



MODELLING STUDENT'S ACADEMIC PERFORMANCE DURING COVID-19: A CASE STUDY BASED ON CLASSIFICATION AND REGRESSION IN SUPPORT VECTOR MACHINE



NOR AIN MAISARAH BINTI SAMSUDIN

UNIVERSITI PENDIDIKAN SULTAN IDRIS

2021





MODELLING STUDENT'S ACADEMIC PERFORMANCE DURING COVID-19: A CASE STUDY BASED ON CLASSIFICATION AND REGRESSION IN SUPPORT VECTOR MACHINE

NOR AIN MAISARAH BINTI SAMSUDIN



**THESIS REPORT IS SUBMITTED TO MEET THE REQUIREMENTS FOR
OBTAINING A BACHELOR OF SCIENCE (MATHEMATICS)
WITH EDUCATION**

**FACULTY OF SCIENCE AND MATHEMATICS
UNIVERSITI PENDIDIKAN SULTAN IDRIS**

2021





DECLARATION

“I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.”

Signature :

Full Name :

Nor Ain Maisarah Binti Samsudin

Matric No. :

D20172080938

Date :

21st February 2021





ACKNOWLEDGEMENT

All praise be to Allah S.W.T because with His bounty and grace, I have successfully completed this study. I would like to express my sincere gratitude to my supervisor, Dr. Shazlyn Milleana Binti Shaharudin, for her motivation, guidance, and support throughout the research. She continuously steered me in the right direction in this final year project.

I would like to thank to the Department of Mathematics from Faculty of Science and Mathematics, UPSI for giving me permission to conduct the research on students from the Department of Mathematics, UPSI. I am really grateful for all the support from Nurul Ainina Filza Binti Sulaiman who was always available to me whenever I needed his guidance other than my supervisor.

My deep sense of gratitude to my mother, Selamah Binti Salleh and my sister Nor Emy Suhada Binti Samsudin who are my moral strength and motivation. I would like to thank Noor Aziera Binti Musa, Muhammad Fakhrollah Bin Mohd Fuad and Muhammad Fareezuan Bin Zulfikri for their constant support and guidance throughout my final year project.

Finally, I would like to thank all my friends, seniors and juniors who made my time here at UPSI very memorable.





MODELLING STUDENT'S ACADEMIC PERFORMANCE DURING COVID-19: A CASE STUDY BASED ON CLASSIFICATION AND REGRESSION IN SUPPORT VECTOR MACHINE

ABSTRACT

This study proposed a statistical investigate the pattern of students' academic performance before and after online learning due to the Movement Control Order (MCO) during pandemic outbreak and a modelling students' academic performance based on classification and regression in Support Vector Machine (SVM). Data sample consist of 228 students were taken from undergraduate students of Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris (UPSI). The instrument used for data collection is a questionnaire. Student's Grade Point Average (GPA) before and after online learning were taken to identify the pattern of students' academic performance before and after Movement Control Order (MCO) while Cumulative Grade Point Average (CGPA) after online learning were obtained to developed model of academic performances during COVID-19 outbreak. The algorithm of Support Vector Machine (SVM) classification and regression was used to develop a model of students' academic performance. For the Support Vector Machine (SVM) classification algorithm, there are two important parameters which are C (misclassification tolerance parameter) and epsilon while for the Support Vector Machine (SVM) regression also need to identify two important parameter which are penalty term and epsilon. All the parameters need to identify before proceed the further analysis. The parameters was applied to four different types of kernel which is linear kernel, radial basis function kernel, polynomial kernel and sigmoid kernel for both model. The result was found that the best accuracy achieved by SVM classification are 73.69% by using radial basis kernel with parameter of misclassification tolerance C is 128 and epsilon is 0.6. Meanwhile, SVM regression model also resulted radial basis function as the best kernel with parameter of penalty term is 4 and epsilon 0.8 with the 0.255659 as the lowest value of RMSE. The pattern of prediction of students' academic performance are similar with the current CGPA. Therefore, Support Vector Machine regression are able to predict students' academic performance.



