Sentiment Analysis on Covid-19 Impact using Python

Varada Lakshmi Soujanya[1], Samta Gajibhiye[2]

Department of Computer Science and Engineering,
Shri Shankaracharya Group of Institutions, Bhilai, C.G., India

***Abstract*: Coronavirus has represented a genuine pandemic danger the executives challenge as far as effect, readiness, reaction, and alleviation by governments, wellbeing associations, non-legislative associations (NGOs), broad communications, and partners. This investigation inspects the pandemic influenced hazard correspondence in questionable settings and its effect on the feelings and estimations got from the semantic examination in the public eye during the COVID-19 pandemic. This investigation inspects the pandemic influenced hazard correspondence in questionable settings and its effect on the feelings and estimations got from the semantic examination in the public eye during the COVID-19 pandemic. The flare-up of Covid sickness (COVID-19) overturned individuals' lives around the world. Coronavirus is brought about by serious intense respiratory disorder Covid 2 (SARS-CoV-2), a novel human microbe that virologists accept rose up out of bats and ultimately leaped to people through a delegate have. This flare-up has made colossal pulverization in day lives. Conclusion investigation is a field which is getting up to speed in the new years and its applications are liable to increment to a more extensive territory in not so distant future.**

***Keywords*: Sentiment analysis, python, covid-19, machine learning**

I. Introduction:

Text mining, likewise alluded to as text information mining, like content investigation, is the way toward getting excellent data from text. It includes "the disclosure by PC of new, already obscure data, via consequently removing data from various composed assets." Composed assets may incorporate sites, books, messages, audits, and articles. Great data is normally acquired by conceiving examples and patterns by means, for example, factual example learning. As per Hotho et al. (2005) we can vary three alternate points of view of text mining: data extraction, information mining, and a KDD (Information Disclosure in Data sets) process. Text mining for the most part includes the way toward organizing the info text (typically parsing, alongside the expansion of some determined semantic highlights and the evacuation of others, and ensuing inclusion into a data set), inferring designs inside the organized information, lastly assessment and translation of the yield. 'Superior grade' in text mining for the most part alludes to a mix of significance, curiosity, and interest. Regular content mining undertakings incorporate content arrangement, text bunching, idea/element extraction, creation of granular scientific categorizations, notion examination, archive rundown, and element connection displaying (i.e., learning relations between named elements).

Text examination includes data recovery, lexical investigation to consider word recurrence appropriations, design acknowledgment, labeling/explanation, data extraction, information mining procedures including connection and affiliation investigation, perception, and prescient investigation. The general objective is, basically, to transform text into information for examination, through use of normal language handling (NLP), various sorts of calculations and insightful strategies. A significant period of this interaction is the translation of the assembled data.

1.1 Text analysis process:

Subtasks—parts of a bigger book examination exertion—ordinarily include:

- Dimensionality decrease is significant method for pre-preparing information. Procedure is utilized to recognize the root word for genuine words and decrease the size of the content information.
- Information recovery or recognizable proof of a corpus is a preliminary advance: gathering or distinguishing a bunch of literary materials, Online or held in a document framework, data set, or substance corpus chief, for examination.
- Although some content examination frameworks apply only progressed factual techniques, numerous others apply more broad regular language preparing, for example, grammatical form labeling, syntactic parsing, and different sorts of semantic investigation.
- Named element acknowledgment is the utilization of gazetteers or measurable strategies to distinguish named text highlights: individuals, associations, place names, stock ticker images, certain truncations, etc.
- Disambiguation—the utilization of logical hints—might be needed to choose where, for example, "Portage" can allude to a previous U.S. president, a vehicle maker, a famous actor, a stream intersection, or some other element.
- Recognition of Example Distinguished Elements: Highlights, for example, phone numbers, email addresses, amounts (with units) can be perceived by means of customary articulation or other example matches.
- Document bunching: ID of sets of comparative content records.
- Coreference: ID of thing phrases and different terms that allude to a similar item.
- Relationship, truth, and occasion Extraction: ID of relationship among substances and other data in text

- Sentiment examination includes knowing abstract (instead of verifiable) material and separating different types of attitudinal data: feeling, assessment, state of mind, and feeling. Text examination methods are useful in dissecting supposition at the element, idea, or subject level and in distinctive assessment holder and assessment object.
- Quantitative content investigation is a bunch of methods coming from the sociologies where either a human appointed authority or a PC removes semantic or linguistic connections between words to discover the significance or expressive examples of, typically, a relaxed individual book with the end goal of mental profiling and so forth.

