Universiti Teknologi MARA

Predicting Breast Cancer Using Ant Colony Optimisation

Siti Sarah Aqilah Binti Che Ani

Report submitted in fulfillment of the requirements for
Bachelor Science (Hons.) Management Mathematics
Faculty of Computer and Mathematical Sciences

January 2021

STUDENT'S DECLARATION

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

……

SITI SARAH AQILAH BINTI CHE ANI

2019328783

JANUARY 26, 2021

ABSTRACT

Breast cancer is one of the common reasons of death for women in every country, including Malaysia, and the number of breast cancer cases is rising every year worldwide. This type of cancer happens when there exist lumps or additional tissue mass in the breast which are also known as tumors. The cancer cells can be dangerous if it is a benign cell but not a cancerous cell if it is a malignant cell. In order to classify the breast cancer cells, an accurate and effective classification model is needed. Hence, the main purpose of this study is to develop a model of classification for breast cancer cell prediction. This study implements a machine learning algorithm called Ant Colony Optimization (ACO) algorithm to develop an accurate classification model for predicting breast cancer cells. In this study, the ACO algorithm will be compared with another machine learning algorithm - the J48 algorithm - to compare which brings more precise and effective results. Besides, in this study, the Ant-Miner system also plays an important role, which can train the data several times to achieve the highest percentage of predictive accuracy and make comparisons to achieve a good classification model. The results of this study have shown that a classification model using ACO is as comprehensible and accurate as one using the J48 algorithm due to the predictive accuracy produced. Since the lowest rules number and condition numbers is produced by the Ant Colony Optimization, this study concludes that it is the most suitable algorithm in developing the classification model for predicting breast cancer cells.

TABLE OF CONTENTS

CONTENTS		PAGE
SUPERVISOR'S APPROVAL		ii
DECLARATION		iii
ACKNOWLEDGEMENT		iv
ABSTRACT		v
TABLE OF CONTENTS		vi
LIST OF FIGURES		viii
LIST OF TABLES		ix
CHAPTER ONE: INTRODUCTION		
1.1	Background of the Study	1
1.2	Problem Statement	3
1.3	Objective of the Study	4
1.4	Scope of the Study	4
1.5	Significance of the Study	4
1.6	Summary	5
CHAPTER TWO: LITERATURE REVIEW		
2.1	Breast Cancer	6
2.2	The Cell Nuclei in a Digitized Image of a Fine Needle Aspirate of a Breast Mass	7
2.3	Ant Colony Optimization	8
2.4	Ant Colony Optimization for Rule Induction	8
2.5	Summary	10
CHAPTER THREE: RESEARCH METHODOLOGY		
3.1	Research Framework	11

3.2 Data Pre-processing 11
3.3 Model Development 13
3.4 Model Validation 19

CHAPTER FOUR: RESULTS AND DISCUSSIONS

4.1 Results of Classification using Ant-Miner 20
4.2 Results of Classification using J48 24

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions 26
5.2 Recommendations 26

REFERENCES **27**

APPENDICES

APPENDIX A: ANT-MINER RESULT WITH NUMBER OF ANTS=2000 31

APPENDIX B: J48 RESULT 42