TABLE OF CONTENTS

	PAGES
DECLARATION	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
TABLE OF CONTENT	v
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi

CHAPTER 1 INTRODUCTION

1.0	Introduction	1
1.1	Preface	2
1.2	Background Study	3
1.3	Problem Statement	4
	1.3.1 Challenges with Online Learning	4
	1.3.2 Online Leaning Affects Student’s CGPA	5
	1.3.3 Studies Related To Online Learning During Pandemic	5
1.4	Conceptual Framework	6
1.5	Purpose of Study	7
1.6	Objective of Study	7

1.7	Research Question	7
1.8	Significant of Study	8
1.9	Limitation of Study	8
1.10	Operational Definition	9
1.10.1	Coronavirus Disease (COVID-19)	9
1.10.2	Movement Control Order (MCO)	9
1.10.3	Online Learning	9
1.10.4	Grade Point Average (CGPA)	10
1.10.5	Cumulative Grade Point Average (CGPA)	10
1.11	Conclusion	11

CHAPTER 2 LITERATURE REVIEW

2.1	Introduction	12
2.2	Student Academic Performance	13
2.3	Online Learning Affecting Students' Academic Performance	13
2.4	Predictive Model	14
2.4.1	Support Vector Machine (SVM)	15
2.4.1.1	SVM Classification	16
2.4.1.2	SVM Regression	17
2.5	Conclusion	17

CHAPTER 3 METHODOLOGY RESEARCH

3.1	Introduction	19
3.2	Design of Study	20
3.2.1	Location and Duration of the Study	20
3.2.2	Population and Samples	20

3.2.2.1	Populations	20
3.2.2.2	Samples	21
3.3	Research Instruments	21
3.4	Data Collection	22
3.5	Data Analysis	22
3.5.1	Support Vector Machine (SVM)	22
3.5.1.1	SVM Classification	24
3.5.1.2	SVM Regression	26
3.7	Conclusion	28

CHAPTER 4 RESULTS AND DISCUSSION

4.0	Introduction	29
4.1	Pattern of Students' Academic Performance	30
4.2	Support Vector Machine (Classification)	31
4.2.1	Turning Parameter in SVM Classification	31
4.2.2	Result of SVM by Classification	34
4.3	Support Vector Machine (Regression)	35
4.3.1	Turning Parameter in SVM Regression	35
4.3.2	Type of Kernel and Value of RMSE	38
4.3.2.1	Linear	38
4.3.2.2	Radial Basis Function (RBF)	39
4.3.2.3	Polynomial	39
4.3.2.4	Sigmoid	40
4.3.3	Result of SVM by Regression	41
4.4	Prediction of Students' Academic Performance by Using SVM Regression	42

4.5	Conclusion	43
-----	------------	----

CHAPTER 5 IMPLICATION, LIMITATION AND FUTURE WORKS

5.0	Introduction	45
5.1	Implication of the Study	45
5.2	Limitation of the Study	45
5.3	Future Work	46
5.4	Conclusion	46
5.5	Summary	47

REFERENCES	48
-------------------	----

ATTACHMENT



LIST OF TABLES

No. of Tables		Pages
4.1	The Result of Turning Parameter Process in SVM Classification.	30
4.2	Result of Support Vector Machine by Classification	33
4.3	The Result of Turning Parameter Process in SVM Regression	35
4.4	Parameters, Number of Support Vectors and RMSE Value for Linear Kernel	38
4.5	Parameters, Number of Support Vectors and RMSE Value for RBF Kernel	39
4.6	Parameters, Number of Support Vectors and RMSE Value for Polynomial Kernel	40
4.7	Parameters, Number of Support Vectors and RMSE Value for Sigmoid Kernel	40
4.8	Result of Support Vector Machine by Regression	41





