



PROJECT REPORT

STOCK PRICES PREDICTION USING MACHINE LEARNING

**SELLY MARGARETHA SUDIYANDI
19.K1.0046**

**Faculty of Computer Science
Soegijapranata Catholic University
2022**



HALAMAN PENGESAHAN

Judul Tugas Akhir: : STOCK PRICES PREDICTION USING MACHINE LEARNING

Diajukan oleh : Selly Margaretha Sudiyandi

NIM : 19.K1.0046

Tanggal disetujui : 22 Desember 2022

Telah setuju oleh

Pembimbing : R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

Penguji 1 : Yonathan Purbo Santosa S.Kom., M.Sc

Penguji 2 : Hironimus Leong S.Kom., M.Kom.

Penguji 3 : R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

Penguji 4 : Rosita Herawati S.T., M.I.T.

Penguji 5 : Y.b. Dwi Setianto S.T., M.Cs.

Penguji 6 : Yulianto Tejo Putranto S.T., M.T.

Ketua Program Studi : Rosita Herawati S.T., M.I.T.

Dekan : Dr. Bernardinus Harnadi S.T., M.T.

Halaman ini merupakan halaman yang sah dan dapat diverifikasi melalui alamat di bawah ini.

sintak.unika.ac.id/skripsi/verifikasi/?id=19.K1.0046

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Selly Margaretha Sudiyandi

ID : 19.K1.0046

declare that this work, titled "STOCK PRICES PREDICTION USING MACHINE LEARNING", and the work presented in it is my own. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at Soegijapranata Catholic University
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given.
5. Except for such quotations, this work is entirely my own work.
6. I have acknowledged all main sources of help.
7. Where the work is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Semarang, January, 03, 2023



Selly Margaretha Sudiyandi

19.K1.0046

HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : Selly Margaretha Sudiyandi
Program Studi : Teknik Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Menyetujui untuk memberikan kepada Universitas Katolik Soegijapranata Semarang Hak Bebas Royalti Noneksklusif atas karya ilmiah yang berjudul "**STOCK PRICES PREDICTION USING MACHINE LEARNING**". Dengan Hak Bebas Royalti Noneksklusif ini Universitas Katolik Soegijapranata berhak menyimpan, mengalihkan media/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan tugas akhir ini selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Semarang, 03 Januari 2023

Yang menyatakan



Selly Margaretha Sudiyandi

19.K1.0046

ACKNOWLEDGMENT

Gratitude and praise the author prays to God Almighty for all the grace offered so that the preparation of this thesis might be finished successfully. The title of the thesis that the author presents is "STOCK PRICES PREDICTION USING MACHINE LEARNING". The thesis was presented to meet graduation requirements at the Faculty of Computer Science, Informatics Engineering Study Program at Soegijapranata Catholic University. Many things occurred to the author while writing this thesis, ranging from hard effort to patience in comprehending the components required to accomplish the thesis in a better quality. The author realizes that many folks have contributed to the completion of the study and the final output. As a result, the author would like to express his profound thanks to :

1. Parents and family members who have provided unwavering support during lectures.
2. Mr. Robertus Setiawan Aji Nugroho S.T., MCompIT., Ph.D as a my supervisor that really assists me for my thesis to be done.
3. All IT lecturers at Unika.
4. Natalia Diva Muljono, Natasha Diva Muljono, and Vivian Davina Hendrawan, my friends who are always there for me when I need aid and who continue to encourage me when I'm down.
5. Everyone who has helped me in so many ways that I can't name them all.

Even though the author attempts to do his or her best, the author recognizes that this thesis is far from flawless. Finally, the writer anticipates criticism and ideas for the realization of excellent things from this thesis. The writer hopes that this thesis can be valuable for readers.

ABSTRACT

Prediction of stock price movements in the future will be an area that is widely researched. There is a hypothesis that it is considered impossible to predict stock prices, but it can also show that stock price forecasts can achieve a fairly high level of accuracy if properly formulated and modeled. This is because equity trading is one of the most important investment activities. Modeling and forecasting future stock prices based on current financial information can be very helpful to investors. They want to know if inventories go up or down in the short or long term. In this research, the author wants to analyze the comparison of accuracy and train the dataset using linear regression, lasso regression, LSSVM, LSTM, and CNN, then the accuracy will be calculated from the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). This can be used by investors in predicting stock prices using a more accurate algorithm. The findings reveal that the CNN model has a substantially lower accuracy value, while LSTM also performs well on specific datasets. However, there is one difference between these two models: the LSTM training time is slower than the CNN model. This is because computations in CNNs may occur in parallel (the same filter is applied to numerous circumstances at the same time), but LSTMs need to be processed sequentially, because the next step depends on the prior one.

