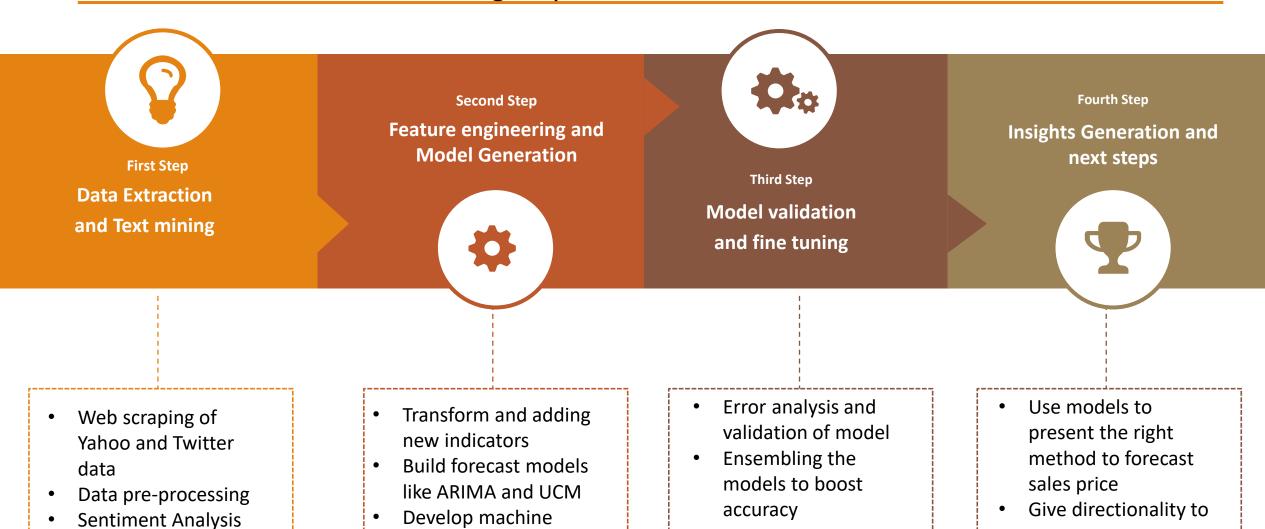
Stock Price Prediction

TEAM 8 - ADITI, SHWETA, ROSHAN

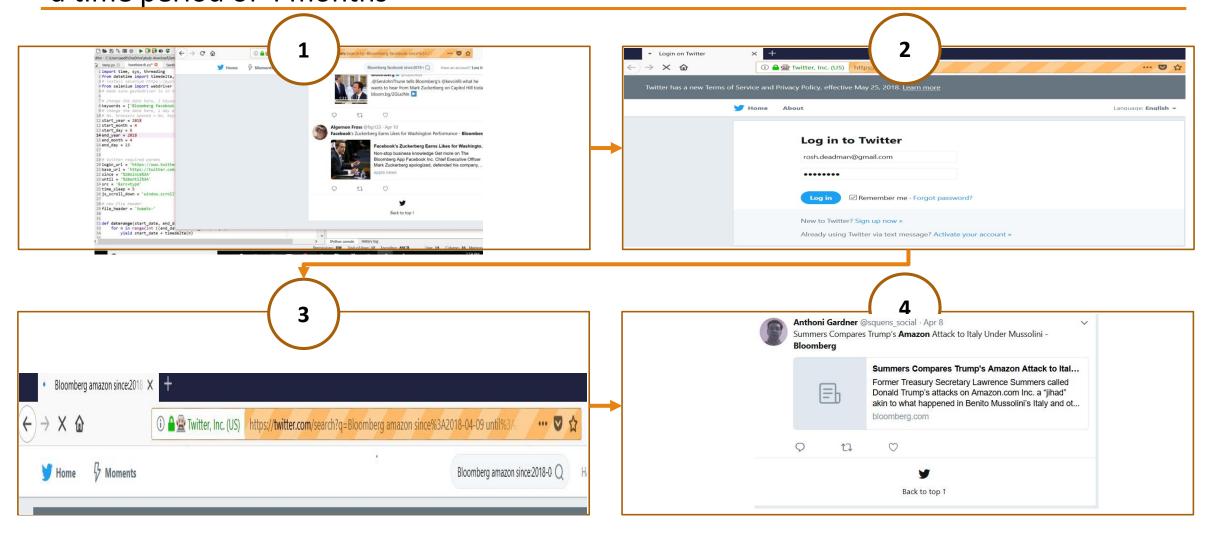
The steps involved in predicting stock price of Amazon, Facebook and Google has been conducted in the following way



