A Forecasting Study on Stock Prices

Abstract:
This paper analyzes different forecasting models on stock price datasets of Microsoft, Netflix, and Tesla. Stock price datasets are highly nonlinear and nonstationary. Daily close prices of Microsoft Corporation (MSFT), Tesla Inc. (TSLA) and Netflix Inc (NFLX) from Yahoo Finance for the time period of 2011 to 2021 are collected. Datasets from 2011 to 2020 are used to build the models and testing is done on the 2021 dataset. We have used five traditional univariate forecasting models, one global forecasting model and five deep learning models. We used five error metrics for evaluation: the Mean Absolute Scaled Error (MASE), Symmetric Mean Absolute Percentage Error (sMAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Modified Symmetric Mean Absolute Percentage Error (msMAPE).
Cross-Validation is done for all forecasting models on the test set using rolling windows. We have also ensembled some models by taking simple averages. This helps to check whether the ensembled model performed better than individuals for these stock price datasets or not.

Models Used:
We have used five traditional univariate forecasting models and one global forecasting model:
- Exponential Smoothing (ETS)
- Auto-Regressive Integrated Moving Average (ARIMA)
- Simple Exponential Smoothing (SES)
- Theta
- Trignometric Box-Cox ARMA Trend Seasonal (TBATS)
- Pooled Regression (PR)

We have also used five deep learning models:
- Simple Feed Forward Estimator (SFFE)
- DeepAR Estimator (DE)

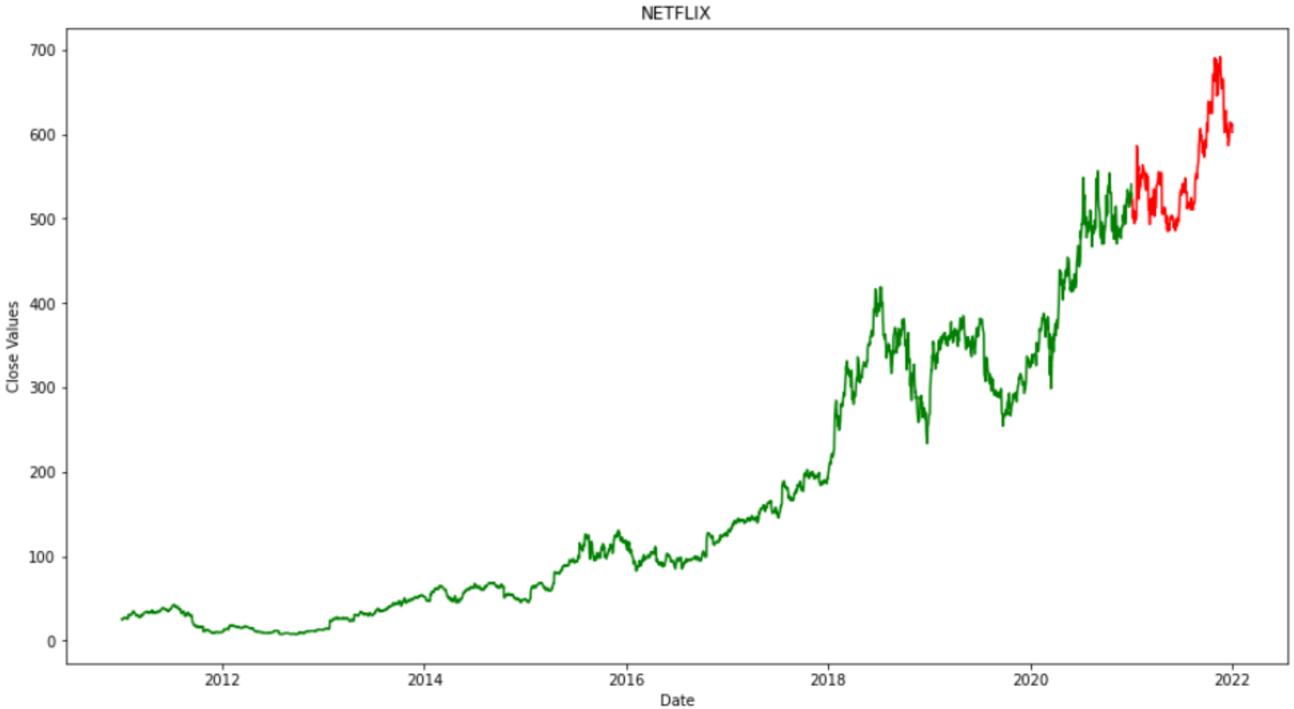
- NBeats Estimator (NBE)
- WaveNet Estimator (WNE)
- Transformer Estimator (TE)

