



Research Project Progress Report

Week [5] – SIT723

Student Name:	Satvik Sharma
Supervisors' Names:	Prof. Maia Angelova
Project Title:	Financial Time Series Analysis with Machine Learning
SIT723 Target Grade:	Distinction I believe I have completed the tasks according to the deadline I chose. I have completed the literature review and have started working on the Project design.
Overleaf Project Link:	https://www.overleaf.com/project/623000d8d7e111b8d806686e
Project Folder Link:	https://deakin365.sharepoint.com/:f:/r/sites/SIT723-FinancialTimeSeriesAnalysiswithMachineLearning/Shared%20Documents/General/Satvik%20Sharma(218595095)?csf=1&web=1&e=laxNjw
Worklog:	162 hours https://deakin365.sharepoint.com/:x:/r/sites/SIT723-FinancialTimeSeriesAnalysiswithMachineLearning/Shared%20Documents/General/Satvik%20Sharma(218595095)/Worklog.xlsx?d=w2cfbd3ab588d42f58966474ef14f85c2&csf=1&web=1&e=2nMoH6
Project Plan	



Summary of the work planned with your supervisor:	In the first four weeks, the first and foremost task was to complete the Literature Review. The advantage of doing the literature review was to understand what kind of machine learning models are used by the people by the researchers. This will help out in creating the research artefact, and the initial plan for the research project has been started. The dataset will be pre-processed and create a machine learning model to predict the outcome for stock market for three years and five years.
Summary of the work done:	The literature review has been completed and a number of papers have been reviewed. The next part, that is, completing of the research design has been started along with the creation of the artefact. The machine learning model will be created in the Jupyter notebook. The data pre-processing has been started and the machine learning model will be made. Research on different machine learning algorithms is still in process and the best machine learning model(s) will be chosen for the project.
Next steps:	The next step is to create a machine learning model with different algorithms and finding out the best model. Then it will be the best idea to recreate the machine learning model by hybridisation.
Overall project progress:	The project is going on with the good speed and with the current progress the project will be completed on time.

Reflection against Assessment Rubric

Add comments and evidence against your target grade for each criterion in the assessment rubric. You should also provide evidence (relevant to your target grade) on how you meet all the relevant criteria for that grade based on your progress so far.

In case you have not started work on a specific criterion, e.g., research evaluation, leave it blank.

Your reflection on work done so far against the assessment rubric criteria	
Literature Review	Distinction
Technical & Academic Writing	Distinction
Research Design	Credit

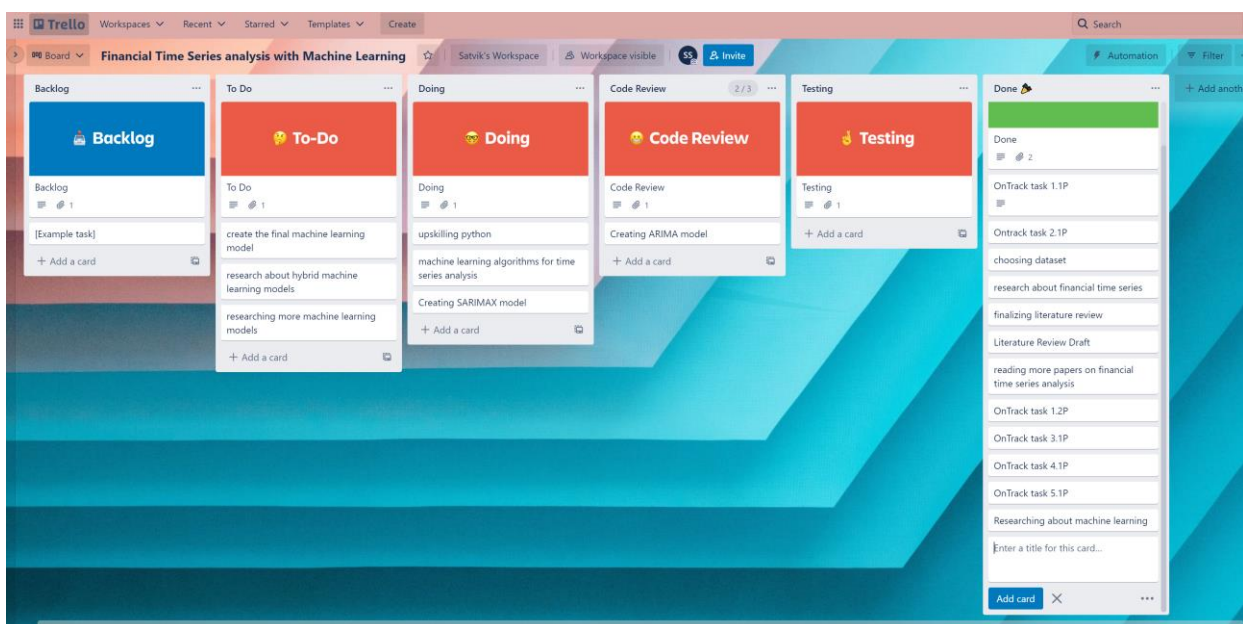


Project Management	Distinction
Artefact Development	Credit
Research Evaluation	[Add your comments, grade and relevant evidence here]
Research Dissemination	[Add your comments, grade and relevant evidence here]
OVERALL	Distinction

Additional Evidence

Add any evidence that supports your case for your target grade. For example, screenshot of git commits log, screenshot of your cloud storage folder, progress in your research report or artefact development, or a visual chart to illustrate project progress so far.

