



Fintech Bootcamp

Insider Trades

Team 2

Content

- 01** Our Team
- 02** Problem Statement
- 03** Solution
- 04** Data Import and Cleansing
- 05** Machine Learning Models and Results
- 06** Results and recommendation
- 07** Lessons Learned and Next Steps

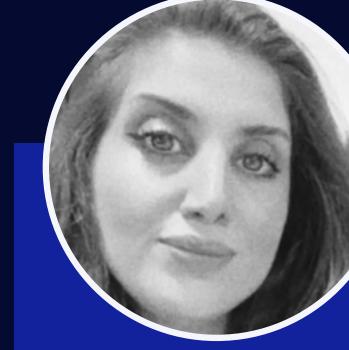
Our Team



Mario
Gunawardena



Joseph
Scorsone



Alghalia
Alsammak



Luis Rosales
Garcia



Nafeezur
Chowdhury

Problem

- Corporate Insider trading is legal if the insider makes a trade and reports it to the Securities and Exchange Commission
- Figuring out whether an insider trade filed with the SEC (Form 4) has an impact on stock price discovery.



Solution

- The Insider application checks the effect of insider trades on the top 10 AI companies closing price after 1 week of the filing date and compares the trend by using machine learning models.
- It downloads 3 years of market daily closing from Yahoo Finance and the insider trade with SEC (Form 4) for the top 10 AI companies

DATA FORMATTING

CREATING AND TUNING
THE MODELS

TESTING & MODELS
COMPARISON

FINAL REPORT



Data Inputs / Cleaning

Sources

- FinViz
- Yahoo Finance
- Sec.gov

Data Cleansing

- Remove \$, ", ", % ,#errors
- Changing the data types to datetime, float, integer
- Drop unnecessary columns and null values
- Applying Standard and other encoders

Aggregation

- Add stock closing data
- Add offsets to stock data
- Filter data based on the stock
- Adding labels to the data

Tools & Technologies

Multiple technologies and statistical models are used to build the The insider Application