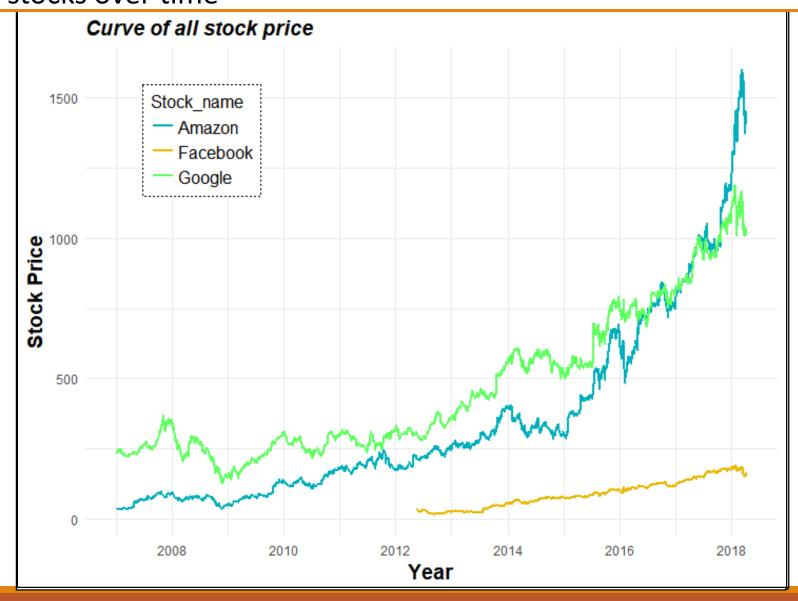
learning models

user

Twitter data was extracted using a Python shell script for all the stocks going back to a time period of 4 months



Exploratory data analysis was conducted to come up with high level insights about the behavior of stocks over time



To predict stock price with trend and sentiment analysis we have to take certain scenarios into consideration

Problem

• We do not carry the tweets for the future, therefore forecasting stock prices with tweets is not possible

Assumption

Date	Open price	High Price	Low Price	Close Price	Volume	pos_tweet	neg_tweet	Tweet_lag
1/16/2018	1273.39	1305.76	1273.39	1305.2	5443700	92%	8%	
1/17/2018	1323	1339.94	1292.3	1304.86	7220700	83%	17%	8%

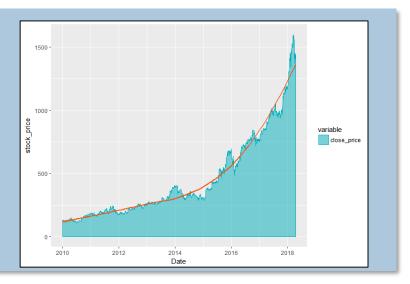
Solution

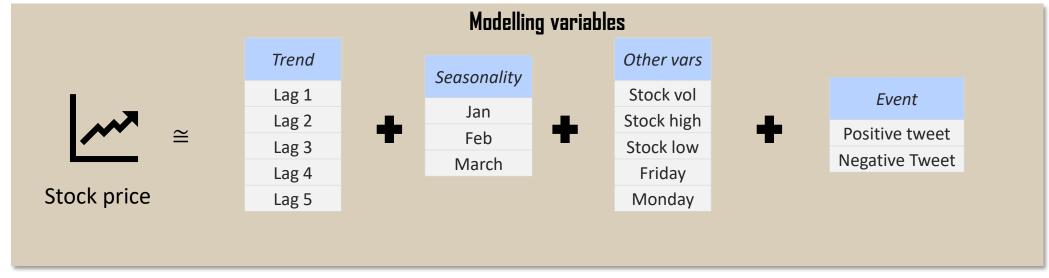
- Predict with lag of tweet Memory = 1 . Forecast for next day
- Predict univariate timeseries. Forecast for the next 30 days

Before getting into the modelling aspect, there are some considerations that went behind the process