Parameters used in Models common to all datasets:

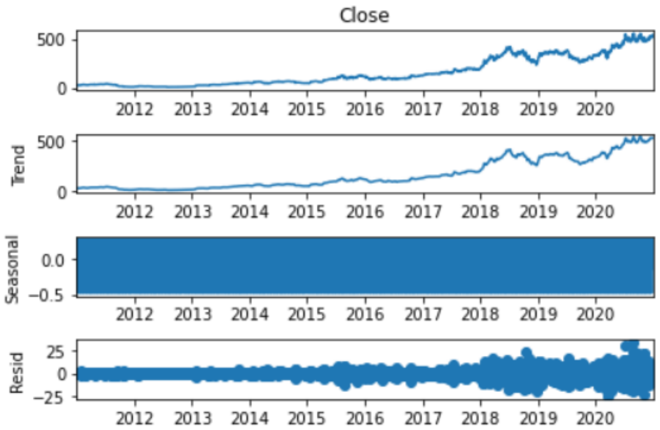
1. For SES, we have used auto-optimization (*alpha* value is not passed).
2. ETS model is applied with the *seasonal_periods* parameter as 7.
3. For Arima order chosen is (10,1,1) for (p,d,q). Here p is the number of lagged terms, d is the differencing factor, and q is the no. of differentials of the lagged terms.
4. For Pooled Regression model, we used simple linear formula with lag terms to be considered as 3.
5. For Theta Model, we considered the period of 7.
6. For TBATS, Default parameters were considered
7. Cross-validation is done on the test set using rolling windows. We have used sklearn *TimeSeriesSplit()* function with the number of splits as 10.

Netflix Dataset:

Netflix stock's close prices change from 2011 to 2021.



For these daily close prices, we plot the trend and seasonality with the period of 7.



- The Trend line does not clearly exhibit linear behaviour so we will use a multiplicative method.
- The additive method can be used for seasonality.
- The Error changes in magnitude as the series goes along so a multiplicative method will be used.
  This leaves us with an ETS(M, M, A) model.

Here are the error metrics values for all models:

	SES	ETS	Tbats	Theta	Arima	PR	SFFE	DE	NBE	WNE	TE
smape	8.0082919	19.36633	7.9971288	7.4113979	8.0364187	7.869929	13.234244	38.1669080	16.380227	12.2780051	27.606239
rmse	59.014252	134.91215	58.534856	51.791265	59.387384	59.38738	91.336816	322.599050	96.751137	74.9551782	198.04396
mae	45.281173	121.98612	45.218336	41.94980	45.434673	44.32972	78.629309	282.788142	86.071584	63.8928832	182.10147
msmape	8.0075812	19.36478	7.9964190	7.410740	8.0357055	7.869228	13.233124	38.1642968	16.378659	12.2768124	27.604131
mase	0.4417364	1.190024	0.4411234	0.4092375	0.4432338	0.432454	0.7670611	2.75871431	0.8339663	0.62330128	1.7764745

Ensembling Three Models : THETA, ARIMA and POOLED REGRESSION on NETFLIX Dataset

smape	7.047685312980247
rmse	47.81599406124055
mae	39.80368911869563
msmape	7.047058312382706
mase	0.3883013124637572

CROSS VALIDATION ON TEST SET USING ROLLING WINDOWS:

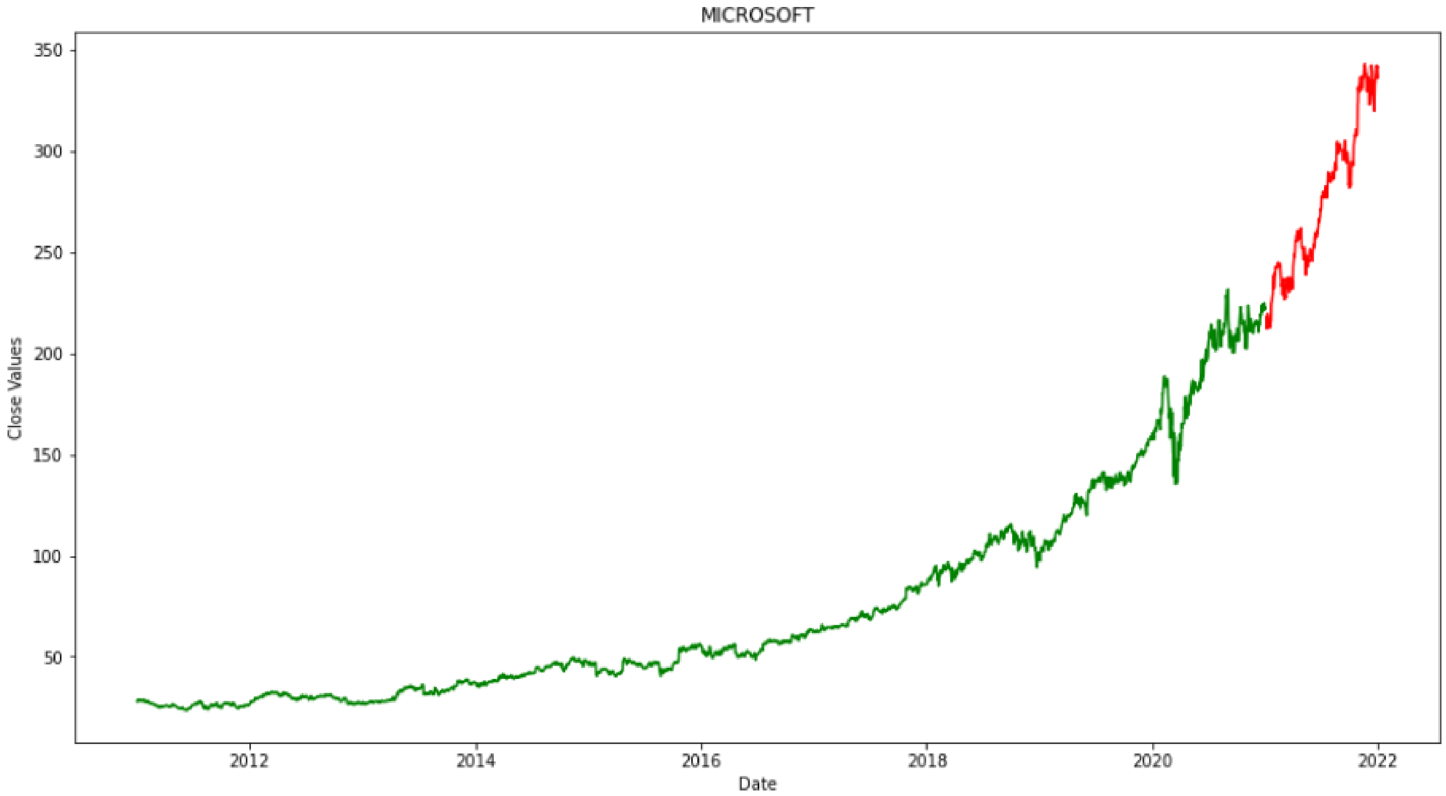
	SES	ETS	Tbats	Theta	Arima	PR
smape	3.915936	4.439498	3.918709	4.183429	4.6800323	5.057823
rmse	25.30951	28.63053	25.326615	26.890279	29.149663	31.460365
mae	22.44203	25.5183	22.46057	23.92601	26.571102	28.634440
msmape	3.915595	4.4391092	3.918364	4.183060	4.679617	5.05737415
mase	0.084530	0.0906442	0.084616	0.088152	0.0986844	0.1086191

Observations:

