



PROJECT REPORT
WINE QUALITY PREDICTION USING ADABOOST
AND RANDOM FOREST

GILBERTUS SURYO RAHARJO
17.K1.0049

Faculty of Computer Science
Soegijapranata Catholic University
2022

APPROVAL AND RATIFICATION PAGE



Judul Tugas Akhir : WINE QUALITY PREDICTION USING
ADABOOST AND RANDOM FOREST

Diajukan oleh : GILBERTUS SURYO RAHARJO

NIM : 17.K1.0049

Tanggal disetujui : 13 January 2023

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 : 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=17.K1.0049

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Gilbertus Suryo Raharjo

ID : 17.K1.0049

declare that this work, titled “Wine Quality Prediction Using AdaBoost 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, 16 January 2023



GILBERTUS SURYO RAHARJO

17.K1.0049

HALAMAN PERNYATAAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan dibawah ini:

Nama : GILBERTUS SURYO RAHARJO

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 “Wine Quality Prediction Using ADABOOST 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, 16 January 2023

Yang menyatakan



GILBERTUS SURYO RAHARJO

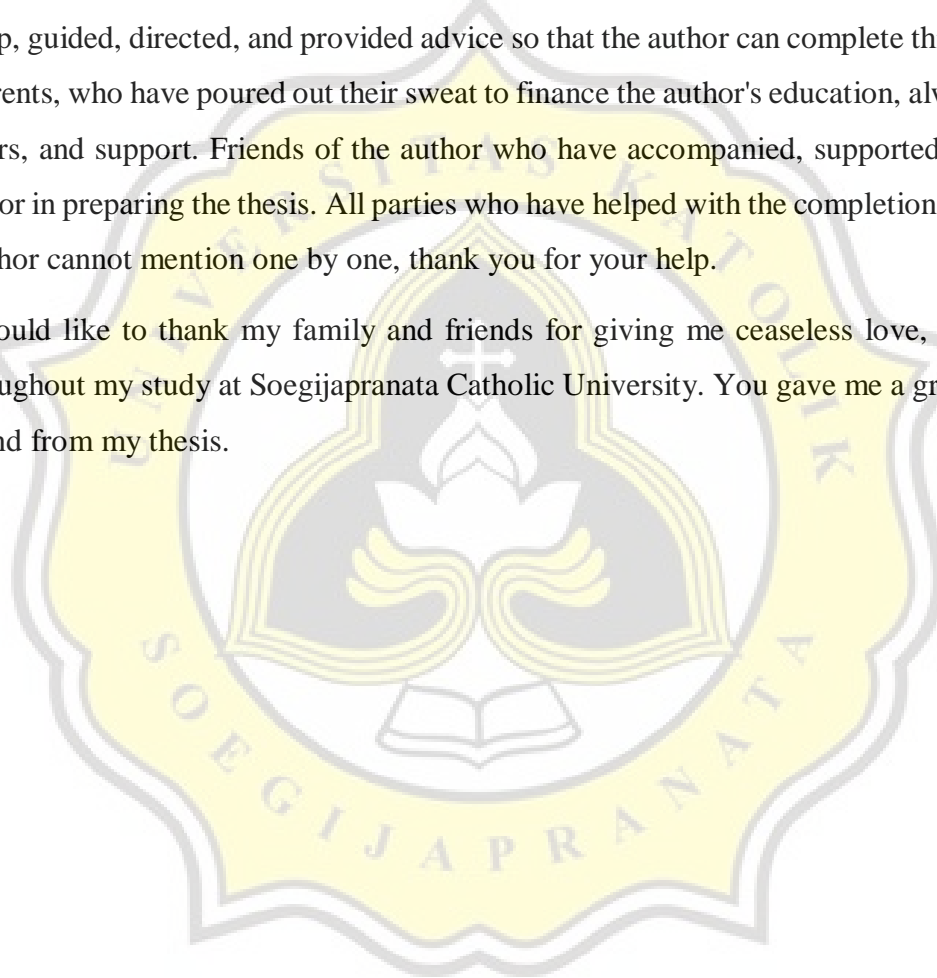
17.K1,0049

ACKNOWLEDGMENT

Praise be to God Almighty who has always given His abundance of grace and grace so that the author can complete the thesis with the title "Wine Quality Prediction Using AdaBoost and Random Forest"

I have received a great deal of support, advice, and help during the writing of this document. I would like to thank my supervisor Rosita Herawati S.T., M.I.T who has taken the time, to help, guided, directed, and provided advice so that the author can complete this thesis. The author's parents, who have poured out their sweat to finance the author's education, always provide love, prayers, and support. Friends of the author who have accompanied, supported, and prayed for the author in preparing the thesis. All parties who have helped with the completion of this thesis that the author cannot mention one by one, thank you for your help.

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 aims to test the accuracy of the AdaBoost and Random Forest methods of the two methods to test the level of good alcohol content in wine. In this study using the data taken is alcohol, and from the data processed and sampled data then the two methods get prediction output results and from the two methods can determine which results are better in determining predictions. From the results of the AdaBoost and Random Forest research in the dataset looking for the best alcohol quality levels in wine, it can be concluded that the AdaBoost method gets more accurate and suitable results than the Random Forest method.

Keyword: AdaBoost, Random Forest, wine quality, comparing method



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	vi
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	1
CHAPTER 2 LITERATURE STUDY	2
CHAPTER 3 RESEARCH METHODOLOGY	5
3.1. Literature study.....	5
3.2. Data and variable.....	5
3.3. Orange Data Mining Tool	5
CHAPTER 4 ANALYSIS AND DESIGN	6
4.1. Collecting data.....	6
4.2. AdaBoost.....	6
4.3. Random Forest	8
4.4. Desain	9

4.5. Result	9
CHAPTER 5 IMPLEMENTATION AND RESULTS.....	10
5.1. AdaBoost Steps	10
5.1.1. Importing CSV data	10
5.1.2. Selecting Attribute	11
5.1.3. Data Sampler	12
5.1.4. Making Prediction.....	13
5.2. Random Forest Steps	15
5.2.1. Importing CSV.....	15
5.2.2. Selecting Attribut.....	16
5.2.3. Data Sampler	17
5.2.4. Making Prediction.....	18
5.3. Results.....	20
CHAPTER 6 CONCLUSION.....	25
REFERENCES	26
.....	a

LIST OF FIGURE

Figure 4.1 Data sample	6
Figure 4.2 Flowchart.....	9
Figure 5.1 Import CSV AdaBoost	10
Figure 5.2 Set target AdaBoost	11
Figure 5.3 Data Sampler AdaBoost	12
Figure 5.4 Making Prediction AdaBoost	13
Figure 5.5 Output AdaBoost	14
Figure 5.6 Import CSV Random Forest.....	15
Figure 5.7 Set target Random Forest	16
Figure 5.8 Data Sampler Random Forest.....	17
Figure 5.9 Prediction Random Forest	18
Figure 5.10 Output Random Forest	19
Figure 5.11 Output AdaBoost and Random Forest	20
Figure 5.12 Graph comparison.....	22
Figure 5.13 RMSE graph.....	23
Figure 5.14 MSE graph.....	23

LIST OF TABLE

Table 5.1. Data Sampler.....	21
Table 5.2. Result.....	22

