

**UNIVERSITI TEKNOLOGI MARA**

**DETERMINE YOUTH SQUAD  
FOOTBALL POSITION FOR NEGERI  
SEMBILAN FOOTBALL  
ASSOCIATION (NSFA) USING  
DECISION TREE TECHNIQUE**

**MOHD AMINNUR HAKIM BIN SABTU**

**BACHELOR OF INFORMATION  
TECHNOLOGY (HONS.) INFORMATION  
SYSTEMS ENGINEERING**

**JANUARY 2017**

## **STUDENT'S DECLARATION**

I certify that this report and the project to which it refers is the product of my own work and that any idea or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.

.....  
MOHD AMINNUR HAKIM BIN SABTU  
2014382959

JANUARY, 2017

## ABSTRACT

The main purpose of this research is to produce a prototype that automatic resulting the suitable football position for the under 19 (U19) players of Negeri Sembilan Football Association (NSFA). This research only focus on basic position in football such as goalkeeper, center back, right and left back, midfielder, right and left winger, attacking midfielder and striker. There are 32 football characteristic that can lead in resulting these football position. By developing this prototype, coach of NSFA can get more confirmation for U19 position based on the player's characteristic. There are lot of rules embedded to this prototype. This set of rules was generated by creating decision tree (DT). DT gives a clear view the rules involved. This project used waterfall model for the guideline for development. There are four (4) phases of methodology involved such as knowledge acquisition, requirement gathering and analysis, design and implementation. In each phase brings own deliverable and objective. All objectives have been achieved. For this project, the first objective started from phase two (2) which is requirement gathering and analysis. This phase contains two (2) objectives. The first one was to gather and analyze common characteristics for each football player and to classify football position characteristic. For the third phase objective in this research is to design the system for NSFA and last objective for the last phase is to design the system for NSFA. This research contains a lot of challenges. This research needs a lot of knowledge in data mining especially for DT classification technique. Other than that, there was a lot of interview conducted in order to get the data from the stakeholder and to understand the business process. This prototype has been develop using JAVA as programming language and oracle as a database for store the data. At the end of this project, Football Position Decision Maker (FPDM) was able to generate recommendation of football position for U19 football squad. This result ease coach in order to get the confirmation of U19 football position. This project needs some enhancement, one of the potential enhancement that can be seen to be developed is increase the system security. This system used email for send verification to the coach. In a future enhancement, this verification can be encrypt and only can be decrypt only by current coach of NSFA in nearly future.

# TABLE OF CONTENT

<b>CONTENTS</b>	<b>PAGE</b>
<b>SUPERVISOR'S APPROVAL</b>	<b>ii</b>
<b>STUDENT'S DECLARATION</b>	<b>iii</b>
<b>ACKNOWLEDGEMENT</b>	<b>iv</b>
<b>ABSTRACT</b>	<b>v</b>
<b>TABLE OF CONTENT</b>	<b>vi</b>
<b>LIST OF FIGURES</b>	<b>xi</b>
<b>LIST OF TABLES</b>	<b>xiii</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xiii</b>
<b>CHAPTER ONE: INTRODUCTION</b>	<b>1</b>
1.1 Background of Study	1
1.2 Problem Statement	3
1.3 Aim	4
1.4 Objectives	4
1.5 Scope	4
1.6 Research Significance	4
1.7 Research Outline of the Thesis	5
<b>CHAPTER TWO: LITERATURE REVIEW</b>	<b>6</b>
2.1 Overview of the Football Position	6
2.1.1 Goalkeeper	7
2.1.2 Defender	7

2.1.2.1	Center Back	7
2.1.2.2	Wing Back	8
2.1.3	Midfielder	8
2.1.3.1	Midfielder	8
2.1.3.2	Attacking Midfielder	9
2.1.4	Forward	9
2.1.4.1	Winger	9
2.1.4.2	Striker	9
2.2	Data Mining	10
2.2.1	Unsupervised Learning	11
2.2.2	Supervised Learning	11
2.3	Overview of Classification Models	12
2.3.1	Classification by Decision Tree Induction	12
2.3.1.1	Advantages and Disadvantages of Classification by Decision Tree Induction	13
2.3.1.2	Gini Index	13
2.3.1.3	Gain Ratio	14
2.3.1.4	Information Gain	15
2.3.1.5	Comparison between Gini and Gain Ratio	16
2.3.2	Bayesian Classification	17
2.3.3	Neural Network	18
2.3.4	Support Machine Learning	18
2.4	Comparison of Classification Models	19
2.5	Methodology	21
2.5.1	Waterfall Model	21
2.5.1.1	The Advantages and Disadvantages of Waterfall Model	22
2.5.2	Agile Model	23