RECRUIT BEST CANDIDATES WITH MACHINE LEARNING

Selvantharajah Thushanth

168272D

Degree of Master of Science

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

June 2018

RECRUIT BEST CANDIDATES WITH MACHINE LEARNING

Selvantharajah Thushanth

168272D

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science specializing in Data Science, Engineering and Analytics

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

June 2018

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: ……..… Date:…….

Selvantharajah Thushanth

The above candidate has carried out research for the Masters of Science thesis under my supervision.

Name of the supervisor: Dr. Charith Chitraranjan

Signature of the supervisor: ……. Date: …….

Abstract

In this research, I propose a robust approach for predicting personality traits of job candidates using machine learning. Relationship between personality traits and job performance has been studied extensively during the past few decades and thus this relationship can be utilized to overcome limitations in choosing the right candidates.

The proposed approach uses scenario-based analysis using machine learning techniques. Candidates will be asked to take part in scenario-based written conversations and their personality traits will be extracted from these conversations using machine learning techniques. Exacted personality traits of the candidates will be compared with the required job related characteristics in order to evaluate the fitness for the position for which candidates are applying. In order to categorize personality traits of candidates, the Five Factor model is used. Existing methods of evaluating personality traits such as standard set of questionnaires are susceptible to candidates providing false information and also time consuming.

Besides candidates' qualifications, knowledge and experience, candidates' personality traits also used to rank the candidates and shortlist them for face-to-face interviews. Thus, this technique not only allows recruiting right candidates to right position but also reduces significant amount of time and cost spent on evaluating candidates' suitability for given a job position by reducing the number of interviews to conduct. Further, this proposed system can be incorporated into existing e-recruitment system thus leveraging its effectiveness. Therefore, it is beneficial for companies since the proposed system helps to reduce cost and time consumption in the recruitment process while assisting them to choose more suitable candidates for a particular job position.

ACKNOWLEDGEMENTS

I would like to acknowledge the support and motivation given by Dr. Charith Chitraranjan, my supervisor, on successfully completing my research work. I also would like to thank him for his continuous suggestions and guidance throughout my research work. I wish to express my sincere gratitude to Dr. Surangika Ranathunga, for her valuable feedbacks during the initial phase of my research work.

Many thanks and appreciations go to all the staff members of Department of Computer Science and Engineering, University of Moratuwa for their generous help and motivation. Finally yet importantly, I would like to thank to my family for continuous support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
MAE	Mean Absolute Error
MBTI®	Myers-Briggs Type Indicator®
MSE	Mean Squared Error
POS	Part-Of-Speech
RAE	Relative Absolute Error
RMSE	Root Mean Squared Error
RSE	Relative Squared Error
SMO	Sequential Minimal Optimization