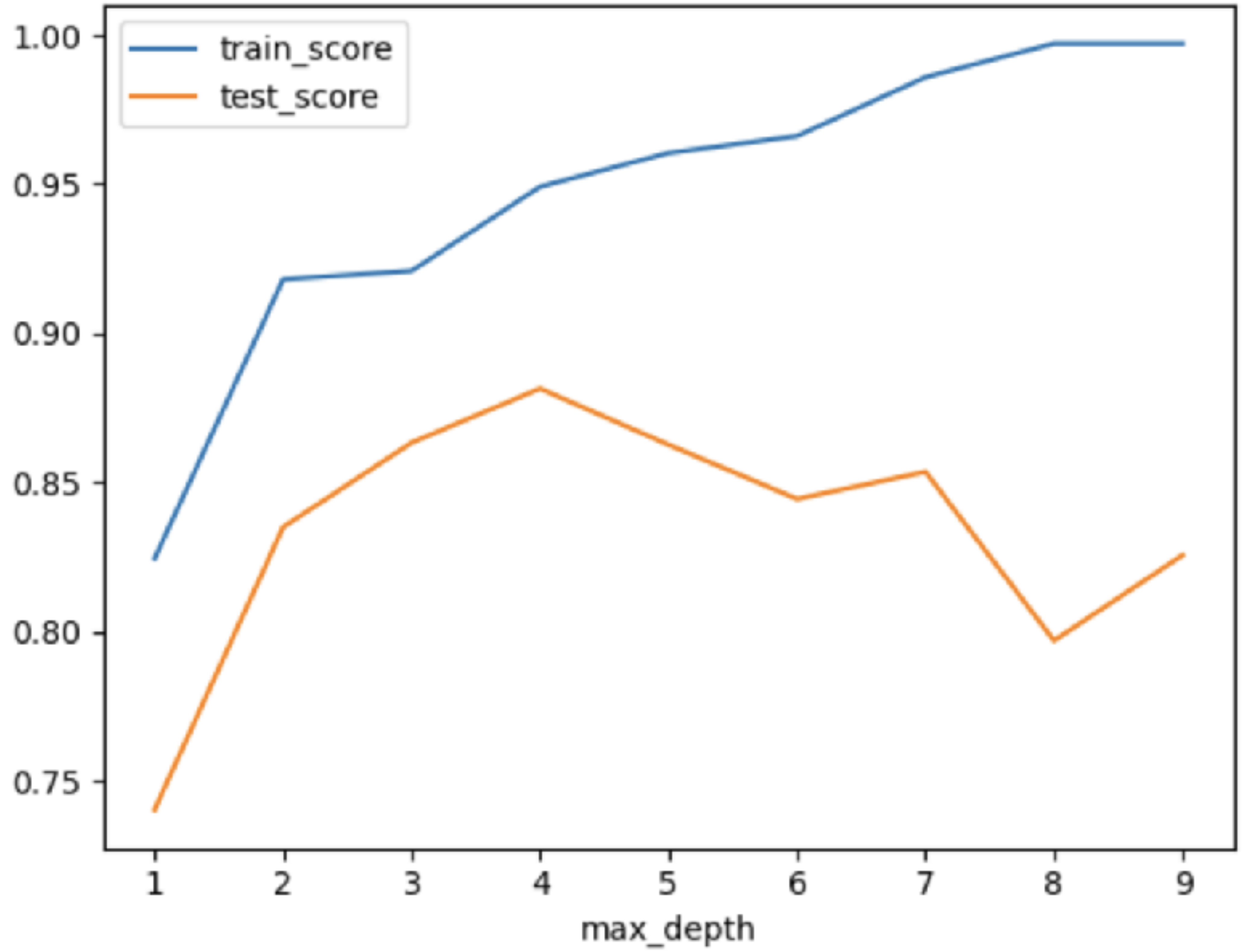
trade value counts: 0 (no) = 232
1 (yes) = 200

Machine Classifier	Training score	Testing score
Support Vector Machine – kernel = poly	0.806	0.666
Support Vector Machine – kernel = liner	0.852	0.824
Logistic Regression	0.870	0.824
Decision Tree	0.546	0.509
Random Forest	1.0	0.870
Random Forest (max depth =4*)	0.941	0.854
KNN (n_neighbors = 13*)	0.691	0.620
Gradient Boosting Classifier	0.9969	0.833
Ada Boost Classifier	0.994	0.843

Presentation Summary

Parameter optimization for the Random Forest classifier model.

