

**UNIVERSITI TEKNOLOGI MARA**

**MODULAR DEEP NEURL  
NETWORK IN REDUCING  
OVERFITTING TO ENHANCE  
GENERALIZATION**

**MOHD RAZIF BIN SHAMSUDDIN**

Thesis submitted in fulfilment  
of the requirements for the degree of  
**Doctor of Philosophy**  
**(Computer Science)**

**College of Computing, Informatics and Mathematics**

**April 2024**

## ABSTRACT

Machine Learning (ML) and Artificial Intelligence (AI) are a hype in this new age. Some researcher may use it to do classification, recognitions, or even predictions. Although NN have been around for so long since it was introduced. There are many of its renditions and variations that was created for different purposes. Some variation of the network design such as Deep Neural Network (DNN) is very practical and applicable in many different disciplines and field of study with massive amount of data. More and more research has been researched and developed that focuses on how the DNN model that can produce accurate results. Most of those research results varies as it uses different data, different network design, different parameters and optimizing algorithm. This research aims to experiment a new DNN model that functions modularly by looking into several features that will affect the NN training dynamics. This research also aims to discuss several issues and problems that is associated to deep networks such as overfitting, scaling issues and training time. This study is conducted through development of a Modular Deep Neural Network (MDNN) and several experiments to enhance its training capabilities. The experiments are conducted in three phases. In the first phase, the NN is tested with two different data formatting to ensure which data format is more suitable for the next consecutive experiments. It is found out that grayscale format works better as it retains the original information of the inputs and produced better precision based on the highest accuracy of 98.11%. Consecutively at the second phase, DNN models with different optimizers, batch size and activation functions are trained and analysed. From these experiments it has shown that when Adamax optimizer with ReLu activation can produce a promising test accuracy of 96% with just a small batch size. Finally, in the third phase, an experiment is executed to enhance the accuracy of the MDNN model with seeded data. By adding seeded dataset to the training data tremendously changes its accuracy with maximum test accuracy of 94%. This also greatly reduced its training time from 20 hours to just 30 minutes. Finally, a T-Test is conducted to compare several MDNN models to see its consistencies in producing good results.

## ACKNOWLEDGEMENT

First and foremost, I'm deeply grateful to Allah S.W.T for granting me the opportunity to embark on my PhD, for my long and challenging journey, in good health and wellbeing until finally successfully finishing it.

I would like to send my highest gratitude and my humble thanks go to my ever supporting and passionate supervisor Prof. Hajjah Dr. Azlinah Haji Mohamed, and not forgetting my current supervisor, Assoc. Prof. Dr. Shuzlina Abd. Rahman for her patience and perseverance towards helping me in my PhD journey. My deepest respect goes to both of you as you've always had been a beacon of me to always persevere in the pursue of finishing my PhD. I thank you both for the support, patience and ideas in assisting me in this challenging research.

I also would like to express my gratitude to the staff of the Faculty of Computer Science and Mathematics, especially Assoc. Prof. Dr. Norjansalika Janom, Dr. Sofianita Mutalib, and Dr. Azliza Binti Mohd Ali for providing the facilities, knowledge and assistance. My special thanks to the Deans of this college, Program Coordinator and the Head of Computer Science whom are always ready to give support in finishing my studies.

My sincere appreciation goes to my family especially to my loving wife Roslinda Binti Ibrahim. My three loveable kids, Muhammad Zakwan Fahim, Nurul Zulaikha Sofia and Muhammad Zakwan Furqaan for the everlasting inspiration they gave me during my PhD journey. Special thanks to my siblings, my colleagues and friends for the spiritual and moral support for me to finish my research.

Finally, this thesis is dedicated to my father and mother of their prayers, vision and determination to educate me. This piece of victory is dedicated to both of you. Alhamdulillah.

# TABLE OF CONTENTS

	<b>Page</b>
<b>CONFIRMATION BY PANEL OF EXAMINERS</b>	<b>ii</b>
<b>AUTHOR’S DECLARATION</b>	<b>iii</b>
<b>ABSTRACT</b>	<b>iv</b>
<b>ACKNOWLEDGEMENT</b>	<b>v</b>
<b>TABLE OF CONTENTS</b>	<b>vi</b>
<b>LIST OF TABLES</b>	<b>x</b>
<b>LIST OF FIGURES</b>	<b>xii</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xviii</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>19</b>
1.1 Research Background	19
1.2 Problem Statement	23
1.3 Research Questions	28
1.4 Research Objectives	29
1.5 Research Scope	29
1.6 Research Significance	31
1.7 Research Contribution	32
1.8 Overview on The Research Methodology	34
1.9 Overview of the Thesis	35
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>36</b>
2.1 Introduction	36
2.2 Artificial Neural Networks	41
2.3 Brief History of Artificial Neural Networks	42
2.4 Beginning of Deep Learning Era	45
2.5 Neural Networks Architecture	46
2.5.1 Feedforward Backpropagation as Deep Learning Methods	47
2.5.2 BPNN components and algorithm	48
2.5.3 Activation functions in Neural Network	50

# CHAPTER 1

## INTRODUCTION

### 1.1 Research Background

Machine Learning (ML) is quickly being deployed in all sorts of industries, making a large need for skilled professionals. This machine education industry is expected to produce from \$100 million in 2018 to \$935 million in 2025 (Fakhruddin, 2018; McCain, 2022). ML applications are used for information analytics, information mining, and design identification. On the consumer point, ML powers system search results, real-time advertisements, and web intrusion detection, to list just few of the many jobs it can do. In addition to the countless listed jobs for us, it is generating businesses across the globe. ML jobs rank among the best emerging jobs and businesses on LinkedIn, with nearly 2,000 business lists posted and these businesses pay great. In 2022, the average wage for the machine learning engineer base pay was USD108,385 per year in the United States (glassdoor.com, 2022).

ML jobs include technologists, developers, investigators, and information scientists. Like AI and machine education, Robotic operation equipment (RPA), is another application that is automating businesses. RPA is the usage of software to automate business operations, e.g., understanding applications, processing transactions, dealing with information, and even responding to e-mails. RPA automates continual jobs that people accustomed to do. These are not only the menial jobs of the low-paid person: Up to 45% of the actions we do will be automated, including the job of business managers, doctors, and CEOs. ML, reorganized as a distinct set, actually have started to grow in the 1990s. This environment shifted its purpose from achieving artificial intelligence to tackling solvable issues of the functional world. It changed emphasis aside from the symbolic advances it got inherited from AI, and toward methods and models borrowed from statistics and quantity hypothesis (Langley, 2011; Petar, 2023). It has also gained its popularity from the increasing accessibility of digitized data, and the power to disseminate it via the internet. Most of ML algorithm and models that are widely used are known as NN models or a derivation of it from its original architecture.

Artificial Neural Network (NN) can be considered as one of modern contemporary Artificial Intelligence (AI) and a subset of ML. It is an exploration of the