LIST OF FIGURES

No. of Figures		Pages
1.1	Conceptual Framework of the Study	6
2.1	Support Vector Machine (SVM)	16
3.1	Support Vector Machine (SVM): Separable Classes	23
3.2	Support Vector Machine Classification	24
3.3	Support Vector Machine Regression	26
4.1	GPA of Academic Performance Students Before and During COVID-19 Outbreak	31
4.2	Performance of Turning Parameter in SVM Classification	34
4.3	Performance of Turning Parameter in SVM Regression	37
4.4	Current and Forecasted CGPA during Online Learning	42





LIST OF ABBREVIATIONS

CGPA	Cumulative Grade Point Average
COVID-19	Coronavirus disease
FSM	Faculty of Science and Mathematics
GPA	Grade Point Average
MCO	Movement Control Order
MOH	Ministry of Health
MOHE	Ministry of Higher Education
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SVM	Support Vector Machine
WHO	World Health Organization
UPSI	<i>Universiti Pendidikan Sultan Idris</i>





ATTACHMENT

A Questionnaire





CHAPTER 1

INTRODUCTION



1.0 Introduction

This chapter will discuss the introduction, background of the study, problem statement, conceptual framework, purpose of the study, objectives of study, research questions, significant of the study to specific parties, limitations of the study and operational definitions that need to be clearly understood. Therefore, researcher can make these things as guidelines to determine the research method that will be used in this study.





1.1 Preface

On December 31, 2019, the world was shocked by a novel coronavirus originating from Wuhan province, China. The outbreak is a new type of Coronavirus (COVID-19) which has the potential to cause serious illness and death (Berita Harian, 2020). The World Health Organization (WHO) declared a 'global health emergency' in January 2020 when the COVID-19 epidemic claimed 170 lives in a short period of time (WHO, 2020). In March 2020, the World Health Organization (WHO) declared COVID-19 as a pandemic in which the epidemic spread across the region and involved large populations into one world (WHO, 2020).

Malaysia is also not immune from the COVID-19 epidemic. The positive rate of COVID-19 Malaysian patients initially in a controlled rate declined where the number of patients increased with the arrival of foreigners to Malaysia (MOH, 2020). On 16 March 2020, the Prime Minister, Tan Sri Muhyiddin Yassin declared the Movement Control Order (MCO) effective from 18 March 2020 until 31 March 2020 and continued for five months and 13 days in an effort to curb the spread of COVID-19 (Berita Harian, 2020).

The effect of this MCO has indirectly caused schools and institutions of higher learning institutions not to be allowed to run during the MCO period. On 27 May 2020, Ministry of Higher Education (MOHE) informed that all Teaching and Learning activities for students must be implemented online until 31 December 2020 (MOHE, 2020). Online learning is not stranger among university students. However, conducting





the concept of digital lectures as a whole is seen to have a huge impact on university students, especially on academic achievement.

1.2 Background Study

University students need to develop the basics of self that is knowledgeable, skilled, creative and have a strong belief in work ethic and great spiritual values so that they can compete on the global stage (Senian Malie & Oariah Akir, 2013). One of the most important aspects for students is to maintain a high level so that they can ensure that they have a brighter future after graduation and start engaging themselves in the world of work.



Since Prime Minister Tan Sri Muhyiddin Yassin has announced MCO and closure of educational institution, Universiti Pendidikan Sultan Idris was also included to restrict students and lecturers from attending classes. Because of that, students and lecturers need to continue their classes through online platforms.

This research mainly focused on the prediction of academic performance of Universiti Pendidikan Sultan Idris (UPSI) undergraduates after they completely attend whole one semester by studying online. Moreover, this research also determines whether online classes will be a bad or good platform for students to study and gain knowledge.





1.3 Problem Statement

All universities or private colleges have implemented and changed the concept of online learning and teaching since the Movement Control Order (MCO) was implemented. However, this change towards virtual learning is indeed a big challenge not only to the students, but also to impact the teaching methods by the lecturers.

1.3.1 Challenges with Online Learning

Online learning is a counter to face-to-face meetings. Online learning is no stranger to the University. Not to mention through the philosophy of cybergogy that is so close to today's generation apart from heutagogy and paragogy (Afandi Ahmad, 2020). Facing a change in learning style like this is not easy. Students must inform the lecturer of their Internet access capabilities. This is very important to ensure that lecturers do not beat in implementing online learning. In the meantime, with limited mobility, students need to make full use of online resources, such as electronic books, YouTube and even scientific calculators. Taking into account the challenges of whether students are still on campus or have been at home, they must be more disciplined, understandably, do not have to chase to the lecture hall, but only need to sit in front of a computer or with a mobile phone to study.





1.3.2 Online Learning Affects CGPA.

Online classes will also pose a problem in terms of assessment. While many tasks can be completed at home, exams are a fairly important component of assessment. Tests are usually done in a room or hall, and it is impossible to gather hundreds of students to sit for their exams as long as there is a risk of COVID-19. The alternative is to hold an online exam. However, the risk of cheating is high especially when they cannot be directly monitored. Ej jasper et al. (2012) stated that there is insufficient evidence to support the claim that the time spent on e-learning is related to CGPA. Although some people suspect that longer time spent on online learning will lead to better academic performance



1.3.3 Studies Related To Online Learning During Pandemic.

Many studies have been conducted to study online learning, but it is very difficult to find studies related to online learning during pandemics (Gonzalez C., 2012). Past studies on online learning have been implemented whether the impact on student achievement, student learning methods or the students' own perceptions of online learning. However, it is quite difficult to find pandemic related studies on students' academic performance.



1.4 Conceptual Framework

In order to find the pattern of students' academic performance before and after online learning and the proposed modelling students' academic performance was shown in Figure 1.1.

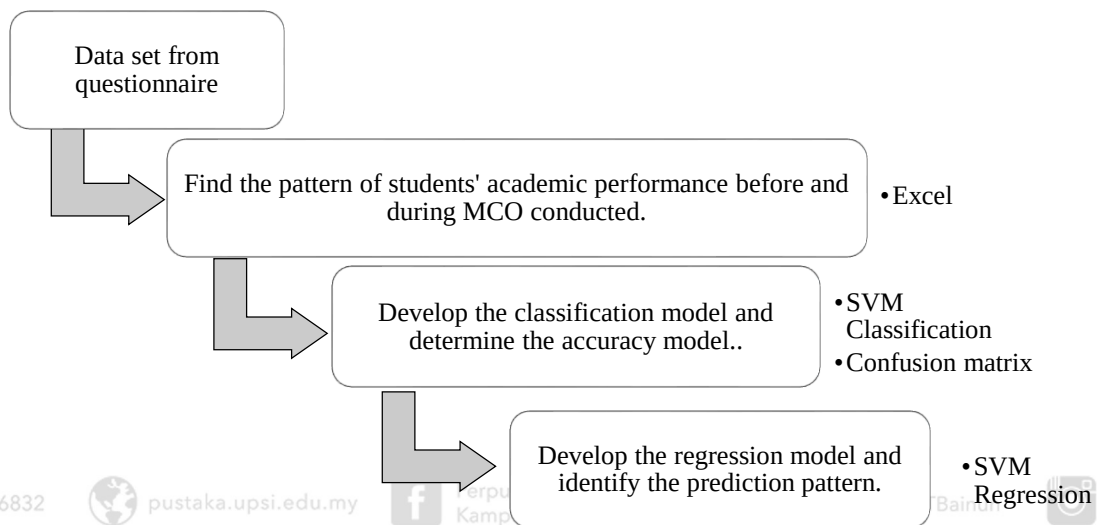


Figure 1.1. Conceptual Framework of the Study.

This study begins with the collection of data through a questionnaire. The data of students' GPA before and after online learning were used to find the pattern of students' academic performance before and after MCO by using Microsoft Excel.

Then the study continued to develop the classification model by using Support Vector Machine (classification). Confusion matrix was used to identify the accuracy model. Then, Support Vector Machine (regression) were applied to develop regression model and identify the prediction pattern of students' academic performance.



