



PROJECT REPORT
PARIS HOUSING PRICE PREDICTION USING LINEAR
REGRESSION AND RANDOM FOREST

YOGA DINDA RIANA
17.K1.0053

Faculty of Computer Science
Soegijapranata Catholic University
2022

APPROVAL AND RATIFICATION PAGE



Judul Tugas Akhir : Paris Housing Price Prediction Using Linear Regression And Random Forest

Diajukan oleh : Yoga Dinda Riana

NIM : 17.K1.0053

Tanggal disetujui : 21 Desember 2022

Telah setuju oleh

Pembimbing : Rosita Herawati S.T., M.I.T.

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

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

sintak.unika.ac.id/skripsi/verifikasi/?id=17.K1.0053

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Yoga Dinda Riana

ID : 17.K1.0053

declare that this work, titled “Paris Housing Price Prediction Using Linear Regression And Random Forest”, 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, 12, 20, 2022



YOGA DINDA RIANA

17.K1.0053

HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : Yoga Dinda Riana

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 “PARIS HOUSING PRICE PREDICTION USING LINEAR REGRESSION AND RANDOM FOREST”. 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, 20 12 2022

Yang menyatakan



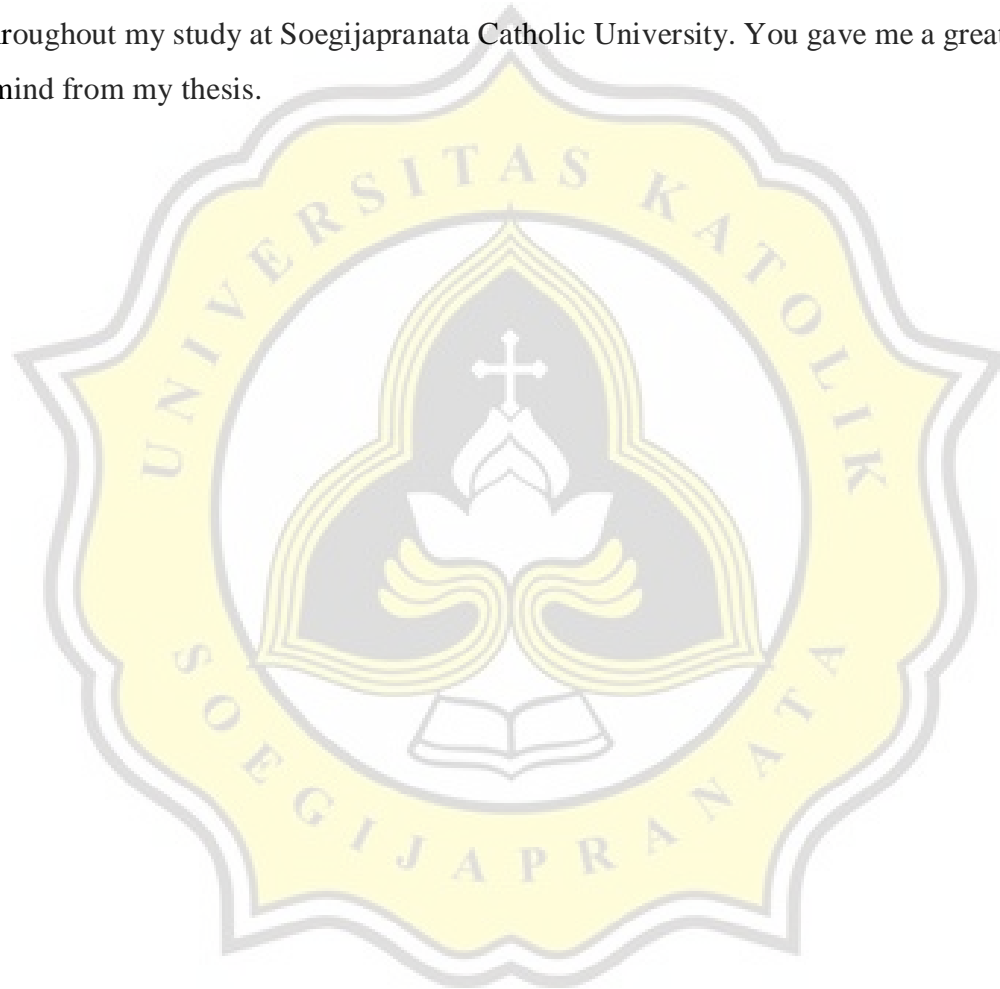
YOGA DINDA RIANA

17.K1.0053

ACKNOWLEDGMENT

I have received a myriad of support, advice, and assistance throughout this document writing. I would like to thank my supervisor Miss Rosita Herawati S.T., M.I.T. for formulating this topic. I would also like to thank my friend Victor Johsua and clementino for guiding me with advice to finish this document.

I would like to thank my family and friends for giving me ceaseless love, support, and advice throughout my study at Soegijapranata Catholic University. You gave me a great escape to rest my mind from my thesis.



ABSTRACT

This project has the purpose to get the result of the accuracy of linear regression and random forest for a test which one of those two has a better result for regression using data "Paris housing price". That data has been processed and already sampling data, then using linear regression and random forest to get a prediction, and then comparing those two methods to get a prediction