



Sentiment Analysis in Kannada

Mr. Gajendra P.¹ and Dr. B. K. Ravindranath²

1. Guest Lecturer and Doctoral Candidate, DOS in Linguistics, Kuvempu Institute of Kannada Studies, University of Mysore, Manasagangotri, Mysuru, Karnataka, India-570006
2. Professor and Chairman (BOS), DOS in Linguistics, Kuvempu Institute of Kannada Studies, University of Mysore, Manasagangotri, Mysuru, Karnataka, India-570006

Abstract

This paper mainly focuses on the idea of Sentiment Analysis in Kannada. As the term 'Sentiment' suggests, it may be a joy, sadness, frustration, angry, unhappy, lamentation, if it is said in totality human feelings or emotions. As we know that man is a social animal. He/ She comes with the relationships and goes with the relationships from the day one to day last. Human cannot live alone and it is inevitable to satisfy some natural basic needs to survive and human cannot break the shackles of society.

Facebook's corpus for Sentiment analysis in Kannada which is a South Dravidian Language. The term 'Sentiment' suggests a vision or opinion that is expressed in any context. As we are all aware of, the human sentiment is a social behavior and at the same time it could be studied from many perspectives. Innumerable researches are being held in the fields of Natural Language Processing and Information Extraction. In the contemporary era, with the furtherance of web technology and its improvement there is a huge volume of data is available in the web for researchers and everyday a lot of data is generated. There are many works that have been taken place towards Sentiment Analysis. The data that has been used in this paper is collected from the social media like Facebook and Kannada & English mixed in nature. The data has been classified based on how the sentiments are expressed- Positive, Overt Positive & Covert Positive, Negative, Overt Negative & Covert Negative and Neutral.

Keywords

Facebook, Social Media, Sentiment Analysis, Kannada, Data Science

INTRODUCTION

Kannada is a South Dravidian Language that is spoken in Karnataka and it is recognized constitutionally. The Kannada Language is not only spoken in Karnataka but also other states and countries as well. Sentiment plays a crucial role in our day-to-day life. Social sites like Facebook, Twitter, Youtube, Google and etc... are getting popularity as they allow people to share and express their views on any topic. There has been lot of work in the field of natural Language processing and Information extraction. “What others think and express” is always a paramount for us. As all we know Sentiment analysis is a Natural Language Processing and Information Extraction work which aims to avail people’s feelings expressed positively or negatively or neutrally by analyzing innumerable comments. In general, Sentiment analysis aims to decide the attitude of a social media users’ with respect to some topics. Nowadays, Internet becomes a platform for online learning, exchanging ideas and sharing opinions and the public opinion is the motivation behind Sentiment analysis today. As far as our knowledge is concerned web is a godown of structured and unstructured data. The investigation of this huge volume of data to extract covert public sentiments is really a challenging task. Sentiment Analysis is also called opinion mining. Opinion is central to human activities and influences our behaviors. Most of the time it is sought others opinions regarding any topic during decision making process; it is not only for individuals but also for institutions. Sentiment analysis puts light on sentiments, evaluations, examinations, attitudes, emotions and etc... because these are all subjects of study of Sentiment Analysis. It is evaluated or examined reviews, comments, discussions, opinions, blogs that are expressed on social media and since huge data opinionated in digital form, it is easy to analyzed social media resources.

PREVIOUS WORK

It has not been found many sources on this topic, although it has been found some articles, books on Sentiment Analysis generally and few literatures could be analyzed here for the purpose of this paper. **“Fundamentals of Sentiment Analysis and its Application”**(Mohsen Farhadloo and Erik Rolland 2016) (– **Research Gate-** <https://www.researchgate.net/publication/300965436> - Chapter · March 2016 – DOI: 10.1007/978-3-319-30319-2_1) The Book says about the problem of identifying people’s opinions expressed in written language is a relatively new and very active field of research. Having access to huge amount of data due to the ubiquity of Internet, has enabled researchers in different fields—such as Natural Language Processing (NLP), Machine Learning and Data Mining, Text Mining, Management and Marketing and even Psychology—to conduct research in order to discover people’s opinions and sentiments from the publicly available data sources. **“SENTIMENT ANALYSIS AND OPINION MINING”** (Bing Liu 2012) April 22, 2012, liub@cs.uic.edu, Due to copyediting, the published version is slightly different, Morgan & Claypool Publishers May 2012. Synthesis Lectures on “Human Language Technologies”- Graeme Hirst, Series Editor), the book discusses various facets of the Sentiment Analysis (SA) from the perspective of structured approach and gives an in-depth introduction to this fascinating problem and to present a comprehensive survey of all important researches and paramount of the latest developments in the field. As evidence of that the book covers more than 400 references from all major conferences and journals.

TWITTER SENTIMENT ANALYSIS”(Afroze Ibrahim Baqapuri 2012) Bachelors in Electrical (Electronics) Engineering Department of Electrical Engineering, School of Electrical Engineering & Computer Science, National University of Sciences & Technology, Islamabad, Pakistan), the Project report, encounters the problems of sentiment analysis in twitter; which is categorizing tweets according to the sentiment expressed in them: positive, negative, and neutral. As we know that Twitter is an online micro-blogging and social-networking platform which allows users to write short status updates of maximum length 140 characters. It is also proved that analyzing the public sentiment is important for many applications such as firms trying to find out the response.

