



UNDERGRADUATE THESIS

PREDICTION OF THE NUMBER OF TOURIST VISITS IN EAST JAVA PROVINCE USING THE GATED RECURRENT UNIT (GRU) METHOD

Ahsan Nashihul Karim
NPM 20081010106

THESIS ADVISORS

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.
Ardhon Rakhmadi, S.Tr.T., M.Kom.

**MINISTRY OF EDUCATION, CULTURE, RESEARCH, AND TECHNOLOGY
NATIONAL DEVELOPMENT UNIVERSITY OF VETERANS EAST JAVA
FACULTY OF COMPUTER SCIENCE
INFORMATICS STUDY PROGRAM
SURABAYA
2026**

APPROVAL SHEET

PREDICTION OF THE NUMBER OF TOURIST VISITS IN EAST JAVA PROVINCE USING THE GATED RECURRENT UNIT (GRU) METHOD

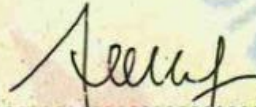
By :

Ahsan Nashihul Karim
NPM. 20081010106

It has been defended in front of and accepted by the Thesis Examiner Team of the Informatics Study Program, Faculty of Computer Science, National Development University of East Java Veterans on date.....

Approved,

Muhammad Muharrom Al Haromainy,
S.Kom., M.Kom.
NIP. 19950601 202203 1 006


..... (Advisor I)

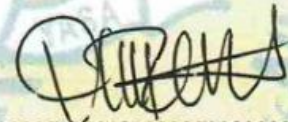
Ardhon Rakhmadi, S.Tr.T., M.Kom.
NIP. 19910805 202406 1 002


..... (Advisor II)

Retno Mumpuni, S.Kom., M.Sc
NIP. 1987 0716 2025 212 045


..... (Head Assessor)

Henni Endah Wahanani, ST. M.Kom.
NIP. 19780922 2021212 005


..... (Assessor I)

Acknowledge by,
Dean of the Faculty of Computer Science



Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

APPROVAL SHEET

PREDICTION OF THE NUMBER OF TOURIST VISITS IN EAST JAVA PROVINCE USING THE GATED RECURRENT UNIT (GRU) METHOD

By :

Ahsan Nashihul Karim
NPM. 20081010106

Approved by,

Informatics Study Program Coordinator

Faculty of Computer Science



Dr. Intan Yuniar Purbasari, S.Kom. MSc.

NIP. 19800602 202521 2 029

STATEMENT OF ORIGINALITY

I am the undersigned:

Name : Ahsan Nashihul Karim
NPM : 20081010106
Programs : Bachelor (S1)
Study Programs : Computer Science
Faculty : Computer Science

Stating that in this scientific document of this Thesis there is no part of other scientific works that have been submitted to obtain an academic degree in a Higher Education institution, and also there are no works or opinions that have ever been written or mentioned in full in the bibliography.

And I declare that this scientific document is free from elements of plagiarism. If in the future indications of plagiarism are found in this thesis, I willing to accept sanctions in accordance with applicable laws and regulations.

Thus I made this statement actually without any coercion from anyone and to be used as it should.

Surabaya, July 01 2026
Declarant,



Ahsan Nashihul Karim
NPM. 20081010106

ABSTRACT

Student Name / NPM : Ahsan Nashihul Karim/ 20081010106
Thesis Title :
Supervisor :
1. Muhammad Muharrom Al Haromainy, S.Kom.,
M.Kom.
2. Ardhon Rakhmadi, S.Tr.T., M.Kom.