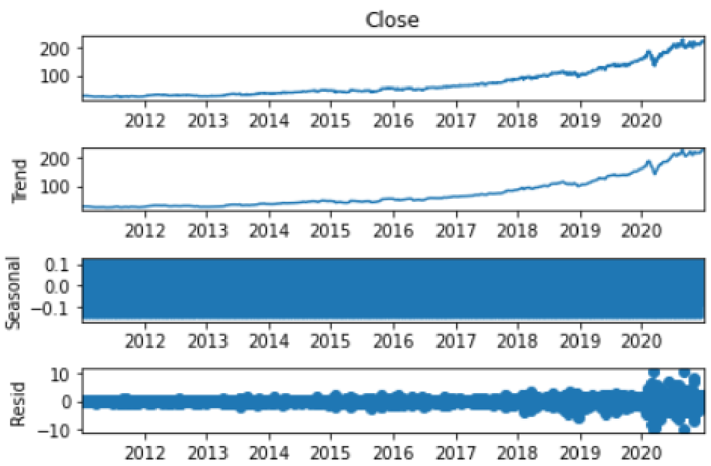
1. When we compare the error matrices of different models, we find that Theta Model has performed better than any other model considered.
2. We also observe that the deep AR estimator has performed the worst as compared to other models on this dataset.
3. The ensemble model (Simple Average) that we had taken which was the combination of Theta, ARIMA and Pooled Regression, performed better than any other model considered so far.
4. Now we compare the cross-validation on the test set using the rolling window on 6 primary models, here we found out that Simple Exponential Smoothing has better results as compared to any other model.

Microsoft Dataset:

Microsoft stock's close prices change from 2011 to 2021.



For these daily close prices, we plot the trend and seasonality with period of 7.



- The Trend line does not clearly exhibit linear behaviour so we will use a multiplicative method.
- The additive method can be used for seasonality.
- The Error changes in magnitude as the series goes along so a multiplicative method will be used. This leaves us with an ETS(M, M, A) model.

Here are the error metrics values for all models:

	SES	ETS	Tbats	Theta	Arima	PR	SFFE	DE	NBE	WNE	TE
smape	20.7767	13.32621	20.786168	19.09836	21.00866	7.0607028	9.5864189	3.41080450	2.9139043	4.77446319	11.888475
rmse	65.0150	44.82229	65.034190	60.56906	65.48782	25.471102	33.709616	11.0252646	10.316971	15.9093997	40.581723
mae	54.1045	36.39045	54.125806	50.24686	54.62817	20.071831	26.803605	9.40327428	8.0998483	13.1664962	33.971436
msmape	20.7727	13.32375	20.782162	19.09471	21.00461	7.0594494	9.5846917	3.41017577	2.9133707	4.77358388	11.886381
mase	1.80598	1.214699	1.8066982	1.677220	1.823467	0.6699898	0.8949386	0.31387761	0.2703697	0.43949249	1.233953

Ensembling Three Models : THETA, ARIMA and POOLED REGRESSION on MICROSOFT Dataset

smape	15.380818737951515
rmse	50.309062739052706
mae	41.41715065761898
msmape	15.37794776650536
mase	1.382488284071927

CROSS VALIDATION ON TEST SET USING ROLLING WINDOWS:

