

Stock Predictions		Tweets	S&P500	Stocks	Inputs/Outputs	Comments
Pre Processing	I will be using a pre trained model, because creating an effective model on my own would be outside the scope of this project. Credit: Default 'Sentiment Analysis' model by transformers		Train_sp500 3,632 Records	Train_Main_Stocks 49,221 Records	TIK	Purpose: - To create a dashboard that uses ML Models to show relevant data used in determine the validity of an investment - Learn ML models, webdev, databases, and Python - Not a financial tool, not suitable for intra-day trading
ML Models	Tweets for the companies will need to be preprocessed by this model. TSA		Test_sp500 MSE: 1.477e-5 Type: Linear Regression	Test_StocksMain MSE: 1.434e-4 Type: Linear Regression		Challenges: - Getting data, especially historical tweet data that I can associate with an increase or decrease in stock price - Stacking ML models in a fashion that produces relevant and reliable predictions
Meta-Model	Each ML model will feed their results into this data set for Meta-Model		Train_tweetstock_data 19,460 Records	Test_tweetstock_data #	Model Input IN Model Output Out	- Each Machine Learning model will produce a result between -1 (don't invest) and 1 (invest). - The tweetstock_data is the only data that links all the models in one dataframe, as during preprocessing, tweetstock_data is ran through each model to apply the value before being inputted into the Meta-Model
		Tweets/SP500	Technical Analysis	Stock Prediction	User Input	Comments
Web Scrapping	TIK Recent Tweets TSA IN SP500 Data		TIK Financial Data List of Tiker's and Industry (and relevant Tech indicators) Financial Data from	Stock Info For Per/day information : the % change is calculated from open yesterday to open today. Per/min IN Per/day	TIK Ticker	User Inputs Company – Starts Process A List of Tickers was used because that was the data I could get on historical tweets. In a more sophisticated model you could extrapolate. Here using any ticker is outside the scope of the project.
Processing	Graph S&P Data for the day possibly choose day, year etc.		Not Attached to Meta – Model directly, support material for informed Decisions Calc. technical indicators/Ratio analysis based on industry standards	Make A Graph of Minute Data for the day that updates every minute Make Predictions from Meta-Model Out Stock Prediction		A prediction on investment health is the core purpose of the model
Further Development	I would like to be able to add relevant news articles: the three most recent articles scraped from the web every 5 minutes are so, all in a hope to get a better understanding of the greater impact on the stock price than just the current price.		This model is not intended to be a financial instrument but an education tool. Still, if the predictions turn out to be better than 70%, I will consider adding a recording function to collect the historical data of all the models and financial data, then re training the model to look at real time financial data, then attach it to the Alpaca API, which allows me to trade paper and then with a real portfolio depending on the efficiency of the model	To make this model suitable as a financial instrument, one limitation is the data I have collected, especially tweet data, which was downloaded from Kaggle and the reason this model has a set number of Tickers. A new model will need to work unsupervised using real time feedback from trading paper money, while using another function to determine which stocks to trade based off a whole new set of criteria.		Possible Layout of Dashboard → TIK – Price now Prediction Industry Relevant Accounts Ratio's Technical Indicators Stock Price graph Future: Relevant News Articles SP500 price Graph