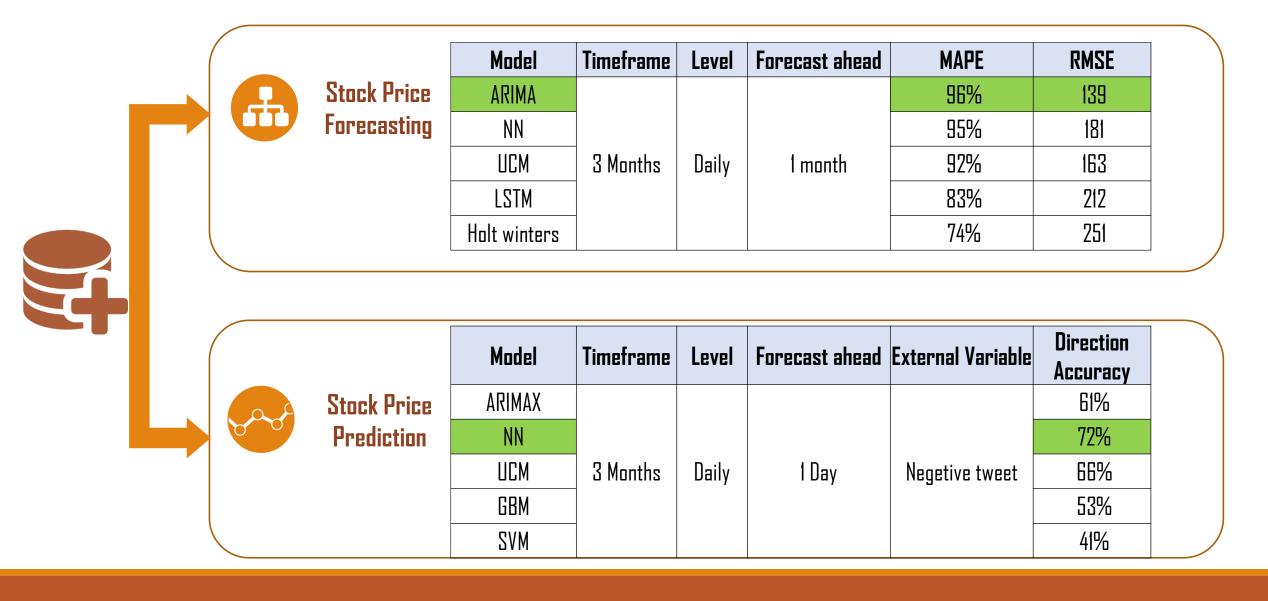
Modelling conditions

- Model has around 2083 points in last 7 years.
- Highly fluctuating and nonstationary
- To avoid generalization we build model on 3 months data and validate on 1 month

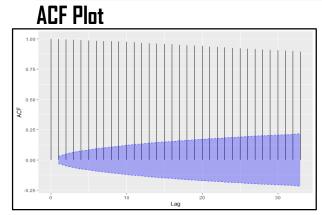


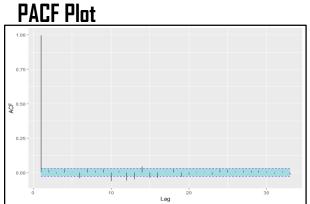


Amazon stock price forecasting - A quick summary on all modelling results for both cases



The Arima Model was built using the Auto Arima condition since the ACF and PACF plot did not show any clear indication of (P,D,Q)

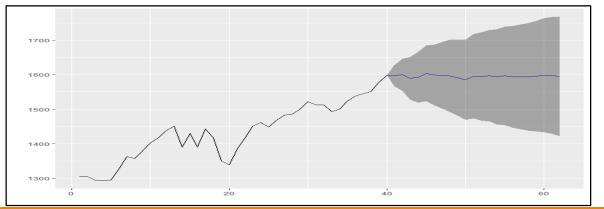




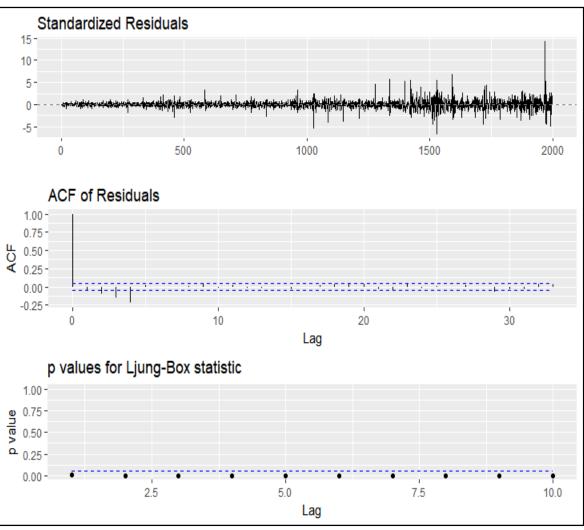
Model Equation		ARIMA (2,1,2) Series: Amazon stock					
		ar1	ar2	ma1	ma2	Neg tweet	
	Coefficients:	0.08	-0.73	0.06	1.00	-33.34	
	SE	0.21	0.15	0.14	0.12	16.20	

Sigma^2	536.4
AIC	370.91
RMSE	129.3
MAPE	6.75%
MAE	105.2

Forecast Plot



Residual Statistics

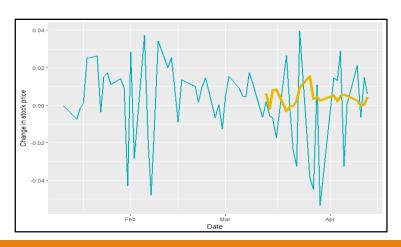


Time Series Forecasting for directionality is done with machine learning models that can better model the trend and has a good memory

Unobserved Component Model

A trend ,slope and irregular component model is considered for this exercise

	Actual State	Predicted State	# correct
Week 1		1111	3
Week 2	11111	11111	3
Week 3	11111	1111	4
Week 4	11441	11111	4

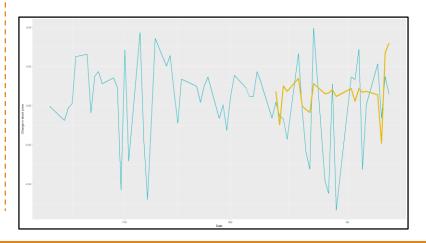


Accuracy	64%
Decision	Increase

Neural Network Model

Average of 20 networks, each of which is a 2-4-1 network with 17 weight with linear output units

	Actual State	Predicted State	# correct
Week 1	1111	1111	4
Week 2	11111	11111	3
Week 3	11111		4
Week 4	11441	1111	4



Accuracy	72%
Decision	Increase

We run the entire activity for the other two stocks and come out with a model for predicting price and forecasting the trend

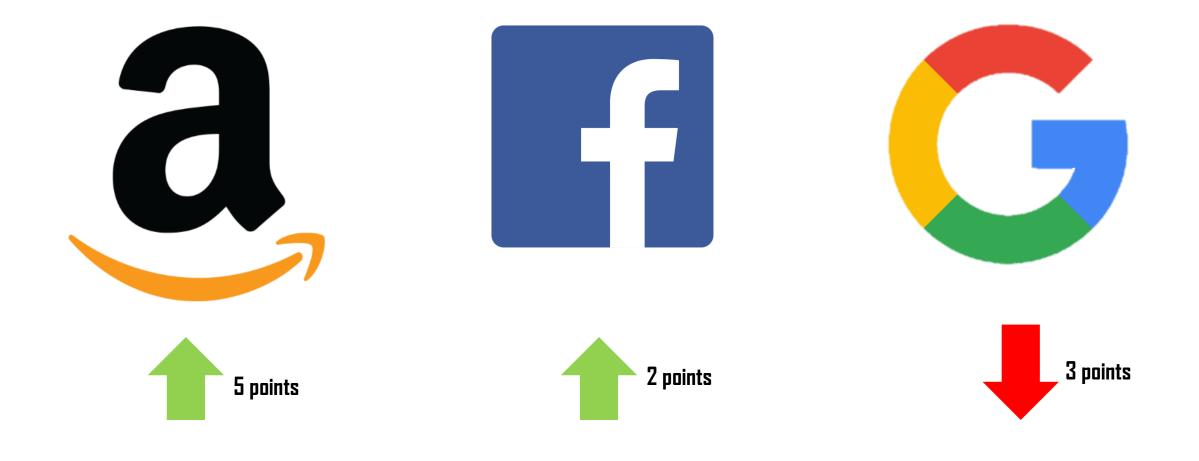
Time Series Forecasting

Stock	Model	Timeframe	Level	MAPE	Model
	ARIMA	3 Months	Daily	96%	Arima
AMAZON	NN			93%	
	ПСМ			92%	
	AMIRA	3 Months	Daily	84%	Arima
FB	NN			88%	
	ПСМ			89%	
	AMIRA	3 Months	Daily	86%	Arima
GOOGLE	NN			90%	
	ПСМ			92%	

<u>Time Series Regression</u>

Stock	Model	Timeframe	Level	Accuracy	Model
	ARIMA	3 Months	Daily	63%	NN
AMAZON	NN			72%	
	UCM			66%	
	AMIRA	3 Months	Daily	53%	NN
FB	NN			59%	
	UCM			55%	
	ARIMA	3 Months	Daily	48%	ШСМ
GOOGLE	NN			55%	
	UCM			59%	

Recommendations – The predicted value and direction of the stocks are as follows



Thank You