Keyword: stock prices, regression, lstm, cnn, lssvm

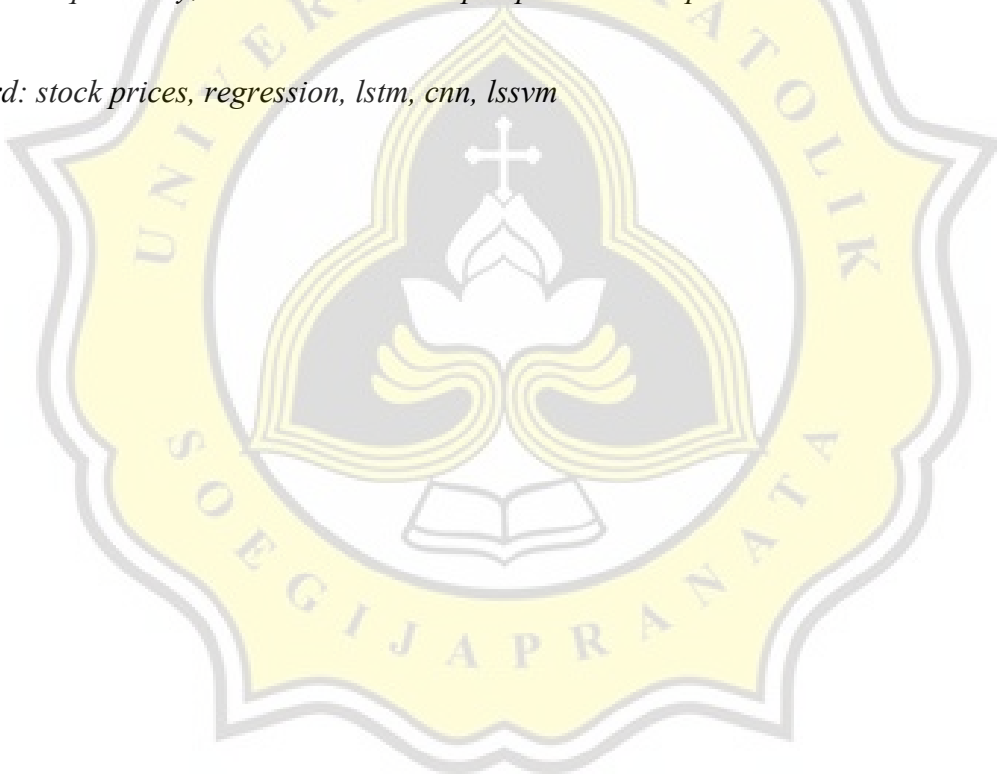


TABLE OF CONTENTS

COVER	i
DECLARATION OF AUTHORSHIP	iii
HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iv
ACKNOWLEDGMENT	v
TABLE OF CONTENTS	vii
LIST OF FIGURE	ix
LIST OF TABLE	xi
CHAPTER 1 INTRODUCTION	1
1.1. Background.....	1
1.2. Problem Formulation	1
1.3. Scope.....	1
1.4. Objective.....	2
CHAPTER 2 LITERATURE STUDY	3
CHAPTER 3 RESEARCH METHODOLOGY	8
3.1. Literature of Research.....	8
3.2. Data Preprocessing and Analysis.....	8
3.2.1. Data Collection	8
3.2.2. Data Selection	8
3.2.3. Data Visualization.....	9
3.2.4. Split Data	11
3.2.5. Feature Scaling	12
3.2.6. Models Evaluation.....	12
3.3. Algorithms	13
CHAPTER 4 DESIGN AND IMPLEMENTATION	15
4.1. Design	15

4.2.	Implementation	16
4.2.1.	Experiment Setup	16
4.2.2.	Data Visualization	16
4.2.3.	Data Selection	28
4.2.4.	Feature Scaling	29
4.2.5.	Split Data	30
4.3.	Deep-Learning Models	32
4.3.1.	LSTM	32
4.3.2.	CNN	36
4.4.	Evaluation Models	39
CHAPTER 5 RESULTS AND ANALYSIS		40
5.1.	Results	40
5.2.	Analysis	41
CHAPTER 6 CONCLUSION		47
REFERENCES		48
APPENDIX		a

LIST OF FIGURE

Figure 3.1 Heatmap.....	10
Figure 3.2 Box Chart.....	10
Figure 3.3 Bar Plot.....	11
Figure 3.1 LSTM Architecture.....	14
Figure 3.2 CNN Architecture.....	14
Figure 4.1 Flowchart.....	15
Figure 4.2 Dataset BAC.....	17
Figure 4.3 Dataset HDB.....	17
Figure 4.4 Dataset RY.....	18
Figure 4.5 Graphic of BAC.....	19
Figure 4.6 Graphic of HDB.....	20
Figure 4.7 Graphic of RY.....	21
Figure 4.8 Rolling 30 days.....	23
Figure 4.9 Open max value of BAC.....	24
Figure 4.10 Open max value of HDB.....	24
Figure 4.11 Open max value of RY.....	25
Figure 4.12 Gap between close price and open price.....	27
Figure 4.13 Heatmap of Dataset.....	28
Figure 4.14 Data Selection.....	29
Figure 4.15 LSTM Model.....	32
Figure 4.16 CNN Model.....	36
Figure 5.1 Mean Absolute Error of BAC.....	41
Figure 5.2 Mean Absolute Error of HDB.....	42
Figure 5.3 Mean Absolute Error of RY.....	43
Figure 5.4 Root Mean Squared Error of BAC.....	43

Figure 5.5 Root Mean Squared Error of HDB	44
Figure 5.6 Root Mean Squared Error of RY	45
Figure 5.7 Mean Absolute Percentage Error of BAC	45
Figure 5.8 Mean Absolute Percentage Error of HDB.....	46
Figure 5.9 Mean Absolute Percentage Error of RY	46



LIST OF TABLE

Table 5.1. Results of BAC	40
Table 5.2. Results of HDB	40
Table 5.3. Results of RY	41