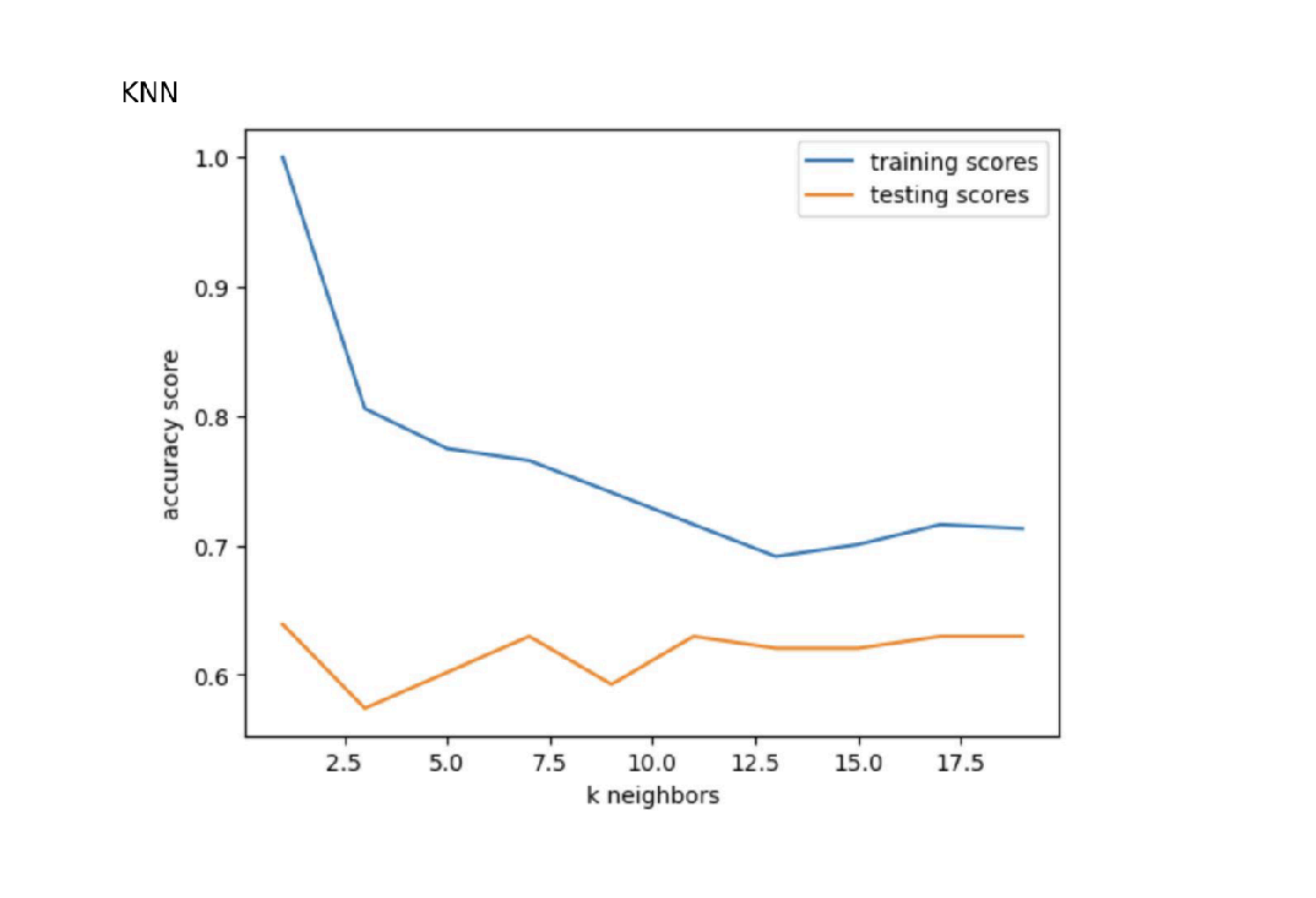
<Axes: xlabel='max_depth'>




```
models_df
```