Python



Pandas, Numpy



FinViz



yfinance



Google Collab



Streamlit



Sklearn



Matplotlib



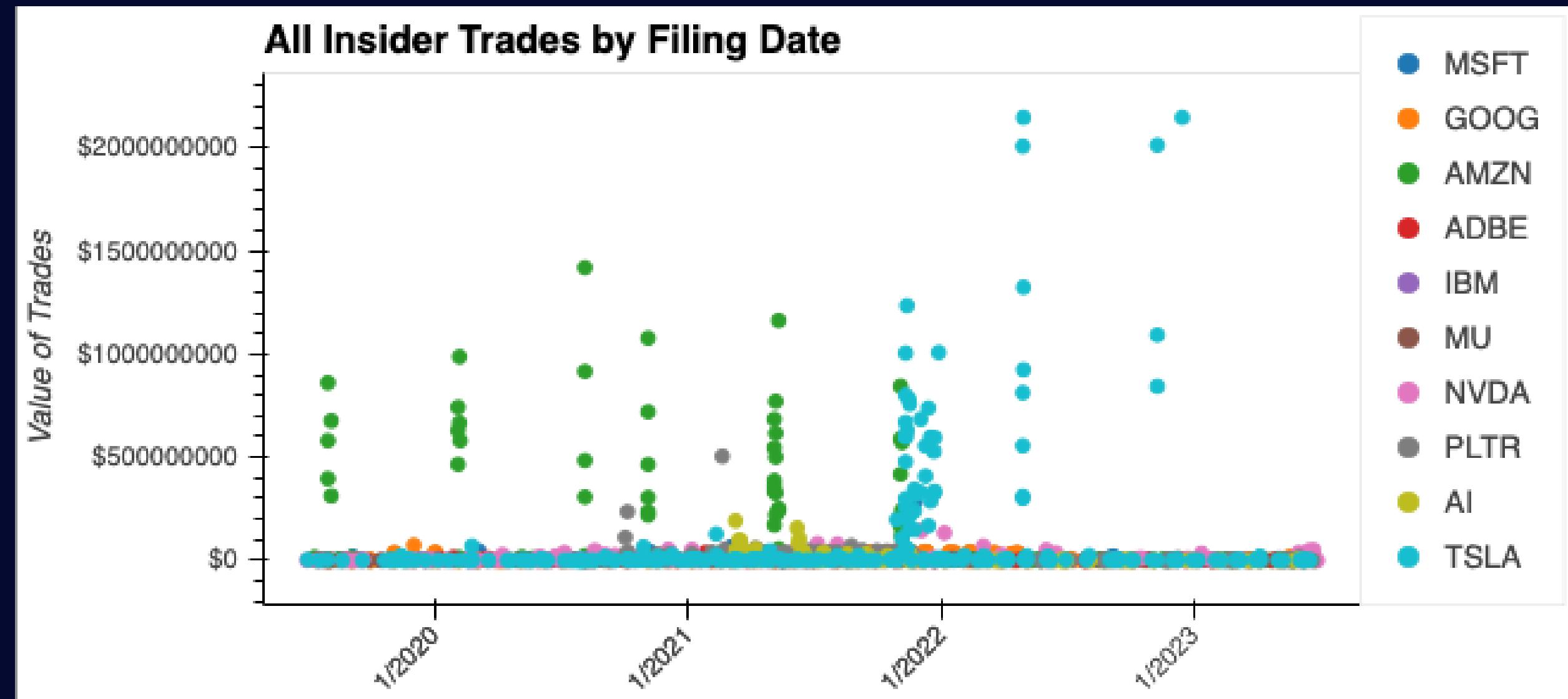
Github

Insider Trading Data - Summary

Insider Trades for AI
related companies for the
last 3 years

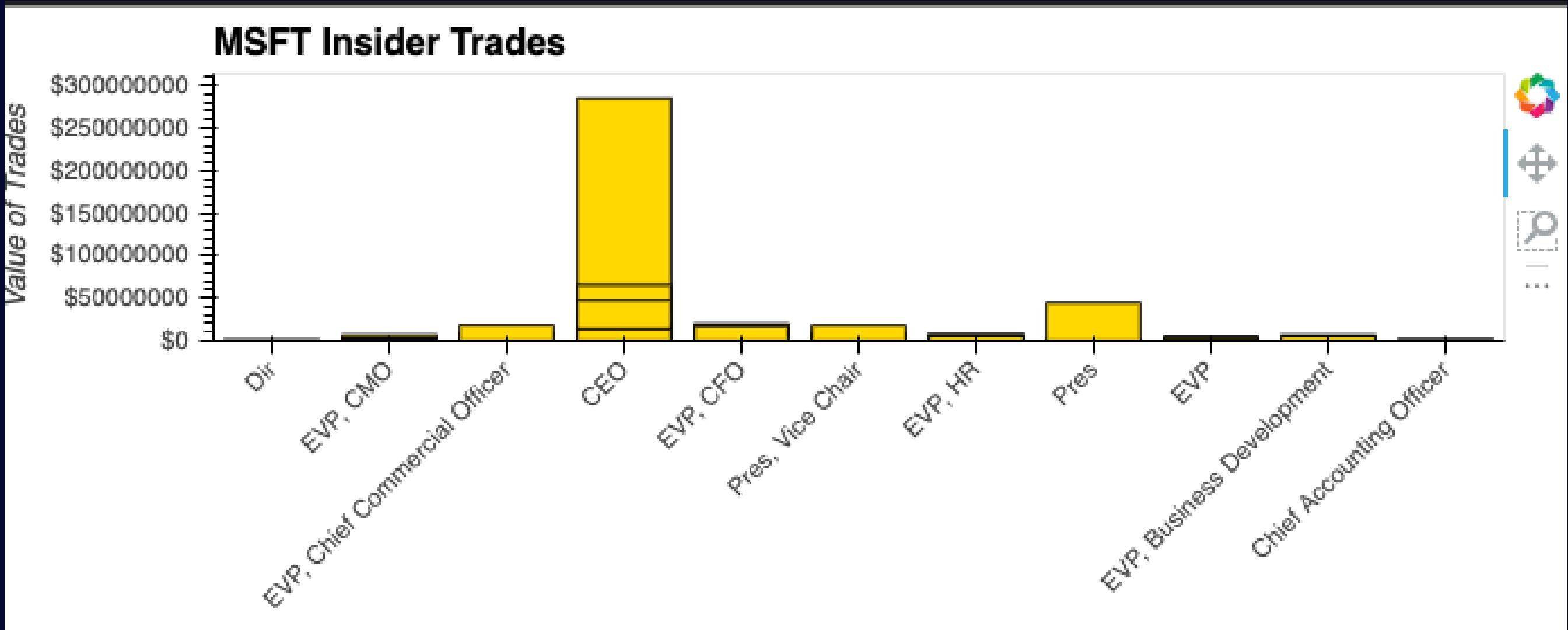
Data

- Filing Date
- Trade Date
- Ticker
- Insider Name
- Insider Title
- Trade Type
- Price
- Qty
- Owned
- Change in Ownership
- Values
- Close Price



