



PROJECT REPORT

BUS ROUTE DEMAND PREDICTION WITH DEEP LEARNING

Stevanus Alditian Lai
18.K1.0084

Faculty of Computer Science
Soegijapranata Catholic University
2022

HALAMAN PENGESAHAN



Judul Tugas Akhir: : BUS ROUTE DEMAND PREDICTION WITH DEEP LEARNING

Diajukan oleh : Stevanus Alditian Lai

NIM : 18.K1.0084

Tanggal disetujui : 14 Januari 2022

Telah setuju oleh

Pembimbing : Yonathan Purbo Santosa S.Kom., M.Sc

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

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

Penguji 3 : Rosita Herawati S.T., M.I.T.

Penguji 4 : Y.b. Dwi Setianto S.T., M.Cs.

Penguji 5 : R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

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=18.K1.0084

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Stevanus Alditian Lai

ID : 18.K1.0084

declare that this work, titled "Bus Route Demand Prediction with Deep 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, 17, 2022



Stevanus Alditian Lai

18.K1.0084

APPROVAL PAGE FOR PUBLICATION OF SCIENTIFIC PAPERS FOR ACADEMIC INTREST

The undersigned below:

Name : Stevanus Alditian Lai
Undergraduate Program : INFORMATICS ENGINEERING
Faculty : COMPUTER SCIENCE
Type of work : THESIS

Approved to give Non-Exclusive Royalty Free Right to Soegijapranata Catholics University Semarang for scientific work entitled “**BUS ROUTE DEMAND PREDICTION WITH DEEP LEARNING**”. Along with the existing tools (if needed). With this Non-Exclusive Royalty Free Right to Soegijapranata Catholics University has the right to store, transfer data/format, manage in form of database, maintain and publish this final project as long as I keep my name as a writer/creator and as a Copyright ownler.

This statement I made in truth.

Semarang, Januari, 17, 2000



Stevanus Alditian Lai

CHAPTER 118.K1.0084

ACKNOWLEDGMENT

Praise and gratitude the author prays to the presence of God Almighty for all the grace given so that the preparation of this thesis can be completed with great results. The title of the thesis that the author proposes is "BUS ROUTE DEMAND PREDICTION WITH DEEP LEARNING".

The thesis was submitted to fulfill graduation requirements at the Faculty of Computer Science, Informatics Engineering Study Program at Soegijapranata Catholic University. Many things that the author went through while writing this thesis, ranging from hard work to patience in understanding the aspects needed to complete the thesis so that it was of higher quality. The author understands that many persons have contributed to the completion of the study and the final project. Therefore, with great gratitude, on this moment the author would like to thank:

- 1 Parents and family who have give full support during lectures.
- 2 Mr. Yonathan as a my supervisor that realy helps me for this thesis to be done.
- 3 All Unika IT Lecturers.
- 4 Close friends who help self-development.
- 5 Unika has become a place for me to be ready in the world of work.

The author realizes that this thesis is far from perfect, even though the author tries to give the best of what the author can do. Finally, the writer expects criticism and suggestions for the achievement of good things from this thesis. The writer hopes that this thesis can be useful for readers, especially for other writers that .

Semarang, Januari, 17, 2000



Stevanus Alditian Lai

18.K1.0084

ABSTRACT

bus companies currently have several obstacles in providing their fleets from one city to another because of the highly dynamic demand from passengers, bus companies must be able to analyze which routes will have a lot of demand so that bus companies can provide more fleets on the routes that will have high demand. Unfortunately the bus company is currently still unable to predict which routes will be in high demand, at this time the bus company can only guess. Currently, to overcome this, the bus company has collected data which will later be analyzed.

Since the deep learning method is relatively new for bus company to predict the bus route demand, this study explores new method to make the bus company more profitable by trying to create and implement LSTM Autoencoder-Bi-LSTM Hybrid Models and Bi-LSTM to forecast bus route demand to support the decision making process in order to optimize bus fleet deployment each route.

The results shows that LSTM Autoencoder-Bi-LSTM Hybrid Models and Bi-LSTM models doesn't differ very much, the loss and metrics value differ a little, and both models performs quite well, but 1 thing that differs these 2 models, that is the training time, the autoencoders training time is very slow compared to models without autoencoders. This is normal for autoencoder to train slower than without it due to more network depth of the models with autoencoder.

