

**UNIVERSITI TEKNOLOGI MARA**

**PREDICTION OF UITM STUDENT  
ACADEMIC PERFORMANCE USING  
NAIVE BAYES ALGORITHM**

**MUHAMMAD IRFAN ZAHIN BIN JAILANI**

**Thesis submitted in fulfillment of the requirements for  
Bachelor of Computer Science (Hons.) College of  
Computing, Informatics and Mathematics**

**January 2024**

## **ACKNOWLEDGEMENT**

Alhamdulillah, praises and thanks to Allah because of his utmost blessings I was able to finish this research within the time duration given. First, my special thanks go to my supervisor, Ts. Dr. Zamri Bin Abu Bakar for his guidance and support throughout this research. Without him I would not have been able to finish this research with success.

I would like to extend my gratitude to Madam Ummu Fatimah, for her support and knowledge in the subject of CSP600 and CSP650 that have greatly assisted me in doing my research. Furthermore, special appreciation also goes to my beloved parents for the unending support they gave to me physically and mentally as well as encouragement and support for my financial situation.

Last but not least, I would like to give gratitude to my dearest friend and everyone who has been a part of this research and has given me with their unwavering support and encouragement. Without them I would not be able to finish this research.

## **ABSTRACT**

This paper proposes a predictive system using the Naive Bayes algorithm to solve the pressing issue of low academic performance among students at Universiti Teknologi MARA (UiTM). The study underlines a variety of factors affecting student outcomes while highlighting the significance of effective academic performance. With the help of customized interventions and early identification of at-risk pupils, the proposed approach seeks to increase graduation rates and overall achievement. The main objectives of this study include studying the Naive Bayes algorithm in student academic performance prediction, designing and developing a student academic performance prediction model utilizing Naive Bayes, and evaluating the accuracy of the prediction prototype using the developed model. The approach, which consists of splitting the training and testing datasets, preparing the data, and applying Naive Bayes, produces remarkable results: 94.85% accuracy, 91.52% precision on average, 96.84% recall, and a 94.08% overall F1 score. As a result, the suggested system promotes improved student achievement and the welfare of society by providing a proactive way of addressing academic difficulties.

## TABLE OF CONTENTS

<b>CONTENT</b>	<b>PAGE</b>
<b>SUPERVISOR APPROVAL</b>	<b>I</b>
<b>STUDENT DECLARATION</b>	<b>II</b>
<b>ACKNOWLEDGEMENT</b>	<b>III</b>
<b>ABSTRACT</b>	<b>IV</b>
<b>TABLE OF CONTENTS</b>	<b>V</b>
<b>LIST OF FIGURES</b>	<b>VIII</b>
<b>LIST OF TABLE</b>	<b>X</b>
<b>LIST OF ABBREVIATION</b>	<b>XI</b>
<b>CHAPTER 1: INTRODUCTION</b>	
1.1 Background of Study	1
1.2 Problem Statement	2
1.3 Objective	4
1.4 Project Scope	4
1.5 Project Significance	5
1.6 Overview of Research Framework	6
1.7 Conclusion	7
<b>CHAPTER 2: LITERATURE REVIEW</b>	
2.1 Artificial Intelligence	8
2.2 Student Academic Performance	9
2.3 Naïve Bayes Algorithm	10
2.3.1 Naïve Bayes and How Does It Work?	11
2.3.2 Advantages of Naïve Bayes Algorithm	14
2.3.3 System development model	15

2.4 Implementation Naïve Bayes Algorithm in Various Problem	17
2.5 Similar works	21
2.6 The Implication of Literature Review	29
2.7 Conclusion	30

### **CHAPTER 3: RESEARCH METHODOLOGY**

3.1. Overview of Research Methodology	31
3.1.1 Detail Research Framework	31
3.2. Preliminary Phase	36
3.2.1. Literature Study	36
3.2.2. Data Pre-Processing	36
3.2.2.1. Data Collection	37
3.2.2.2 Data Cleaning	38
3.3 Design Phases	39
3.3.1. System Architecture	40
3.3.2 Flowchart	41
3.3.3 Naive bayes algorithm	42
3.3.4 User Interface Design	45
3.4 Development	46
3.5 Evaluation Phases	47
3.5.1 Performance Evaluation	47
3.6 Gantt Chart	49
3.7 Conclusion	51

### **CHAPTER 4: RESULT AND FINDINGS**

4.1 Conceptual Framework	52
4.2 Data preprocessing	54
4.2.1 Data cleaning	54
4.2.2 Feature selection	56
4.2.3 Data transformation	57
4.3 Naïve Bayes implementation	58
4.3.1 Data partitioning	58
4.3.2 Construct Naïve Bayes algorithm.	59
4.4 Prototype Interfaces	61
4.5 Functionality Test	65