ACKNOWLEDGEMENTS

I would like to take this opportunity to express my gratitude to those who helped me with various aspects of conducting research and writing this thesis.

Firstly, I would like to express my deepest gratitude to **Dr. Mie Mie Thet Thwin**, Rector, the University of Computer Studies, Yangon, for her kind permission to develop this thesis.

I am grateful to my thesis supervisor **Dr. Khin Mar Soe**, Professor, NLP Lab, for her suggestions and encouragement to do this thesis and also for her close supervision, invaluable suggestions, kind guidance and constant encouragement during the course of this work.

I also like to express my thanks to our dean **Dr. Thi Thi Soe Nyunt**, Professor and Head of Faculty of Computer Science, for her kind help in this work.

I would like to extend my deepest gratitude to **Daw Ni Ni San**, Lecturer, English Department, for her advice and language editing.

I heartily appreciate the suggestions and recommendations of the teachers who attended all my seminars.

Sincere thanks are also due to my close friends for their kind help, understanding and cooperation throughout the work. Finally, I am most indebted to my beloved family for their support, patience and constant encouragement throughout my studies.

ABSTRACT

Text Classification is the task of automatically assigning a set of documents into certain categories (class or topics) from a predefined set. This also plays an important role in natural language processing and also crossroads between information retrieval and machine learning. The dramatic growth of text document in digital form news website makes the task of text classification more popular over last ten year. The application of this method can be found in spam filtering, question and answering, language identification. This paper presents the idea of text classification process in term of using machine learning technique and illustrates how Myanmar news documents are classified by applying genetic algorithm. The applied system will use Myanmar online news articles from Myanmar news website for the purpose of training and testing the system. Term frequency inverse document frequency (TF-IDF) algorithm was used to select related feature according to their labelled document which is also applied in many text mining methods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS