

Universiti Teknologi MARA

**Job Position Prediction Based on Skills
and Experience Using Machine
Learning Algorithm**

Ezaryf bin Hamdan

**Thesis submitted in fulfilment of the requirements
for Bachelor of Computer Science (Hons.)**

JULY 2024

ACKNOWLEDGEMENT

First and foremost, 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. Without His support and guidance, this work would not have been possible. I would like to express my deepest appreciation to all those who contributed to the completion of this research.

Firstly, my special thanks go to my supervisor, Mrs. Norulhidayah Isa for her endless guidance, encouragement, support, and knowledge that had been delivered throughout this project. Then, thank you to my CSP600 and CSP650 lecturer, Mrs. Ummu Fatimah binti Mohd Bahrin that also give a comments and recommendation for all of us and teach us to complete this project.

No one has been more crucial to me in the pursuit of this project than my family members. I owe a special thank you to my beloved parents for their unwavering support, both physically and mentally. They have provided me with encouragement and even financial assistance when I needed it the most. I am grateful to all those with whom I have had the pleasure of working during this project. I would like to express my appreciation to my dearest classmates for their invaluable assistance and emotional support in successfully completing the final year project.

ABSTRACT

In response to the intensifying competition in the job market, job seekers often grapple with the challenge of identifying the most suitable positions based on their skills and experience. This paper proposes a sophisticated Job Position Prediction system utilizing Machine Learning algorithms and leveraging data from LinkedIn profiles. The objective is to develop an innovative and user-friendly platform offering accurate job position predictions to aid job seekers in finding optimal career opportunities. The proposed system integrates data processing modules to preprocess and analyse LinkedIn job posting pages, extracting crucial information such as job titles, company names, years of experience, qualifications, and skills. Text preprocessing ensures consistent data representation and facilitates validation. The Machine Learning algorithm, comprising Random Forest, Linear Regression, XGBoost, SVM, and Stacking Ensemble, is embedded in the system for job position predictions based on the analysed data. The algorithm undergoes rigorous training with a vast dataset to ensure high prediction accuracy and reliability. Accessible through a desktop application, the Job Position Prediction System prioritizes user-friendliness for job seekers. Users input their job skills and years of experience, receiving personalized job position predictions as the system analyses the input and provides the most suitable job positions based on LinkedIn data. The success of this project is evaluated using various metrics, including prediction accuracy, precision, recall, and F1 score. Cross-validation techniques are employed to validate the model's performance and ensure robustness. The development and evaluation of the prediction system adhere to a comprehensive research methodology encompassing data understanding, description, preprocessing, feature extraction, and model evaluation. This research culminates in the creation of an efficient and accurate Job Position Prediction System, empowering job seekers with valuable insights to enhance their job search process.

TABLE OF CONTENT

CONTENTS	PAGE
SUPERVISOR APPROVAL	iii
STUDENT DECLARATION	iv
ACKNOWLEDGEMENT	v
ABSTRACT	vi
TABLE OF CONTENT	vii
LIST OF FIGURES	xi
LIST OF TABLES	xiv
LIST OF ABBREVIATIONS	xv
 CHAPTER ONE: INTRODUCTION	
1.1 Background of Study	17
1.2 Problem Statement	20
1.3 Objective	22
1.4 Project Scope	23
1.5 Project Significance	25
1.6 Overview of Research Framework	27
1.7 Conclusion	30
 CHAPTER TWO: LITERATURE REVIEW	
2.1 Job Position Placement: An Overview	31
2.2 Machine Learning	33
2.2.1 Machine Learning and Human Learning	37
2.2.2 Classification of Machine Learning Methods	38
2.2.3 Classification of Prediction Algorithm	39
2.2.3.1 Comparison of Different Approaches of Algorithm	41
2.2.4 Job Position Prediction Techniques	42

2.2.4.1 Existing Techniques and Methodologies in Job Prediction	43
2.3 Random Forest Algorithm	46
2.3.1 Overview of Random Forest Algorithm	48
2.3.2 Suitability of Random Forest Algorithm	50
2.4 Logistic Regression Algorithm	51
2.4.1 Overview of Logistic Regression Algorithm	51
2.4.2 Suitability of Logistic Regression Algorithm	53
2.5 XGBoost Algorithm	54
2.5.1 Overview of XGBoost Algorithm	54
2.5.2 Suitability of XGBoost Algorithm	57
2.6 Support Vector Machine Algorithm	58
2.6.1 Overview of Support Vector Machine Algorithm	59
2.6.2 Suitability of Support Vector Machine Algorithm	61
2.7 Algorithm Ensemble Method	63
2.7.1 Overview of Stacking Ensemble	64
2.7.2 Suitability of Stacking Ensemble	65
2.8 LinkedIn as a Source of Data	66
2.8.1 Introduction to LinkedIn	66
2.8.2 Advantages of Using LinkedIn Data	66
2.8.3 Challenges and Limitations of LinkedIn Data	68
2.9 Implementation of Random Forest in Various Problem	69
2.10 Related Work in Job Position Prediction	76
2.11 The Implication of Literature Review	87
2.12 Conclusion	89

CHAPTER THREE: METHODOLOGY

3.1 Overview of Research Methodology Framework	91
3.1.1 Research Methodology Framework	92
3.1.2 Significance of the Chosen Machine Learning Algorithm	98
3.2 Preliminary Study	99
3.2.1 Problem Statement Formulation	100