APPLICATION OF DATA MINING FOR STUDENT SATISFACTION ANALYSIS OF OFFLINE LEARNING PERFORMANCE IN THE DEPARTMENT OF PUBLIC HEALTH WITH THE C4.5 ALGORITHM METHOD

THESIS

Submitted in partial fulfillment of the requirements for achieving a Bachelor of Computer Science

SUBMITTED BY:

FARID MAJID DARMAJATI

1811102441108

DEPARTMENT OF INFORMATICS

FACULTY OF SCIENCE AND TECHNOLOGY

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

SAMARINDA

2022

Application of Data Mining for Student Satisfaction Analysis of Offline Learning Performance in the Department of Public Health With the C4.5 Algorithm Method

THESIS

Submitted in partial fulfillment of the requirements for achieving a Bachelor of Computer Science

Submitted By:

Farid Majid Darmajati

1811102441108

DEPARTMENT OF INFORMATICS

FACULTY OF SCIENCE AND TECHNOLOGY

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

SAMARINDA

2022

APPROVAL SHEET

Application of Data Mining For Student Satisfaction Analysis of Offline Learning Performance in The Department of Public Health With The C4.5 Algorithm Method

COMPLETED BY :

Farid Majid Darmajati

1811102441108

Have done and passed the thesis exam,
On 29 June 2022

Supervisor

Wawan Joko Pranoto, S.Kom, M.Ti
NIDN : 1102057701

Examiner

Rofilde Hasudungan, S.Kom, M.Sc
NIDN : 1107048601

Dean

Prof. Ir. Sarjito, MT., Ph.D.
NIDN : 0610116204

Head of Department

Asslia Johar Latipah, M.Cs
NIDN : 1124098902

STATEMENT OF THESIS AUTHENTICITY PAGE

The undersigned below,

Student Name : Farid Majid Darmajati

NIM : 181110244108

Concentration : Machine Learning

Declare that the thesis with the following title:

Application of Data Mining for Student Satisfaction Analysis of Offline Learning Performance in the Department of Public Health With the C4.5 Algorithm Method

Supervisor : Wawan Joko Pranoto, S.Kom, M.Ti

1. This paper is completely ORIGINAL and has NEVER been submitted for a Bachelor's Degree in Informatics, either at the Universitas Muhammadiyah Kalimantan Timur (UMKT) or at any other tertiary institution.
2. This paper is MY idea, formulation and, research, without any help from other parties except the direction of the Advisory Lecturer
3. In this paper there are no works or opinions of others, except in writina g that is clearly stated as a reference in the manuscript by mentioning the name of the author and mentioned in the Bibliography of this paper.
4. The software used in this research is entirely my responsibility, not the responsibility of the Universitas Muhammadiyah Kalimantan Timur (UMKT)
5. I make this statement in truth, if in the future there are irregularities and untruths in this statement, then I am willing to accept ACADEMIC SANCTIONS with the revocation of the title that has been obtained, as well as other sanctions in aby norms that apply at Muhammadiyah Higher Education.

Samarinda, 20 June 2022,

That state,

FARID MAJID DARMAJATI

1811102441108

FOREWORD

Alhamdulillah, by giving thanks to Allah SWT, because with His guidance, this thesis with the title " Application Of Data Mining For Student Satisfaction Analysis Of Offline Learning Performance In Department Of Public Health With C4.5 Algorithm Method" can be completed on time, as one of the requirements to complete lecture assignments and to achieve the graduation requirements of the Universitas Muhammadiyah Kalimantan Timur, Samarinda.

The writer is very aware of all the shortcomings in writing this thesis, but with the moral and material assistance from various parties, all the completeness in writing it can be fulfilled. For this reason, on this occasion, the author should express his gratitude and highest appreciation to various parties who directly or indirectly have provided assistance.

1. Mr. Prof.Dr.Bambang Setiaji as the rector of Universitas Muhammadiyah Kalimantan Timur, and Prof. Ir. Sarjito, MT, Ph.D as the Dean of the Faculty of Science and Technology, Universitas Muhammadiyah Kalimantan Timur, who has been pleased to provide the author with the opportunity to complete his studies at the Universitas Muhammadiyah Kalimantan Timur.
2. Mrs. Asslia Johar Latipah, S. Kom., M.Cs as the head of the Informatics Engineering Department of.
3. Mr. Wawan Joko Pranoto, S.Kom., M.Ti as the supervisor who has provided guidance, input and suggestions that are very helpful in perfecting this thesis.
4. All lecturers of the Department of Informatics and all lecturers of the Universitas Muhammadiyah Kalimantan Timur who have provided a lot of knowledge to the author so that they can complete thesis writing as one of the requirements for obtaining a computer degree.
5. My beloved parents and brothers who always give love, prayers, advice, and for their extraordinary patience in dealing with every step of the author's life, which is the greatest gift in life.

6. Colleagues in the Department of Informatics, Universitas Muhammadiyah Kalimantan Timur, who have provided a lot of thought and motivation so that the author can complete it.

Samarinda, June 2022

Farid Majid Darmajati

1811102441108

ABSTRACT

Students become one of the most vital assets for a higher education institution, and in this case, students as consumers who enjoy educational services in higher education need to get consideration whether they are satisfied or dissatisfied with the services provided. One of the ways to increase learning motivation is to provide good educational services quality to students. Because if students are satisfied with the campus services they receive, they will be more active in attending lectures, as well as in participating in other student activities. Previously at the UMKT, there had never been a measurement of student satisfaction at the study program level. To determine the level of student satisfaction needs to do analyzed. Analysis can be done with classification techniques using the decision tree algorithm/C4.5. Because the C4.5 Algorithm is widely used, it has major advantages over other algorithms. The advantages of the C4.5 Algorithm are that it can produce a decision tree that is easy to interpret, has an acceptable level of accuracy, and is efficient in handling discrete and numeric type attributes. In this research, a decision tree was formed and gave information that the main variable that affects student satisfaction in the Department of Public Health is "Administration services in the Department are easy and fast". And the result of calculating the accuracy value is 83%.

Keywords: Student satisfaction, C4.5 Algorithm, classification, decision tree.

TABLE OF CONTENTS

TITLE PAGE ..... **i**

APPROVAL SHEET ..... **ii**

FOREWORD ..... **iv**

ABSTRACT ..... **vi**

TABLE OF CONTENTS ..... **vii**

LIST OF FIGURES ..... **ix**

LIST OF TABLES ..... **x**

LIST OF ATTACHMENT ..... **xi**

CHAPTER 1 INTRODUCTION **1**

1.1 Background 1

1.2 Problem Statement 3

1.3 Research Objectives 3

1.4 Research Scope 3

CHAPTER 2 LITERATURE REVIEW **4**

2.1 Student Satisfaction 4

2.2 C4.5 Algorithm 6

2.3 Confusion Matrix 7

CHAPTER 3 RESEARCH METHODS **8**

3.1 Stages of Data Processing 8

3.2 Data collection 8

3.3 Validity and Reliability test 10

3.4 C4.5 Algorithm 11

3.5 Performance Evaluation 12

CHAPTER 4 DISCUSSION **13**

4.1 Data Collection 13

4.2 Validity and Reliability Test 13

4.3 C4.5 Algorithm 15

4.4 Performance Evaluation 24

CHAPTER 5 CONCLUSIONS AND FUTURE WORKS **26**

5.1 Conclusion 26
5.2 Future Work 26
REFERENCES 27
ATTACHMENT 30

LIST OF FIGURES

Figure 3.1 Research stage 8

Figure 4.1 Decision tree Student Satisfaction 23

LIST OF TABLES

Table 2.1 Previous Research 4
Table 2.2 Previous Research (continued) 5
Table 2.3 Confusion Matrix 7
Table 3.1 Questionnaire Questions 9
Table 3.2 Category scale and score 10
Table 3.3 Cronbach alpha distribution 11
Table 4.1 Collected data 13
Table 4.2 summary of validity test 14
Table 4.3 summary of reliability test 15
Table 4.4 Confusing matrix data testing 24
Table 4.5 Accuracy test results 25

LIST OF ATTACHMENT

Attachment 1 Data of Respondent 29
Attachment 2 r Table 30
Attachment 3 Validity Test 31
Attachment 4 Reliability Test 32
Attachment 5 Dataset 33
Attachment 6 Entropy & gain test of Reliability 34
Attachment 7 Entropy & gain test of Responsive 34
Attachment 8 Entropy & gain test of Assurance 35
Attachment 9 Entropy & gain test of Empathy 36
Attachment 10 Entropy & gain test of Tangible 36
Attachment 11 programming Script 37
Attachment 12 Percentage graph for each question 39