



Bahria University

Discovering Knowledge

FINAL YEAR PROJECT REPORT

COMPARATIVE ANALYSIS OF UBER AND CAREEM SERVICES BASED ON THEIR CUSTOMER REVIEWS USING OPINION MINING

In fulfillment of the requirement
For degree of Bachelors in Information Technology (BS IT)

By

**ILYAS NAWAZ
NIDA MASOOD**

**39111
41229**

SUPERVISED

BY

FATIMA BASHIR

BAHRIA UNIVERSITY (KARACHI CAMPUS)

2019

ACKNOWLEDGEMENTS

I would like to thank everyone who had contributed to the successful completion of this project. I would like to express my gratitude to my research supervisor, Ms. Fatima Bashir for her invaluable advice, guidance and her enormous patience throughout the development of the research.

In addition, I would also like to express my gratitude to our loving parent and friends who had helped and given us encouragement.

Comparative Analysis of Uber and Careem Services based on their Customer Reviews using Opinion Mining

ABSTRACT

Opinion mining is extracting opinions and feelings from a contextual information that defines certain things, products or services around the globe. Sentiment analysis on the extracted data and information about the people's sentiments and feeling about that specific product or service. Our goal was to mine the opinions from social media and perform sentiment analysis on the widespread services of well-known taxi companies namely as Uber and Careem. Uber and Careem came and conquered their audience in no time. They had remarkable services to pick up their customer from their current places in no time and drop them off to their desired destination in a respectable manner and ensure that they had a safe and pleasant journey all the way long. Applying different algorithms and methods on the refined data set collected from the social media gave us the opportunity to analysis the accuracy of the applied algorithms and to know which company served their customers better.

Table of Contents

DECLARATION	1
APPROVAL FOR SUBMISSION	2
ACKNOWLEDGEMENTS	4
ABSTRACT	5
Table of Contents	6
TABLE OF FIGURES	8
LIST OF TABLES	9
LIST OF APPENDIX	10
CHAPTER 1	11
INTRODUCTION	11
1.1 Background	11
1.2 Problem Statement	12
1.3 Objectives	13
1.4 Motivation	13
1.5 Scope of the Project	13
CHAPTER 2	14
LITERATURE REVIEW	14
CHAPTER 3	19
DESIGN AND METHODOLOGY	19
CHAPTER 4	23
IMPLEMENTATION	23
4.1 Data preprocessing	23
4.1.1 Stop word removal	23
4.1.2 Content Based Spam Classification in Twitter	23
4.1.3 Detailed Overview regarding Twitter dataset Collection	24
4.1.3.1 Account Creation	24
4.1.3.2 Keys required:	26

4.1.3.3 Code minutiae:	27
4.1.3.4 For collecting application keys	27
4.1.3.5 Searching for tweets:	28
DataSets in CSV Format	29
4.3 Applying Machine Learning Algorithms on the Datasets	32
CHAPTER 5	33
ANALYSIS AND RESULTS	33
5.1 Performance of Classifier:	33
CHAPTER 6	38
CONCLUSION	38
REFERENCES	40
APPENDIX A	42
SOURCE CODE	42
APPENDIX B	50
CODE EXPLAINED BLOCK WISE	50
Figure 4.1: Comparative Analysis	27
Figure 5.1: Logistic Sentiment Analysis of User Datasets	33
Figure 5.2: Logistic Sentiment Analysis of Custom Dataset	38
Figure 5.3: User - Custom Accuracy with respective machine learning algorithms for each dataset	33