

Bapurao Bandgar

Social Networks and their Opinion Mining

YOUR KNOWLEDGE HAS VALUE



- We will publish your bachelor's and master's thesis, essays and papers
- Your own eBook and book - sold worldwide in all relevant shops
- Earn money with each sale

Upload your text at www.GRIN.com
and publish for free



Bibliographic information published by the German National Library:

The German National Library lists this publication in the National Bibliography; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de> .

This book is copyright material and must not be copied, reproduced, transferred, distributed, leased, licensed or publicly performed or used in any way except as specifically permitted in writing by the publishers, as allowed under the terms and conditions under which it was purchased or as strictly permitted by applicable copyright law. Any unauthorized distribution or use of this text may be a direct infringement of the author s and publisher s rights and those responsible may be liable in law accordingly.

Imprint:

Copyright © 2019 GRIN Verlag
ISBN: 9783346101068

This book at GRIN:

<https://www.grin.com/document/512868>

Bapurao Bandgar

Social Networks and their Opinion Mining

GRIN - Your knowledge has value

Since its foundation in 1998, GRIN has specialized in publishing academic texts by students, college teachers and other academics as e-book and printed book. The website www.grin.com is an ideal platform for presenting term papers, final papers, scientific essays, dissertations and specialist books.

Visit us on the internet:

<http://www.grin.com/>

<http://www.facebook.com/grincom>

http://www.twitter.com/grin_com

**THE STUDY AND ANALYSIS OF EVOLUTION OF
SOCIAL NETWORK AND THEIR OPINION MINING
USING SENTIMENTAL ANALYSIS**

By

**Dr. Bapurao Mayappa Bandgar
Associate Professor, Dept. of MCA
JSPM's Rajarshi Shahu College of Engineering
Tathawade, Pune**

DECEMBER 2019

ACKNOWLEDGEMENT

I thank the almighty for his boundless blessings showered on me. If words are considered as symbols of approval and tokens acknowledgement, then the words play heralding role of expressing my gratitude to all who have helped me directly and indirectly during my Research work.

I would like to express my heartfelt thanks to my respectable and friendly guide **Dr. S. SHEEJA**, M.C.A., M. Phil, Ph.D., Associate Professor Department of Computer Applications, Karpagam University for her kind supervision. Her wise academic advice and ideas have played an extremely vital role in the work presented in this research work.

My especial sincere thanks and gratitude to co-guide **Dr. Binod Kumar**, Director, JSPM's JICA, Pune-33, without whose instructions, scholarly guidance and ceaseless encouragement I would never succeed in completing this thesis.

I also take an opportunity to thank **Dr. Veni**, Associate Professor & HOD, Department of Computer Science, Karpagam University and **Dr. T. Purushottam**, Professor, Government College of Engineering, Coimbatore for their kind cooperation during the Doctorate Committee Meetings and their valuable guidance and helpful suggestions during the research work.

I am equally indebted to our Hon. Founder Secretary, Prof. T. J. Sawant, JSPM Group of Institute, Pune for permitting me to carry out the research work and granting me leave for the same, as and when required.

I also thank to Dr. Ajay Kumar, Director, JSPM JTC, Pune and Dr. N. S. Nehe, Director, JSPM's NTC for their constant support and motivation in the adjustment of duties and giving leave during the visit to the university time to time and providing the college infrastructure for carrying out the research work such as Digital e-Library and Internet facilities etc.

My special thanks to my wife Ms. Sheetal for her untiring support and good humor that helped to boost my spirit and made the research a joyful activity and I thank my son Ku. Shantanu and my daughter Ku. Shweta for giving me time to study and complete my academic journey delightfully and smoothly.

I would like to specially thank to Mr. Ramanand Potdar, Mr. Rajesh Dhudhal and Mr. Amar Kokare for their valuable guidance and help during this work.

Lastly I, take an opportunity to thank all the Computer Science Department staff and Director, Research, Karpagam University and their staff those who have helped me a lot during the research work directly or indirectly.

BANDGAR BAPURAO MAYAPPA

CONTENTS

S. NO	TITLE	PAGE NO
1	INTRODUCTION	1
1.1	Introduction	1
1.2	What is Social Network Analysis?	1
1.3	Motivation and Problem Statement	4
1.4	Objectives	5
1.5	Outline of Thesis	5
2	LITERATURE REVIEW	7
2.1	Introduction	7
2.2	Apriori-Based Approach	7
2.3	Pattern-Growth Approach	7
2.3.1	Survey and Techniques of Frequent Pattern Mining	7
2.4	Survey and Techniques of Online Social Networks	17
2.5	Survey on Opinion Mining	24
2.6	Summary	34
3	AN ANALYSIS OF SOCIAL NETWORK DATA USING GEPHI	35
3.1	Introduction	35
3.2	Features	35
3.2.1	Real-time Visualization	36
3.2.2	Layout	36
3.2.3	Metrics	37
3.2.4	Dynamic Network Analysis	38
3.2.5	Create Cartography	39
3.2.6	Clustering and Hierarchical graphs	39
3.2.7	Dynamic filtering	39
3.2.8	User-centric	39