Insider Trading Info MSFT - By Title

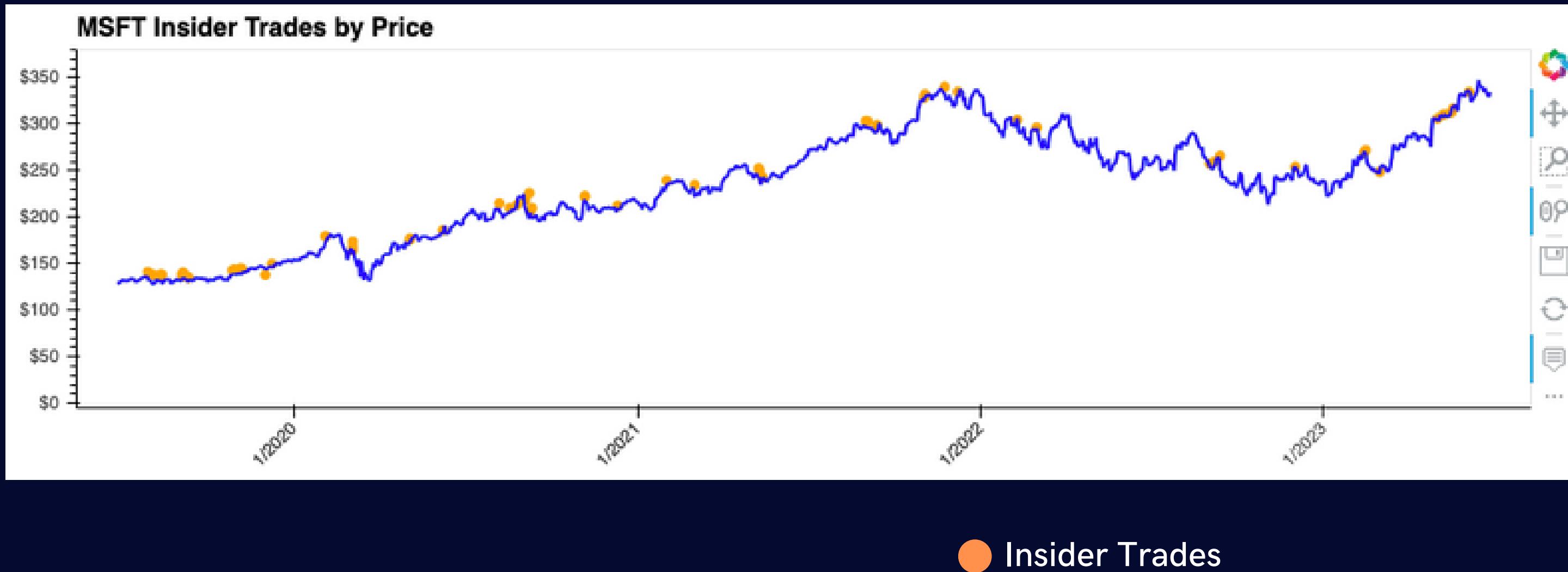
- Insider Trades and the value over the last 3 year time period



Insiders consist of directors and officers of the company, as well as any shareholders, owning 10% or more of the company's outstanding stock. The forms ask about the reporting person's relationship to the company and about purchases and sales of such equity securities.