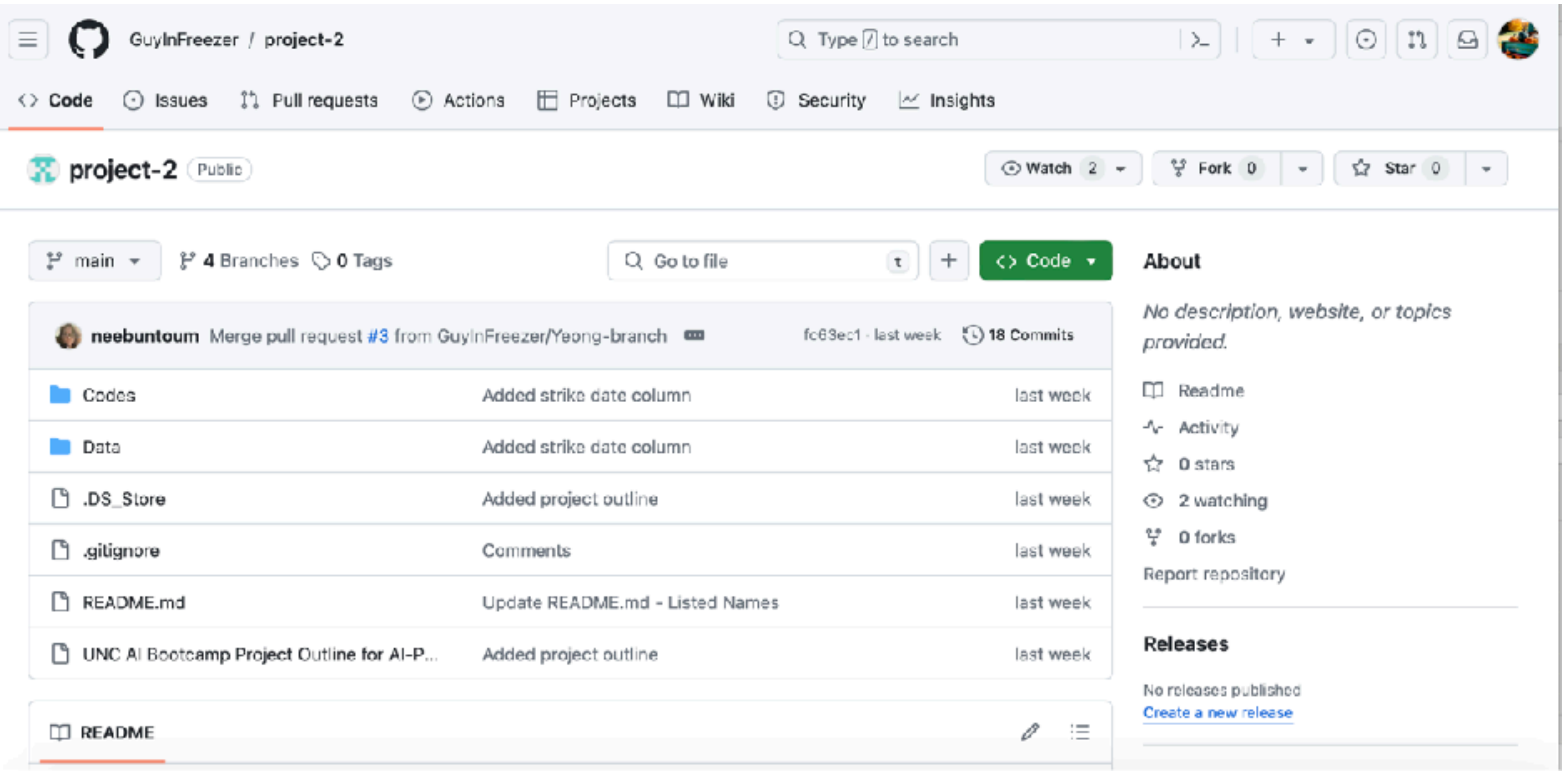
	train_score	test_score
max_depth		
1	0.824513	0.740309
2	0.918079	0.834991
3	0.920904	0.863293
4	0.949153	0.881475
5	0.960452	0.862607
6	0.966102	0.844425
7	0.985876	0.853516
8	0.997175	0.796913
9	0.997175	0.825557

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL



Presentation Summary

Parameter optimization for the KNN classifier model.



GitHub Repository Management

Our GitHub repository is a testament to our organized collaboration. It's the digital blueprint of our project, where code meets clarity, and documentation weaves the story of our development journey.

https://github.com/GuyInFreezer/project-2

Results & Insights

Using our Random Forest Classifier Model:

	Days	SPY open	SPY ROI*	SPY + CC ROI*
Overall	432	446.12	9.6%	14.2%
Bear	122	446.12	-37.8%	11.4%
Bull	310	361.99	33.1%	21.3%
RFC model	107	446.12**	19.8%	GUESS!!

* Annualized
** The SPY should be a weighted average between the bull and bear market, but we chose a higher number to give a worst-case scenario.

Results & Insights

	Days	SPY open	SPY ROI*	SPY + CC ROI*
Overall	432	446.12	9.6%	14.2%
Bear	122	446.12	-37.8%	11.4%
Bull	310	361.99	33.1%	21.3%
RFC model	107	446.12**	19.8%	47.4%

Presentation Summary

Today, we unveiled an AI-driven approach to optimizing covered calls, providing investors with a strategic edge. Here's the essence of our journey:

- **Data-Driven:** Our rigorous data collection from Polygon.io set the stage for accurate analysis.
- **Day 0 Focus:** Concentrating on 'Day 0' SPY covered calls, we've crafted a daily strategy for consistent income.
- **Methodical Analysis:** We processed and engineered features from over two years of market data, ensuring a robust model foundation.
- **Precision Modeling:** By optimizing models like Random Forest and KNN, we achieved a nuanced understanding of when to execute trades.
- **Proven Results:** The models demonstrate promising strategies, reducing losses in bear markets and capturing gains in stable conditions.
- **Next Steps:** Our path forward includes refining models and exploring dynamic strike price adjustments.

In conclusion, our AI model stands as a beacon, guiding retirees through the complexities of the stock market towards a horizon of steady income.

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL