

UNIVERSITI TEKNOLOGI MARA

**EARLY AUTISM SPECTRUM
DISORDER DETECTION USING
MACHINE LEARNING**

MUHAMMAD SHAHMI BIN SHAHRON NIZAN

2020898828

BACHELOR OF COMPUTER SCIENCE (Hons.)

AUGUST 2023

ACKNOWLEDGEMENT

Alhamdulillah, praises and thanks to Allah because of His Almighty and His utmost blessings, I was able to finish this research within the time duration given. Firstly, my special thanks go to my supervisor, Puan Nor Azylia Binti Ahmad Azam, for her utmost support and guidance to ensure everything is completed amazingly with high quality work in this thesis. I have learned a lot from her, having work with her throughout my journey, and the single most crucial matter is that to always understand the task at hand and be detailed in each aspect to not leave any valuable information behind as it is crucial in the thesis. I am thankful for her patience in helping me with corrections that I must undergo at times.

Next, the second person that follows suit is none other than my lecturer Dr Raihah Binti Aminuddin for her dedication in teaching the subject of CSP600 to ensure students will understand and help with the production of this thesis. Her spirit of dedication inspires me to complete this thesis with determination and diligence.

The best should be reserved for last. My family is the source of my motivation and strength, and they deserve more than just a modest acknowledgement like this. I would not have survived while doing this thesis without their efforts, no matter what situation they were in. They have done nothing but show unending support to me, and I will always find a way to thank them through my future achievements. In this thesis, I also want to thank my friends for pushing me to the brink of success and never giving up on me, even when I was fatigued. Thank you very much for your help and consideration.

ABSTRACT

This project aims to develop a web-based application utilizing the Random Forest Classification Algorithm to aid concerned parents in detecting potential Autism Spectrum Disorder (ASD) symptoms in their children aged 1-6 years in Malaysia. The application considers various factors, including children's gender, parental income range, and responses to the Modified Checklist for Autism in Toddlers (M-CHAT) questionnaire, to provide risk categorization and recommend nearby support facilities. By offering an online platform, the project addresses the increasing prevalence of ASD and helps parents seek professional support for their children. It also assists parents in preparing their ASD-affected children for primary school by suggesting appropriate assistance options like Program Pemulihan Dalam Komuniti (PDK), Program Pendidikan Khas Integrasi (PPKI), and Program Pendidikan Inklusif (PPI). The project follows a modified waterfall approach, focusing on creating a user-friendly interface and integrating the Random Forest Classification Algorithm for accurate detection. The results show the algorithm's impressive performance with an 86% precision in predicting ASD traits. In conclusion, this web-based application provides a reliable and accessible tool for early ASD detection, empowering parents to assess their children's risk and seek appropriate support. However, the project acknowledges limitations such as a small dataset and subjective questionnaire-based assessments, calling for further attention. Future work involves data expansion techniques, integrating objective measures alongside questionnaires, and collaborating with relevant organizations to enhance the system's capabilities and effectiveness in detecting ASD in Malaysian children.

TABLE OF CONTENT

CONTENT	PAGE
SUPERVISOR APPROVAL	ii
STUDENT DECLARATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENT	vi
LIST OF FIGURES	x
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xv

CHAPTER ONE: INTRODUCTION

1.1	Background of Study	1
1.2	Problem Statement	2
1.2.1	The Rising Case of Autism Spectrum Disorder (ASD)	2
1.2.2	Challenges Faced by Parents in Malaysia	2
1.2.3	Current ASD Detection System Problems	3
1.3	Project Objectives	3
1.4	Project Scopes	4
1.5	Project Significance	4

CHAPTER TWO: LITERATURE REVIEW

2.1	Overview of Autism Spectrum Disorder	7
2.1.1	Autism	7
2.1.2	Main Causes of Autism	8
2.1.3	Main Types of Autism	8

2.2	Overview of Big Data in Healthcare	10
2.2.1	Big Data	10
2.2.2	Healthcare	12
2.2.3	Big Data in Healthcare	12
2.3	Data Mining Techniques Used in Prediction	13
2.3.1	Data Mining	13
2.3.2	Data Mining Techniques	13
2.4	Algorithm Used in Prediction Model	18
2.4.1	Prediction Model in Healthcare	18
2.4.2	Algorithm in Prediction Models	19
2.4.3	Random Forest Algorithm	19
2.4.4	K-Nearest Neighbours Algorithm	21
2.4.5	Logistic Regression Algorithm	21
2.5	Comparative Analysis on Prediction Model Algorithm	22
2.6	Related Work	23
2.6.1	Evaluation on Machine Learning Algorithms for Classification of Autism Spectrum Disorder (ASD)	23
2.6.2	A Machine Learning Approach to Predict Autism Spectrum Disorder	27
2.6.3	A Machine Learning Strategy for Autism Screening in Toddlers	28
2.7	Feature Analysis of Related Work	29
2.8	Conclusion	30

CHAPTER THREE: METHODOLOGY

3.1	Introduction	32
3.2	Modified Waterfall Model	33
3.3	Requirement Phase	34
3.4	System Design Phase	34
3.5	Implementation Phase	35
3.5.1	Data Collection	36