1.5 Purpose of Study

This study is a quantitative study through questionnaire data. The main purpose of this study is to make a prediction of students' academic performance during COVID-19 using the predictive model which is Support Vector Machine (SVM). These study also involving the correlation between online learning and the academic performance of university students.

1.6 Objective of Study

This study was conducted to:

1. Find the pattern of students' academic performance before and during online learning due to the MCO.
2. Analyse the prediction of students' academic performance throughout COVID-19 based on classification and regression in Support Vector Machine.

1.7 Research of Questions

In particular, the scope of the study leads to the following questions;

1. How is the academic performance of UPSI students before and after undergoing online learning for one semester during COVID-19?
2. What is the prediction of students' academic performance after undergoing online learning?





1.8 Significant of the Study

At the end of this study, it is hoped that it can produce a prediction of student academic performance using a predictive model, namely Support Vector Machine (SVM) which brings benefits to university students, lecturers and researchers in Malaysia. In addition, this study can help universities' students, academic community and other parties who are interested to improve the academic performance of students on online learning studies. Researchers will also be able to review learning attitudes and environmental factors that affect students' academic performance and the correlation between online learning and student academic performance during COVID-19.



1.9 Limitation of Study



In this study, there are several limitations of the study identified namely:

- i. This study is limited to UPSI students which is only focused on undergraduate students from Faculty of Science Mathematic.
- ii. The scope of the prediction for students' academic performance is only focused on student performance before and during COVID-19.
- iii. Limited to students who have internet connection only to answer the questionnaire given.





1.10 Operational Definition

1.10.1 Coronavirus Disease (COVID-19)

From World Health Organization (WHO), coronaviruses are a kind of infection. There are a wide range of sorts, and some reason malady. A recently recognized coronavirus, SARS-CoV-2, has caused an overall pandemic of respiratory sickness, called COVID-19. Coronavirus disease (COVID-19) is an irresistible ailment brought about by a newfound coronavirus. (WHO, 2020).

1.10.2 Movement Control Order (MCO)



The Malaysian Movement Control Order (MCO) 2020 refers to preventive measures by the Malaysian federal government against the case of the coronavirus pandemic (COVID-19) on 16 March 2020. (Hadei et al., 2020; Hsiao et al., 2020). Therefore "movement control" is implemented nationwide. The incident is often referred to by some media, both local and international media as "lockdown" or "partial lockdown".

1.10.3 Online Learning

Online Learning (e - learning) is any form of teaching and learning delivered through the use of digital technology (Wan Aziaris Aziz, 2015). Teaching and learning materials presented using this media have visual graphics, words, animation, video or





audio. In addition, it also provides facilities such as facilities in the form of group learning and assisted by instructors in certain fields online.

1.10.4 Grade Point Average (GPA)

Grade Point Average (GPA) is a measurement of academic performance (Galiher, 2006; Darling, 2005; Broh, 2002; Stephens & Schaben, 2002). Grade Point Average (GPA) is marks and grades for each semester. It is calculated starting from the marks for each subject taken for a semester. Scores from each of these subjects will be multiplied and divided by the amount of credit taken.



1.10.5 Cumulative Grade Point Average (CGPA)

Cumulative Grade Point Average (CGPA) calculates the overall average of students' academic performance for all examinations' grade for all semesters during the tenure years in the college or university (Nor Adibah Abu Hasan; Nurhafizah Ahmad; Noor 'Aina Abdul Razak, 2017). CGPA is the overall score for all semesters that have passed. Number of pointers from semester one and will be mixed all semesters up to last semester.





1.11 Conclusion

Prediction of students' academic performance throughout the MCO using the predictive model, SVM is expected to achieve the set objectives. Various types of students' academic performance predictions have been implemented and it is hoped that this predictions using the SVM model can make an important contribution in the process of improving learning in Malaysia especially online learning during pandemic. As this study is limited to groups of students from the Faculty of Science and Mathematics, UPSI, the findings of the study cannot represent the whole Malaysia universities students.