result, which one of those two has a better prediction. The results of research Comparing the Linear Regression and Random Forest Algorithms in the regression of the Paris housing dataset. It can be concluded that linear regression is accurate works in this study and is also it suitable for this dataset, but for the random forest, the answer is no because random forest output has too many errors.

Keywords: Regression, comparing method, linear regression, random forest, housing price



TABLE OF CONTENTS

COVER	i
APPROVAL AND RATIFICATION PAGE	ii
DECLARATION OF AUTHORSHIP	iii
HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	iv
ACKNOWLEDGMENT	v
ABSTRACT	6
TABLE OF CONTENTS	vii
LIST OF FIGURE	ix
LIST OF TABLE	x
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	6
3.1. Literature Study.....	6
3.2. Data and Variable.....	6
3.3. Orange Data Mining Tool.....	6
CHAPTER 4 ANALYSIS AND DESIGN	7
4.1. Collecting data.....	7
4.2. Linear Regression	7
4.3. Random Forest	9

4.4. Desain	10
4.5. Result	11
4.6. Writing report.....	11
CHAPTER 5 IMPLEMENTATION AND RESULTS	12
5.1. Linear regression steps.....	12
5.1.1. Importing CSV data.....	12
5.1.2. Selecting attribute	13
5.1.3. Data sampler	14
5.1.4. Making Prediction.....	15
5.2. Random forest steps.....	17
5.2.1. Importing CSV.....	18
5.2.2. Selecting attribute	19
5.2.3. Data sampler	20
5.2.4. Making prediction.....	21
5.3. Results.....	23
CHAPTER 6 CONCLUSION.....	28
REFERENCES	29

LIST OF FIGURE

Figure 4.1 Data sample	7
Figure 4.1 Flowchart	10
Figure 5.1 CSV Import Linear Regression	12
Figure 5.2 Attributes select Linear Regression	13
Figure 5.3 Data Sampler Linear Regression.....	14
Figure 5.4 Making Prediction Linear Regression	15
Figure 5.5 Linear regression prediction output	16
Figure 5.6 CSV Import Random Forest.....	18
Figure 5.7 Selecting Attributes Random Forest	19
Figure 5.8 CSV Import Random Forest.....	20
Figure 5.9 Making Prediction Random Forest	21
Figure 5.10 Prediction Output Random Forest.....	22
Figure 5.11 Linear Regression and Random Forest Orange Output	24
Figure 5.12 Graph comparison	25
Figure 5.13 RMSE & MAE Graph.....	26

LIST OF TABLE

Table 5.1. Comparison MSE, RMSE, and MAE.....	24
Table 5.2. Comparison MAE and RMSE	25