The tourism sector in East Java Province faces serious challenges in predicting the number of tourist visits due to fluctuations influenced by the holiday season, religious holidays, and external shocks such as the COVID-19 pandemic. The conventional method that has been used by the East Java Provincial Tourism Office, namely trend extrapolation and simple historical averages, is not able to capture seasonal patterns or produce accurate projections up to the district/city level. This study applied the Gated Recurrent Unit (GRU) method to build a prediction model of the number of archipelago tourist visits per district/city in East Java Province based on monthly data from the Central Statistics Agency for the period January 2019 to December 2025, which covers all 38 districts/cities with a total of 3,192 data points. The model was built with an integrated panel data approach using a sliding window technique with a window size of 12 months, Min-Max Scaler normalization, and the sharing of training data (2019–2024) and test data (2025) based on time series split. The model architecture consists of a GRU layer with 64 units, a dropout layer with a rate of 0.8, and a dense layer. The training process showed good convergence with a train loss value of 0.015233 and a best val loss of 0.014702 with no indication of overfitting. Based on the results of the hyperparameter test scenario, the best configuration (S4-C) resulted in an RMSE value of 95,216.35 with a GRU unit configuration of 64, a window size of 12 months, a learning rate of 0.001, a dropout rate of 0.8, an epoch of 100, and a batch size of 32. Evaluation by region showed that the model worked best in cities with stable visit patterns such as Mojokerto City (RMSE 18,434) and Probolinggo City (RMSE 23,410), while areas with extreme fluctuations such as Malang Regency and Surabaya City produced higher RMSEs. The results of the study show that the GRU method is effectively used for the prediction of tourist visits in East Java Province and has the potential to be the basis for data-based strategic decision-making in the regional tourism sector.

Keywords : Gated Recurrent Unit, tourist visit prediction, deep learning, time series, East Java

ABSTRACT

Student Name / NPM : Ahsan Nashihul Karim/ 20081010106
Thesis Title:
Advisor: 1. Muhammad Muharrom Al Haromainy, S.Kom.,
M.Kom.
2. Ardhon Rakhmadi, S.Tr.T., M.Kom.

The tourism sector in East Java Province faces serious challenges in predicting tourist visit numbers due to fluctuations driven by holiday seasons, religious celebrations, and external shocks such as the COVID-19 pandemic. Conventional methods previously used by the East Java Provincial Tourism Office, namely trend extrapolation and simple historical averages, have proven inadequate for capturing seasonal patterns or producing accurate projections at the regency/city level. This study applies the Gated Recurrent Unit (GRU) method to build a prediction model for the number of domestic tourist visits per regency/city in East Java Province, using monthly data from the Central Bureau of Statistics (BPS) for the period January 2019 to December 2025, covering all 38 regencies/cities with a total of 3,192 data points. The model was developed using an integrated panel data approach with a sliding window technique (window size of 12 months), Min-Max Scaler normalization, and a time series split into training data (2019–2024) and testing data (2025). The model architecture consists of a GRU layer with 64 units, a dropout layer with a rate of 0.8, and a dense layer. Training showed good convergence with a train loss of 0.015233 and a best validation loss of 0.014702, with no indication of overfitting. Based on hyperparameter testing scenarios, the best configuration (S4-C) achieved an RMSE of 95,216.35 using GRU units of 64, window size of 12 months, learning rate of 0.001, dropout rate of 0.8, 100 epochs, and batch size of 32. Per-region evaluation showed the model performed best in cities with stable visitation patterns such as Mojokerto City (RMSE 18,434) and Probolinggo City (RMSE 23,410), while regions with extreme fluctuations such as Malang Regency and Surabaya City produced higher RMSE values. The results indicate that the GRU method is effective for tourist visit prediction in East Java Province and has strong potential as a data-driven strategic decision-making tool for regional tourism management.

Keywords: Gated Recurrent Unit, tourist visit prediction, deep learning, time series, East Java

FOREWORD

Praise be to Allah SWT for all His graces, guidance, and gifts to the author so that the thesis with the title "PREDICTION OF THE NUMBER OF TOURIST VISITS IN EAST JAVA PROVINCE USING THE GATED RECURRENT UNIT (GRU) METHOD" can be completed properly and on time. The author expresses his deepest gratitude to all parties who have helped both directly and indirectly during the process of preparing this thesis. In particular, the author would like to thank the:

1. Mr. Prof. Dr. Ir. Akhmad Fauzi, MMT., IPU. as the Rector of the "Veteran" National Development University of East Java.
2. Mrs. Prof. Dr. Ir. Novirina Hendrasarie, MT. as the Dean of the Faculty of Computer Science, National Development University "Veteran" East Java.
3. Mrs. Dr. Intan Yuniar Purbasari S.Kom. M.Sc. as the *Coordinator of the Informatics Study Program, Faculty of Computer Science, National Development University "Veteran" East Java.*
4. Mr. Muhammad Muharrom Al Haromainy, S.Kom., M.Kom, as Supervisor I who has provided briefing, support, and valuable input during the research process.
5. Mr. Rakhmadi, S.Tr.T., M.Kom., as Supervisor II who has provided briefing, support, and valuable input during the research process.
6. Mrs. Retno Mumpuni, S.Kom., M.Sc as the Chief Examiner who has provided input in compiling this thesis.
7. Mrs. Henni Endah Wahanani, ST. M.Kom. as a member of the examiner who has provided input in compiling this thesis.
8. All lecturers of the Informatics Study Program, Faculty of Computer Science, National Development University, East Java, "Veterans", for all the knowledge given to the author during the lecture period. Hopefully the writer can practice the knowledge that has been given and will be a useful provision for writers in the future.
9. My Parents who have provided material and Emotional support all the time, always praying for the writer in any condition. So that the author can complete the lecture from the beginning to the completion of this thesis.

The author realizes that in the preparation of this thesis there are still shortcomings. Therefore, constructive criticism and suggestions from various parties are highly

expected for the perfection of this work. Finally, with all the limitations that exist, the author hopes that this thesis report can provide benefits for all parties, especially in the development of science in the field of artificial intelligence.

Surabaya, July 01, 2026

Author

TABLE OF CONTENTS

APPROVAL SHEET	ii
APPROVAL SHEET	iii
STATEMENT OF ORIGINALITY	iii
ABSTRACT	v
FOREWORD	viii
TABLE OF CONTENTS.....	ix
IMAGE LIST	xi
TABLE LIST	xii
CHAPTER I INTRODUCTION	1
1.1. Background	1
1.2. Problem Formulation.....	4
1.3. Research Objectives	5
1.4. Research Benefits	5
1.5. Problem Limitations	6
CHAPTER II LITERATURE REVIEW	8
2.1. Previous Research	8
2.2. Tourism Visits and Tourism Indicators.	15
2.3. Time Series.	17
2.4. Deep Learning.	18
2.5. Long Short-Term Memory (LSTM).....	19
2.6. Gated Recurrent Unit (GRU).	21
2.7. Model Evaluation	22
CHAPTER III RESEARCH METHODOLOGY	23
3.1. Research Stages.....	24
3.2. Data Types and Sources.....	24
3.3. Literature Studies	25
3.4. Model Design	26
3.4.1. GRU	28
3.5. Model Evaluation	34
3.6. Testing Scenarios	36
3.7. Output	37

3.8. Results Analysis	38
CHAPTER IV RESULTS AND DISCUSSION	39
4.1. Data Preparation	39
4.2. Preprocessing Data	40
4.2.1. Data Normalization.....	41
4.2.2. Sequence formation	42
4.2.3. Data transformation	45
4.3. Dataset Sharing	46
4.4. Model GRU.....	48
4.5. Model Training.....	51
4.6. Model Evaluation	56
4.7. Testing Scenarios	63
4.8. Actual Comparison and Prediction.....	69
4.9. Results Analysis	70
CHAPTER V CONCLUSION & ADVICE	73
5.1. Conclusion	73
5.2. Suggestions	75
BIBLIOGRAPHY	76

IMAGE LIST

Figure 3. 1 Research Flow.....	23
Figure 3. 2 East Java Visitor Number Dataset	25
Figure 3. 3 Research Design.....	28
Figure 4. 1 Data input and drop results	40
Figure 4. 2 Data normalization	42
Figure 4. 3 Formation of squence	45
Figure 4. 4 <i>Data Transformation</i>	46
Figure 4. 5 Dataset Sharing	48
Figure 4. 6 GRU Build.....	51
Figure 4. 7 Training Results	54
Figure 4. 8 <i>Learning Curve GRU</i>	55
Figure 4. 9 Evaluation Results.....	59
Figure 4. 10 Bar Chart RMSE Per City District	60
Figure 4. 11 Actual Graph vs Prediction.....	61
Figure 4. 12 Best Scenario Outcomes.....	68
Figure 4. 13 Streamlit Display Models	70

TABLE LIST

Table 3. 1 Lumajang Regency Tourism Visit Data in 2019.....	32
Table 3. 2 Testing Scenarios	36
Table 4. 1 Test Scenario Results	65
Table 4. 2 Mojokerto City	69
Table 4. 3 Probolinggo City	69