Insider Trading w/ Stock Price - MSFT

- Insider Trades during the last 3 years for MSFT and stock closing price



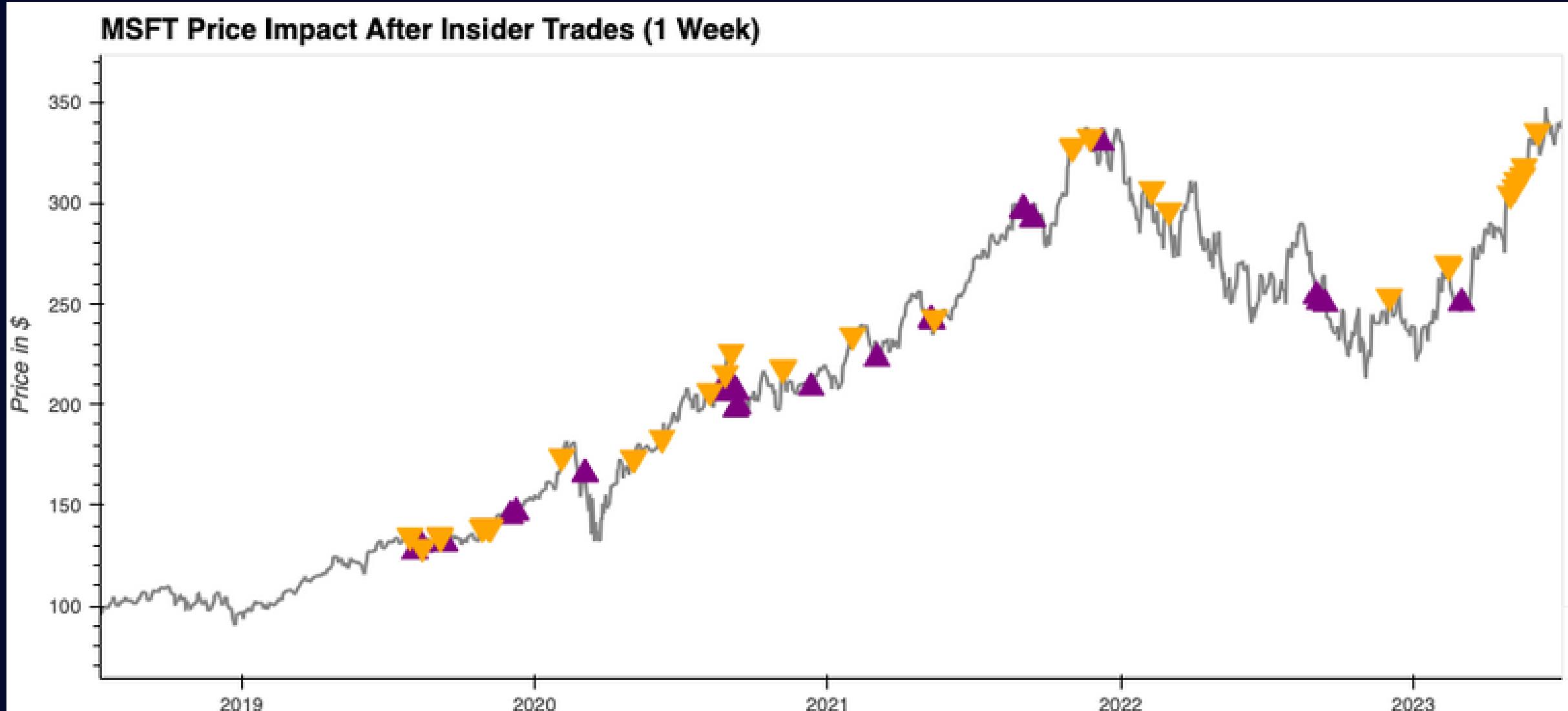
Insider Trading - Price Impact After 1 week - MSFT

Cleaned Data

- Filing Date
- Price (Insider Trade Price)
- Qty
- Owned
- Change in Ownership
- Values
- Close Price (Day of Trade)

Label

- Trend (0 - Stock price decreased, 1- Stock price increased)



- ▲ Price Increased - 1
- ▼ Price decreased - 0

Machine Learning Models Used

- Logistic Regression
- Support Vector Machine (SVM)
- Long short-term memory (LSTM)
- eXtreme Gradient Boosting (XGBoost)

Logistic Regression Model Results - Testing Data

Features		Precision	Recall	F1 Score	Accuracy
• Price (Insider Trade Price)	0	.59	.72	.65	.61
	1.0	.64	.50	.56	
• Qty	0	.61	.71	.66	.62
	1.0	.63	.53	.58	
• Owned	0	.61	.88	.72	.59
	1.0	.50	.17	.25	
• Change in Ownership	0	.65	.87	.75	.65
	1.0	.63	.32	.43	
• Values	0	.61	.71	.66	.62
	1.0	.63	.53	.58	
• Close Price (Day of Trade)	0	.61	.88	.72	.59
	1.0	.50	.17	.25	
Label					
• Trend (0 - Stock price decreased, 1- Stock price increased)	0	.61	.88	.72	.59
	1.0	.50	.17	.25	
Data Encoding / Splitting					
• Tran, Test, Split	0	.65	.87	.75	.65
	1.0	.63	.32	.43	
• Standard Scaler	0	.61	.71	.66	.62
	1.0	.63	.53	.58	