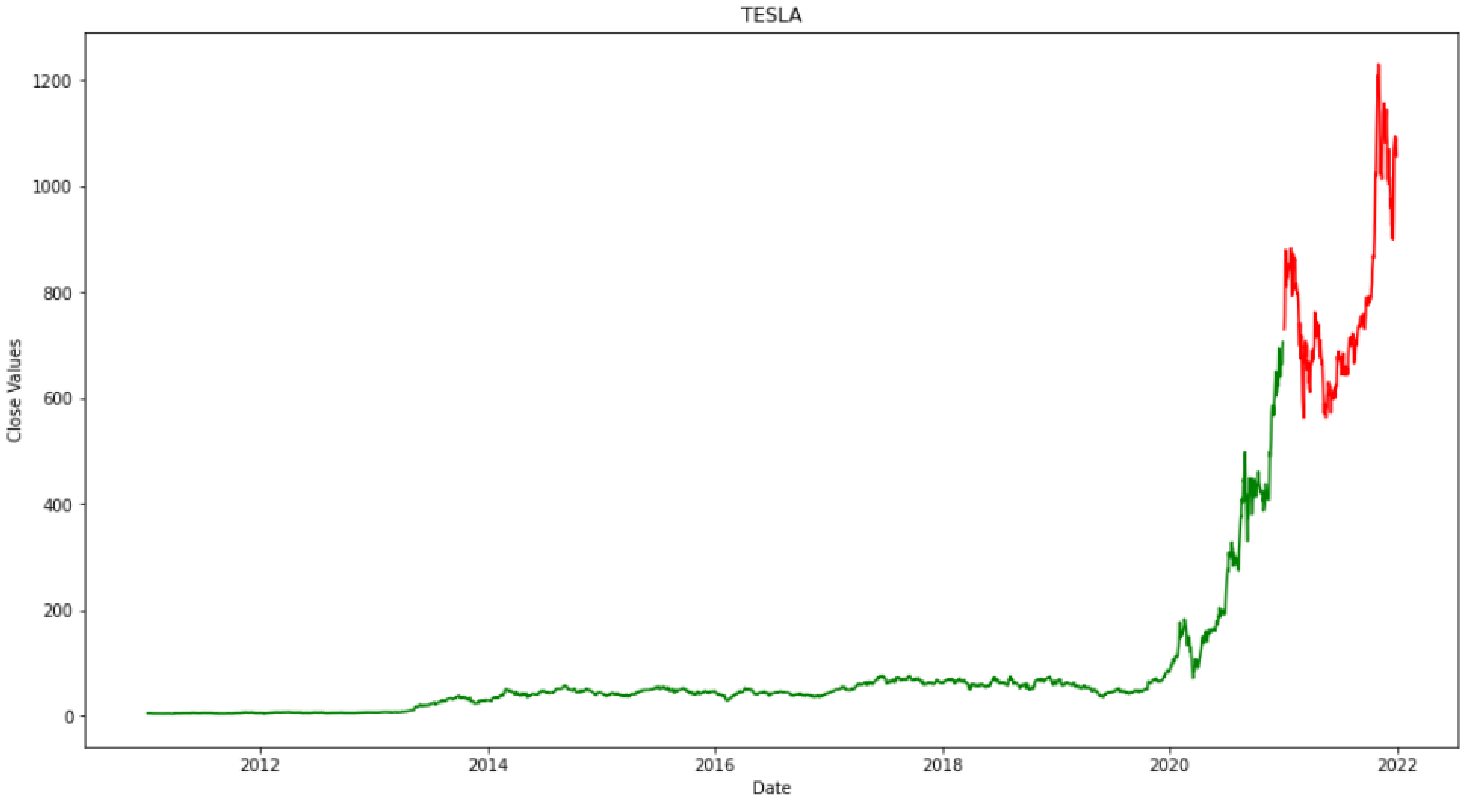
	SES	ETS	Tbats	Theta	Arima	PR
smape	3.4820870528897	3.4554863656092	3.4900405003291946	3.4687363896680	3.08847822756394	3.583788264970006
rmse	11.227761214835	10.819238375030	11.251671417670458	10.951478086471	10.1220393803803	11.47542091273699
mae	9.5781882410484	9.3645859677399	9.602191753129635	9.4367807645640	8.50566760622054	9.751924230743914
msmape	3.4814462057453	3.4548416971141	3.4893983650595652	3.4680914861571	3.08791034122463	3.583122262617063
mase	0.1195697563908	0.1135876638730	0.1199560127248126	0.1155003810700	0.10670866190971	0.120122424942506

Observations:

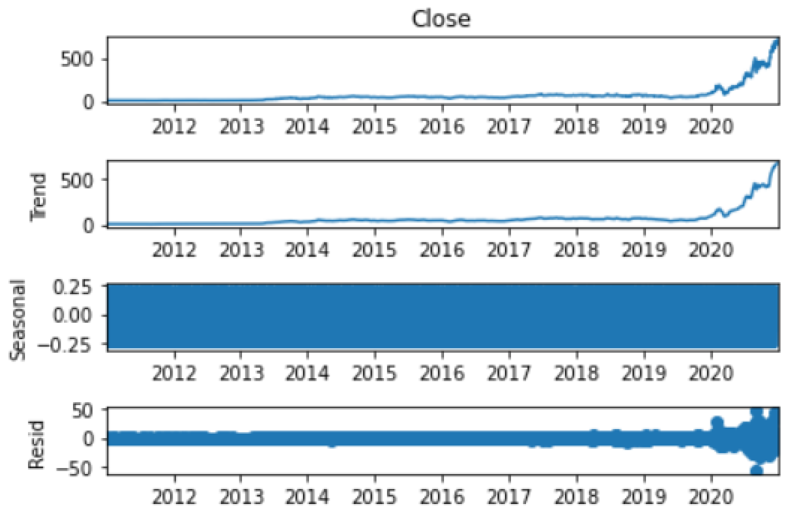
1. When we compare the error matrices of different models, we found out that the NBeats Estimator has performed better than any other model considered.
2. Globally Pooled Regression Model has performed better than other primary models.
3. We also observe that SES, Theta and ARIMA have performed the worst compared to all models on this dataset.
4. The ensemble model(Simple Average) that we had taken which was the combination of Theta, ARIMA and Pooled Regression, performed better than ARIMA and Theta but not as good as Pooled Regression Model.
5. Now we compared the cross-validation on the test set using the rolling window on 6 primary models, here we found out that ARIMA has better results as compared to other primary models.