*This is your research project, and you need to demonstrate (**with evidence**) how you meet criteria for a certain grade. Make this project report easy to read and help anyone reading this report assess, with least effort, the progress you have made.*





General

Posts

Files

+

Meet

+ New

Upload

Sync

Copy link

Download

Open in SharePoint

All Documents

Documents

General

Satvik Sharma(218595095)

Name	Modified	Modified By
dataset	March 28	SATVIK SHARMA
Ontrack task	March 28	SATVIK SHARMA
Bibliography.docx	March 28	SATVIK SHARMA
literature review draft.docx	April 9	SATVIK SHARMA
Worklog.xlsx	34 minutes ago	SATVIK SHARMA



literature review draft.docx

Search

File Home Insert Layout References Review View Help Open in Desktop App Tell me what you want to do Editing Copy Link Comments Catch up Conversation Cl

Calibri (Body) 1 A⁺ A⁻ B I U Styles Find Reuse Files Dictate Editor Designer

show a small summary of what machine learning models were used and what was the conclusion of the papers

Model Name	Page	Performance criteria	Feature dataset	Time period	Environment	Conclusion
ARIMA, LR, MLP, SVM, DAE	[4]	-	S&P, Nasdaq, Dow 30	5-years (2012-2016)	Python Statmodels, Sklearn, Theano	Machine learning models perform better than ARIMA and SVM outperforms all other models
ARIMA, LSTM	[5]	% Reduction in RMSE	Dow Jones Index	1-month	Python with Keras and Theano	LSTM based models work better than ARIMA
SVM, ANN, KNN	[6]	RMSE	DAX 30, S&P500	10-years (2004-2014)	-	KNN outperforms other models
KNN, SVM, Gaussian process, MLP	[7]	-	M3 time series competition data	Monthly for thousand of time series	-	Best two models are MLP and gaussian progression
ANN, ARMAX, 3-D Hydrodynamic model	[8]	MAE, RMSE, R	Water level fluctuation in alpine lake	1-year (2009-2010)	Python Statmodels	The best working models are ANN and ARMAX
ARIMA, ANN, LSTM	[9]	P values, graphical observations	Dell's stock price	1 year (2010)	-	ANN model is better than ARIMA model and LSTM, but the hybrid of ARIMA, GARCH model can be used for more accuracy.
ANN, LSTM	[10]	RMSE, MAPE, R ²	Prediction of Solar irradiance	Hourly	Google Colab, Python, scikit-learn, Keras	Due to small dataset, the ANN model works better than the LSTM, but both works much better than the persistence model
ARIMA, GARCH, regime-switching	[11]	RMSE	Home price indices by the OFHEO	20-years (1980-2000)	-	Regime switching models perform better than ARIMA and GARCH
ARIMA-ANN, ARIMA-Kalman	[12]	MAE, MAPE, MSE	Wind speed prediction	-	-	Both hybrid models have good forecasting accuracy and suitable for wind samplings
ARIMA-SVM, ARIMA-ANN, ARIMA-Random Forest	[13]	-	Indian stock trend	5-years (2004-2009)	MATLAB6.1, SPSS13.0	The hybrid model ANN-ARIMA, was able to predict great values than other models
AR, ARDL, KNN, SVR, Naive, VAR, MLP	[14]	RMSE, R ²	Inflation forecasting	30-years (1984-2014)	-	SVR and ARDL outperforms other models and machine learning models work



jupyter

Financial analysis of Infosys ARIMA

Last Checkpoint: 20 hours ago (autosaved)

Python 3 (ipykernel)

FileEditViewInsertCellKernelWidgetsHelp

Trusted

+

⌕

📄

📁

⬆

⬇

▶ Run

■

↺

▶

Code

⌵

📧

In [739]:

from statsmodels.tsa.arima_model import ARIMA

In [775]:

```
model=ARIMA(train,order=(0,1,0))

pred=model.fit()

pred.summary()
```

C:\Users\Dell\anaconda3\lib\site-packages\statsmodels\tsa\arima_model.py:472: FutureWarning: statsmodels.tsa.arima_model.ARMA and statsmodels.tsa.arima_model.ARIMA have been deprecated in favor of statsmodels.tsa.arima.model.ARIMA (note the . between arima and model) and statsmodels.tsa.SARIMAX. These will be removed after the 0.12 release.

statsmodels.tsa.arima.model.ARIMA makes use of the statespace framework and is both well tested and maintained.

To silence this warning and continue using ARMA and ARIMA until they are removed, use:

import warnings
warnings.filterwarnings('ignore', 'statsmodels.tsa.arima_model.ARMA', FutureWarning)
warnings.filterwarnings('ignore', 'statsmodels.tsa.arima_model.ARIMA', FutureWarning)

warnings.warn(ARIMA_DEPRECATION_WARN, FutureWarning)

Out[775]:

ARIMA Model Results

Dep. Variable:	D.Prev Close	No. Observations:	3499			
Model:	ARIMA(0, 1, 0)	Log Likelihood	-18830.148			
Method:	csm	S.D. of Innovations	52.596			
Date:	Mon, 25 Apr 2022	AIC	37664.295			
Time:	00:25:16	BIC	37676.616			
Sample:	01-04-2011	HQIC	37668.692			
	- 08-02-2020					
	coef	std err	z	P> z	[0.025	0.975]
const	-0.7091	0.889	-0.798	0.425	-2.452	1.034