S. NO	TITLE	PAGE NO
	3.2.9 Modular	39
	3.2.10 Plug-in center	40
3.3	Implementation of Gephi to Social Network data	40
	3.3.1 Data Collection and Experimental Details	40
3.4	Results and Discussion	40
3.5	Summary	50
4	TOOLS AND TECHNIQUES USED FOR EXTRACTION AND PROCESSING OF THE REAL TIME TWEETS	51
4.1	Introduction	51
4.2	Twitter 4j Library	51
	4.2.1 System Requirements	51
	4.2.2 How To Use	51
	4.2.3 Download	51
	4.2.4 Source Code	52
	4.2.5 Maven Integration	52
4.3	Tweet search method by GET search/tweets in API V1.1	53
4.4	JSOUP JAVA Libraries	56
	4.4.1 Parsing and traversing a Document	57
	4.4.2 Parse a document from a String	57
	4.4.3 Parsing a body fragment	58
	4.4.4 Extract attributes, text, and HTML from elements	59
	4.4.5 Sanitize untrusted HTML (to prevent XSS)	59
4.5	Netbeans IDE 8	60
	4.5.1 Working with NetBeans Modules	61
4.6	Processing of tweets	62
4.7	SentiWordNet	63
	4.7.1 History and team members	63
	4.7.2 Database contents	64

S. NO	TITLE	PAGE NO
	4.7.3 Knowledge structure	65
	4.7.4 Psycholinguistic aspects of WordNet	66
	4.7.5 WordNet as a lexical ontology	67
	4.7.6 Limitations	68
5	EXTRACTION AND THE PROCESSING OF REAL TIME TWEETS	69
	5.1 Introduction	69
	5.2 Experimental Details	71
	5.3 Results and Discussions	74
	5.4 Summary	76
6	CLASSIFICATION OF TWEETS USING SENTIMENTAL ANALYSIS	81
	6.1 Introduction	81
	6.2 Machine learning	81
	6.2.1 Naive Bayes Classification	81
	6.2.2 Maximum Entropy	82
	6.3 Support Vector Machines	83
	6.3.1 What SVM is used for?	84
	6.3.2 How SVM Works	84
	6.4 Evaluation of Sentiment Classification	88
	6.5 Enhanced Emoticon Classification	91
	6.5.1 EEC score calculation	91
	6.6 Improved Polarity Classification	92
	6.6.1 PC score calculation	92
	6.7 SentiwordNet Classification	93
	6.7.1 SWNC score calculation	93
	6.8 Method Details	94
	6.9 Results And Discussions	95
	6.10 Summary	100

S. NO	TITLE	PAGE NO
7	CLASSIFICATION OF TWEETS USING HYBRID MODEL	101
7.1	Introduction	101
7.2	Methods	101
7.3	Results And Discussions	102
7.4	Summary	103
8	CONCLUSIONS AND FUTURE WORK	104
8.1	Conclusions	104
8.2	Future Work	106
9	REFERENCES	107

LIST OF TABLES

TABLE NO	TITLE	PAGE NO
3.1	Obtained metric values of the for social network data by the Gephi	41
5.1	The sample of extracted real time and raw tweets on the ‘Narendra Modi’ Keyword [Mon Jun 16 20:52:20 IST 2014]	77
5.2	The sample of preprocessed tweets related to the search keyword ‘Narendra Modi’	78
5.3	The positive emoticons used for the conversion in to text	79
5.4	The negative emoticons used for the conversion in to text	79
5.5	The sample abbreviations used of the for the conversion in to text	79
5.6	The sample of completely processed tweets sample on the Narendra Modi	80
6.1	Simple data in 1-Dimension	85
6.2	Simple data in 2-Dimension	86
6.3	Calculation Results of positive, negative, and hyperplane	87
6.4	Confusion Matrix	88
6.5	Sample of the positive word used in IPC model	96
6.6	Sample of the negative word used in IPC model	97
6.7	Comparisons of the results of EEC, IPC and SWNC model using confusion matrix and precision, recall and accuracy	100

LIST OF FIGURES

FIGURE NO	TITLE	PAGE NO
3.1	General view of the Ghephi tool	36
3.2	Data preview in Ghephi	37
3.3	Metric parameter of the given social network data	38
3.4	General graphical view of given social network data	41
3.5	The degree distribution of given social network data	42
3.6	Average weighted degree distribution of given social network data	43
3.7	Connected component size distribution of given social network data	43
3.8	The clustering coefficient distribution of the given network data	44
3.9	Modularity class distributions of the social network data	45
3.10	The close centrality measure of the social network data	46
3.11	The Betweenness centrality measure of social network data	46
3.12	Eigen vector centrality distribution of the given social network data	47
3.13	Eccentricity distribution of the social network data	47
3.14	The page rank distribution of the given social network data	48
3.15	The authority distribution of the given social network data	49
3.16	The hub distribution of the given social network data	49
5.1	Detailed architecture of real time tweet Extraction and Processing	72
5.2	Flow chart for real time tweet Extraction and Processing the for obtaining meaningful message	73

FIGURE NO	TITLE	PAGE NO
5.3	The dash board for the tweet extraction	75
5.4	The dash board for the tweet pre-processing	75
5.5	The dash board of completely processed tweets which will used for further classification	76
6.1	Different boundary decisions are possible to separate two classes in two dimensions	83
6.2	SVM process flow	84
6.3	Simple data in an input space	85
6.4	Simple Data in a feature	86
6.5	Simple data in a feature space separated by a hyperplane	87
6.6	Detailed architecture of the framework used for tweet classification	95
6.7	The sample classification results obtained by EEC Model	97
6.8	The sample classification results obtained by IPC Model	98
6.9	The sample classification results obtained by SWNC and their hybrid Model	99
6.10	SWNC Model classification % results in pie graph	99
7.1	Flow chart of hybrid model for real time tweets	101
7.2	Hybrid model classification % results in pie graph	102
7.3	SWNC model classification % results in pie graph	103