Tesla Dataset:

Tesla stock's close prices change from 2011 to 2021.



For these daily close prices, we plot the trend and seasonality with the period of 7.



- The additive method can be used for trend and seasonality.
- The Error changes in magnitude as the series goes along so a multiplicative method will be used. This leaves us with an ETS(M, A, A) model.

Here are the error metrics values for all models:

	SES	ETS	Tbats	Theta	Arima	PR	SFFE	DE	NBE	WNE	TE
smape	15.175736	31.342726	15.170814	14.978155	15.091835	80.321017	14.438652	95.1604278	30.101821	81.650621	40.66140
rmse	177.33579	312.76168	177.24805	174.61867	170.75325	1703.54528	151.95105	544.201622	280.85606	499.71675	319.1014
mae	122.26106	283.42569	122.22458	120.80397	121.53699	1360.41732	116.08754	495.082257	214.89113	456.977336	275.13739
msmape	15.174779	31.340958	15.169857	14.977210	15.090882	80.3182901	14.437736	95.1508852	30.099680	81.6431765	40.658349
mase	2.1799224	5.0534980	2.1792720	2.1539424	2.1670122	24.2563263	2.0698482	8.82747757	3.8315224	8.14793463	4.9057170

Ensembling Three Models : THETA, ARIMA and POOLED REGRESSION on TESLA Dataset

smape	40.85607196343323
rmse	502.0318216712393
mae	429.35593408277151
msmape	40.854017321009106
mase	7.65907004289107

CROSS VALIDATION ON TEST SET USING ROLLING WINDOWS:

	SES	ETS	Tbats	Theta	Arima	PR
smape	7.930283974341518	8.588445123680632	9.15944206208291	8.459276005437797	8.295613105049005	8.935861249676188
rmse	75.38403163162234	81.77606296622819	89.00547181092199	79.82880682419187	79.32036034650335	83.03106311652434
mae	63.24351312850516	68.91854409221173	76.72114805610224	67.25300041194632	67.79442601514242	72.49670708252918
msmape	7.92976783076824	8.587889222499104	9.158868396418256	8.458722102535782	8.295081839966954	8.935290474327186
mase	0.101625610306345	0.109648611057281	0.126059472372557	0.107589392615708	0.109507198868172	0.11796813391321956

Observations:

1. When we compare the error matrices of different models, we find that Simple Feed Forward Estimator has performed better than any other model considered.
2. Theta, ARIMA, and SES have comparatively performed better.
3. We also observe that Pooled Regression Model has performed the worst compared to all other models.
4. The ensemble model(Simple Average) that we had taken which was the combination of Theta, ARIMA and Pooled Regression, performed worse than all the models.
5. Now we compared the cross-validation on the test set using the rolling window on 6 primary models, here we found out that Simple Exponential Smoothing has performed better than other primary models.

Conclusion:

Overall, the performance of TBATS and Theta is consistent with all stock price datasets whereas other models have varying results depending on the result. Deep learning Models perform horribly on very dynamic and fluctuating data. Note that the performance of models vastly depends on the parameters passed to them, stock price data is highly non-stationary and it is impossible for any model to work accurately on all stock price datasets. Thus in this paper, we have tried to approximate the performance of forecasting models on 3 stock prices datasets and present a benchmark for judging these models based on the results obtained.

TEAM MEMBERS:

Pratik Ratnadeep Ahirrao (IMT2019064)
Pratyush Upadhyay (IMT2019066)
Rohan Thakkar (IMT2019071)
Fahed Shaikh (IMT2019079)