Logistic Regression Model Oversampled Results - Testing Data

Features		Precision	Recall	F1 Score	Accuracy
• Price (Insider Trade Price)	0	.66	.69	.68	.66
	1.0	.67	.63	.65	
• Qty	0	.61	.66	.58	.53
	1.0	.54	.41	.46	
• Owned	0	.58	.64	.61	.54
	1.0	.50	.44	.47	
• Change in Ownership	0	.48	.34	.40	.45
	1.0	.42	.56	.48	
• Values	0	.61	.66	.58	.53
	1.0	.54	.41	.46	
• Close Price (Day of Trade)	0	.58	.64	.61	.54
	1.0	.50	.44	.47	
Label					
• Trend (0 - Stock price decreased, 1- Stock price increased)	0	.58	.64	.61	.54
	1.0	.50	.44	.47	
Data Encoding / Sampling					
• Tran, Test, Split	0	.48	.34	.40	.45
	1.0	.42	.56	.48	
• Standard Scaler	0	.61	.66	.58	.53
	1.0	.54	.41	.46	
• Random Over sampler	0	.58	.64	.61	.54
	1.0	.50	.44	.47	

SVM Model Results - Testing Data

Features

- Price (Insider Trade Price)
- Qty
- Owned
- Change in Ownership
- Values
- Close Price (Day of Trade)

Label

- Trend (0 - Stock price decreased, 1- Stock price increased)

Data Encoding / Splitting

- Time sliced data (30 months Training)
- Standard Scaler

		Precision	Recall	F1 Score	Accuracy
TSLA	0	.45	1.0	.62	.50
	1.0	1.0	.14	.25	
AMZN	0	.50	.36	.42	.56
	1.0	.58	.71	.64	
MSFT	0	.67	.29	.40	.48
	1.0	.41	.78	.54	
GOOG	0	.54	.97	.69	.56
	1.0	.83	.15	.25	

LSTM Model Results - Testing Data

Features			Precision	Recall	F1 Score	Accuracy
<ul style="list-style-type: none"> • Price (Insider Trade Price) • Qty • Owned • Change in Ownership • Values • Close Price (Day of Trade) 	TSLA	0	0.00	0.00	0.00	0.58
		1.0	0.58	1.00	0.74	
	AMZN	0	0.71	0.56	0.63	0.62
		1.0	0.56	0.71	0.63	
	MSFT	0	0.67	1.00	0.80	0.67
		1.0	0.00	0.00	0.00	
Label	GOOG	0	0.60	0.67	0.63	0.56
		1.0	0.50	0.43	0.46	
Data Encoding / Splitting						
<ul style="list-style-type: none"> • Tran, Test, Split • Standard Scaler 						

XGboost Model Results - Testing Data

Features			Precision	Recall	F1 Score	Accuracy
• Price (Insider Trade Price)		0.0	0.74	0.79	0.77	
• Qty	TSLA	1.0	0.79	0.73	0.76	
• Owned		0.0	0.95	0.72	0.82	
• Change in Ownership	AMZN	1.0	0.79	0.96	0.87	0.85
• Values		0.0	0.82	0.69	0.75	
• Close Price (Day of Trade)	MSFT	1.0	0.2	0.33	0.25	0.62
Label		0.0	0.70	0.75	0.73	
• Trend (0 - Stock price decreased, 1- Stock price increased)	GOOG	1.0	0.63	0.57	0.60	0.68
Data Encoding / Splitting						
Train Test Split						

Analysis Report

- The purpose of the analysis is to examine how corporate insider trades change the trading environment.
- Insiders are generally contrarian investors and will sometimes purchase their own company's stock following recent stock underperformance that they may feel is unjustified.
- We found that Insider trades affect the closing price within 1 week after the filing date.
- comparing the 3 models the XGboost has outperformed the other evaluated models.
- Our findings reinforced the important role of insider trades in providing fundamental information and aiding price discovery.

Lessons Learned and Next Steps

Challenges faced

- Certain libraries tested did not get the data necessary
- Insufficient data available per Ticker for good training
- Data cleanup challenges
- Not all trades are created equal (Some of the stocks have minimal insider trades)

Future Improvements

- Adding simulations to backtest the strategies
- Immediate alerting when new Insider Trades are executed
- Comparing insider activities to sector and market conditions

Thank You!

