



Bahria University
Discovering Knowledge

FINAL YEAR PROJECT REPORT

STACK OVERFLOW: CROWD SOURCING RECOMMENDER SYSTEM

By

MUHAMMAD BILAL LANGAH	(39258)
MOIZ ANWER	(39252)
M FAROOQ ASGHAR KHOKHAR	(39262)
MUHAMMAD AHMED	(39245)
SYED HAMMAD TOUQEER	(39332)

SUPERVISED BY
(MISS FASIHA IKRAM)

BAHRIA UNIVERSITY (KARACHI CAMPUS)

2018

ACKNOWLEDGEMENTS

We would like to thank everyone who had contributed to the successful completion of this project. We would like to express my gratitude to my research supervisor, Miss Fasiha Ikram for her invaluable advice, guidance and her enormous patience throughout the development of the research.

With these lines we wish to acknowledge team members for their hard work and untiring efforts they put in to make this report a success. Without them it would not have been possible.

We would like to take opportunity of this platform to show our gratitude to The Directors of Bahria University and our Head of Department for providing us with the opportunity to work for this final project. We are obliged to be a part of this esteemed institute and are thankful for your support and cooperation.

We are deeply gratified to the Almighty Allah and our wonderful parents for their constant guidance, support and encouragement.

We would also like to present our thanks to Google, Wikipedia and several other sites that helped us through all the data analysis and information for which they are highly appreciated.

In the last again, expressing our profound gratitude to our mentor and Supervisor, Miss Fasiha Ikram, "Ma'am, you make beautiful things grow in our mind with your guidance and teaching. Thank you for your support and patience and for believing in us. As our supervisor and as a person, you have inspired us throughout".

STACK OVERFLOW: CROWD SOURCING RECOMMENDER SYSTEM

ABSTRACT

The objective of this project is to develop a recommender system with help of machine learning algorithms. The recommender system will be able to recommend jobs itself by comparing tags of the user present in the dataset. This report explores techniques used or involved in designing recommender system. Different stages involving data pre-processing stage, data cleaning and database design will be studied and discussed. Finally the end product will be the obviously the recommender system which will recommend jobs which is basically the main functionality and we students will try to integrate it with user friendly interface or will develop a website which will be our secondary task.

This project uses the Machine learning technique (binary classification) to develop the system. The main advantage of using this technique is that it provides the result by simply calculating its probability of occurrence.

The system first checks the user field of interest (from the dump dataset of stack overflow) that in which field he is answering the questions most in stack overflow then the system will automatically recommend the job that are present in our dataset and displays to the user.

TABLE OF CONTENTS

	DECLARATION	ii
	APPROVAL FOR SUBMISSION	iv
	ACKNOWLEDGEMENTS	vi
	ABSTRACT	vii
	TABLE OF CONTENTS	viii
	LIST OF FIGURES	x
	LIST OF TABLES	xii
	LIST OF APPENDICES	xii
CHAPTER		
1	INTRODUCTION	14
	1.1 Background	14
	1.2 Problem Statements	16
	1.3 Aims and Objectives	16
	1.4 Scope of Project	17
2	LITERATURE REVIEW	18
	2.1 Introduction to Recommendation System	18
	2.2 Introduction to Techniques Used In Recommender System	18
3	DESIGN AND METHODOLOGY	20
	3.1 Scrum Software Model	20
	3.2 Pre-Processing	22
	3.3 Working Mechanism	26
	3.4 Block Diagram	27

3.5	Bayesian Equation for Processing	28
3.6	Principal Component Analysis (PCA)	28
3.7	Database Design ER-Diagram	30
3.7.1	Tables Identified (Screenshots)	30
4	IMPLEMENTATION	40
4.1	Implementation	40
4.2	Source Code	41
5	RESULTS AND DISCUSSIONS	53
5.1	Results	53
6	CONCLUSION AND RECOMMENDATIONS	54
6.1	Conclusion	54
6.2	Future Work	54
	REFERENCES	55
	APPENDICES	58