Keyword: Autoencoders, deep learning, LSTM, Bi-LSTM

TABLE OF CONTENTS

APPROVAL PAGE FOR PUBLICATION OF SCIENTIFIC PAPERS FOR ACADEMIC INTREST.....	iv
ACKNOWLEDGMENT.....	v
ABSTRACT.....	vi
CHAPTER 1 INTRODUCTION.....	1
1.1. Background.....	1
1.2. Problem Formulation.....	3
1.3. Scope.....	3
1.4. Objective.....	4
CHAPTER 2 LITERATURE STUDY.....	5
CHAPTER 3 RESEARCH METHODOLOGY.....	10
3.1. Literature Review.....	10
3.2. Data Anlaysis and Preprocessing.....	10
3.2.1. Data Selection and Variable Analysis.....	10
3.2.2. Data Visualization.....	11
3.2.3. Split Data.....	13
3.2.4. Feature Scaling.....	13
3.4. Creating Models.....	14
3.4.1. LSTM Autoencoder-Bi-LSTM Hybrid Models.....	15
3.4.2. Bi-LSTM Model.....	16
3.5. Models Evaluation and Comparison.....	16
CHAPTER 4 ANALYSIS AND DESIGN.....	18
4.1. Data Preprocessing.....	18
4.1.1. Data Selection and Variable Analysis.....	18
4.1.2. Feature Extraction.....	20

4.1.3. Split Data.....	22
4.1.4. Data Analysis and feature scaling.....	22
4.2. Deeplearning Model.....	26
4.2.1. LSTM Autoencoder-Bi-LSTM Hybrid.....	26
4.2.2. Bi-LSTM Model.....	29
CHAPTER 5 IMPLEMENTATION AND RESULTS.....	31
5.1. Implementation.....	31
5.1.1. Pre-Processing.....	31
5.1.2. LSTM-Autoencoders-Bi-LSTM hybrid model.....	33
5.1.3. Bi-LSTM model.....	37
5.1.4. Prediction and inverse transformation.....	38
5.2. Result.....	41
5.2.1. LSTM-Autoencoders-Bi-LSTM hybrid model.....	41
5.2.2. Bi-LSTM model.....	45
5.2.3. Comparisons.....	48
CHAPTER 6 CONCLUSION.....	51
REFERENCES.....	53
APPENDIX.....	a

List of Figures

Figure 3.1: Date Time variable analysis.....	11
Figure 3.2: Heatmap.....	12
Figure 3.3: Gaussian distribution.....	13
Figure 3.4: Normalization.....	14
Figure 3.5: standarization formula.....	14
Figure 3.6: Autoencoder architecture.....	15
Figure 3.7: RMSE and MSE formula.....	16
Figure 4.1: Data Selection.....	18
Figure 4.2: start-destination transformation.....	19
Figure 4.3: Pivoting each unique startDest.....	20
Figure 4.4: date time feature extraction.....	21
Figure 4.5: data visualizarion.....	22
Figure 4.6: Heatmap.....	23
Figure 4.7: Data distribution.....	24
Figure 4.8: yeo-johnson formula.....	25
Figure 4.9: After using powertransformer.....	25
Figure 4.10: One Hot Encoder.....	26
Figure 4.11: LSTM encoder Layers.....	27
Figure 4.12: decoder-output Layer.....	28
Figure 4.13: LSTM-Autoencoders-Bi-LSTM hybrid model architecture.....	28
Figure 4.14: Bi-LSTM Model Layer.....	29
Figure 4.15: Bi-LSTM model Architecture.....	30
Figure 5.1: First prediction batch.....	39
Figure 5.2: prediction per sequence.....	39
Figure 5.3: prediction output.....	40
Figure 5.4: predicted output after inversed.....	41
Figure 5.5: training Loss and metric (a) first route, (b) second route, (c) third route, (d) fourth route, (f) fifth route, for LSTM-Autoencoders-Bi-LSTM hybrid model.....	42
Figure 5.6: prediction result compared with real value.....	44
Figure 5.7: training Loss and metric (a) first route, (b) second route, (c) third route, (d) fourth route, (f) fifth route, for Bi-LSTM model.....	45
Figure 5.8: prediction result compared with real value.....	47
Figure 5.9: Model 1(LSTM-Autoencoders-Bi-LSTM hybrid), Model 2(Bi-LSTM).....	48
Figure 5.10: loss and metrics visualization for both models.....	48
Figure 5.11: Average Loss and Metrics.....	49
Figure 5.12: prediction result with autoencoder (a), without autoencoder (b).....	49

Index of Tables

Table 5.1: Lowest Loss and metrics per route.....	43
Table 5.2: Lowest Loss and metrics per route.....	46