“TWITTER EMOTION ANALYSIS” (Supervisor, Dr David Rossiter, Marc Lamberti - marclamberti.ml@gmail.com MSc, Information Technology, Hong Kong University of Science and Technology, China- 2015), The project report says about the sentiment analysis field of machine learning since it allows approaching natural language processing which is a very hot topic actually. Following author’s previous experience where it was about classifying short music according to their emotion, author applied the same idea with tweets and try to figure out which is positive or negative. Sentiment analysis, also refers as opinion mining, is a sub machine learning task where we want to determine which the general sentiment of a given document is.

SOURCES OF DATA

The data that has been used in this paper are collected from the social media like Facebook.

DISCUSSION

Sentiment Analysis evaluates public opinions that are expressed on social media. It involves categorizing opinions like positive, negative or neutral. It can also be done as subjectivity analysis, opinion mining and appraisal extraction and so on. The words like opinion, sentiment, view and belief are used interchangeably but there are differences between them. Opinion says about a conclusion open to dispute, View says on subjective opinion, Belief says on deliberate acceptance and intellectual assent and Sentiment says on opinion representing one’s feelings. Through the collected data which mainly focuses on the sentiments could be carved up into categories like Positive, Overt Positive, Covert Positive, Negative, Overt Negative, Covert Negative and Neutral and classified according to polarity.

1	FB comment	Nod-beda	bidappa	Ninobba	nodod-bitre	enu	loss	illa
	Gloss	See- don’t	leave	You-alone	See-left	what	loss	no
	Translation	‘There is no loss at all if you don’t see’						
	Polarity	Negative						
2	FB comment	sangolli		rayannan		hesr-alli		od-thu
	Gloss	Sangolli		Rayanna		name-in		ran-it
	Translation	‘It succeeded because of the name Sangolli Rayanna’						
	Polarity	Covert Positive						

3	FB comment	kan	munde	preetha-davaru	na	bidka-dode	pain	ad-kinta	pain	en	ede	life	ali
	Gloss	eye	front	love- they	particle	Leave-go	pain	that-than	pain	what	there	life	in
	Translation	‘There is no deep pain than leave our beloved one’											
	Polarity	Positive											

4	FB comment	yappa	ah cc	matra	nodak-agalla
	Gloss	alas	-----	only	to see - never
	Translation	‘It can’t be seen’			
	Polarity	Negative			

5	FB comment	firstu.....	ellaru	heavy	talented	agok-agala
	Gloss	First	everyone	heavy	talented	become-never
	Translation	‘First everyone can’t become a heavy talented’				
	Polarity	Covert Negative				

6	Sentence	correct sir kumar	ivrella	kannda	film	indale	kannada	ankond	bittavre
	Gloss	correct Kumar	sir these people	Kannada	film	from	Kannada	thought	left
	Translation	‘Correct Kumar sir....they thought that Kannada remains only through Kannada films’							
	Polarity	Covert Positive							

CONCLUSION

To sum up this paper, it is provided different types of sentiments which expressed on Facebook that is source for this paper and relevant literature have been studied. Many research papers could be penned on this topic and one source is relied here that is Facebook, but there are many sources, like Twitter, YouTube, Google+ and etc...these can also be taken as sources. In the contemporary time huge volume of data available for linguists as I said before, so a new research view in the background of Linguistics is expected and analysis of social interaction is a new avenue for Linguistics research.

REFERENCES

- Albert Biffet, Eibe Frank, (2010), Sentiment Knowledge Discovery in Twitter Streaming Data. Discovery Science, Lecture Notes in Computer Science, Volume 6332/2010, 1-15, DOI: 10.1007/978-3-642-16184-1_1
- Alec Go, Richa Bhayani, Lei Huang, (2009), Twitter Sentiment Classification using Distant Supervision, Project Technical Report, Stanford University
- Alexander Pak and Patrick Paroubek, (2010), Twitter as a Corpus for Sentiment Analysis and Opinion Mining, In Proceedings of international conference on Language Resources and Evaluation (LREC)
- Andranik Tumasjan, Timm O. Sprenger, Philipp G. Sandner and Isabell M. Welp, (2010), Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment, In Proceedings of AAAI Conference on Weblogs and Social Media (ICWSM)
- Adeep Hande, Ruba Priyadharshini, Bharathi Raja Chakravarthi (2020), KanCMD: Kannada CodeMixed Dataset for Sentiment Analysis and Offensive Language Detection, Conference Proceedings, Third Workshop on Computational Modelling of People's Opinions, Personality, and Emotions in Social media, pages 54-63
- Bo Pang, Lillian Lee and Shivakumar Vaithyanathan (2002), Thumbs up? Sentiment Classification using Machine Learning Techniques, In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)
- Chenhao Tan, Lillian Lee, Jie Tang, Long Jiang, Ming Zhou and Ping Li (2011), User Level Sentiment Analysis Incorporating Social Networks, In Proceedings of ACM Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD)
- Efthymios Kouloumpis, Theresa Wilson and Johanna Moore (2011), Twitter Sentiment Analysis: The Good the Bad and the OMG, In Proceedings of AAAI Conference on Weblogs and Social Media (ICWSM)
- Sadaf Rani, Muhammad Waqas Anwar, (2020), Resource creation and Evaluation of Aspect Based Sentiment Analysis in Urdu, Conference proceedings, 10th International conference on Natural Language Processing: Students research workshop
- Vishal A.Kharde, Sonawane S.S., (2016), Sentiment Analysis of Twitter Data: A Survey of Techniques, International Journal of Computer Applications (0975-8887), Volume 139-No.11