II. Proposed methodology

Natural learning process:

Regular language preparing (NLP) is a subfield of etymology, software engineering, and man-made brainpower worried about the connections among PCs and human language, specifically how to program PCs to measure and break down a lot of normal language information. The outcome is a PC prepared to do "understanding" the substance of records, including the relevant subtleties of the language inside them. The innovation can then precisely remove data and bits of knowledge contained in the records just as classify and arrange the actual reports.

Difficulties in regular language handling every now and again include discourse acknowledgment, normal language comprehension, and normal language age.

During the 2010s, portrayal learning and profound neural organization style AI strategies got broad in normal language preparing, due to some degree to a whirlwind of results showing that such procedures can accomplish best in class brings about numerous regular language undertakings, for instance in language displaying, parsing and numerous others.

Techniques: rules, insights and neural organizations:

In the good 'ol days, numerous language-preparing frameworks were planned by representative strategies, i.e., the hand-coding of a bunch of rules, combined with a word reference query, for example, by composing sentence structures or conceiving heuristic principles for stemming.

Later frameworks dependent on AI calculations enjoy numerous upper hands over hand-created rules:

- The learning methods utilized during AI consequently center around the most widely recognized cases, though when composing rules by hand it is normal not in any way clear where the exertion ought to be coordinated.
- Automatic learning systems can utilize measurable derivation calculations to create models that are hearty to new information (for example containing words or constructions that have not been seen previously) and to wrong information (for example with incorrectly spelled words or words inadvertently excluded). By and large, taking care of such information effortlessly with written by hand rules, or, all the more for the most part, making frameworks of transcribed guidelines that settle on delicate choices, is very troublesome, blunder inclined and tedious.
- Systems dependent on naturally learning the principles can be made more precise essentially by providing more info information. Notwithstanding, frameworks dependent on transcribed principles must be made more exact by expanding the intricacy of the standards, which is a substantially more troublesome assignment. Specifically, there is a breaking point to the intricacy of frameworks dependent on written by hand runs, past which the frameworks become increasingly unmanageable. Nonetheless, making more information to contribution to AI frameworks basically requires a relating expansion in the quantity of worker hours worked, for the most part without huge expansions in the intricacy of the comment interaction.

Notwithstanding the ubiquity of AI in NLP research, emblematic techniques are still (2020) generally utilized

- when the measure of preparing information is inadequate to effectively apply AI strategies, e.g., for the machine interpretation of low-asset dialects, for example, given by the Apertium framework,
- for preprocessing in NLP pipelines, e.g., tokenization, or
- for postprocessing and changing the yield of NLP pipelines, e.g., for information extraction from syntactic parses.

1. Statistical strategies:

Since the alleged "factual upset" in the last part of the 1980s and mid-1990s, much regular language preparing research has depended intensely on AI. The AI worldview calls rather for utilizing measurable induction to consequently learn such standards through the investigation of huge corpora (the plural type of corpus, is a bunch of reports, conceivably with human or PC comments) of ordinary genuine models.

A wide range of classes of AI calculations have been applied to normal language-handling assignments. These calculations take as information an enormous arrangement of "highlights" that are produced from the information. Progressively, in any case, research has zeroed in on measurable models, which make delicate, probabilistic choices dependent on joining genuine esteemed loads to each info highlight. Such models enjoy the benefit that they can communicate the overall sureness of a wide range of potential answers as opposed to just one, creating more solid outcomes when a particularly model is incorporated as a segment of a bigger framework.

The absolute soonest utilized AI calculations, for example, choice trees, delivered frameworks of hard assuming principles like existing manually written standards. Notwithstanding, grammatical form labeling presented the utilization of covered up Markov models to normal language handling, and progressively, research has zeroed in on measurable models, which make delicate, probabilistic choices dependent on connecting genuine esteemed loads to the highlights making up the information. The reserve language models whereupon numerous discourse acknowledgment frameworks currently depend are instances of such factual models. Such models are for the most part more vigorous when given new information, particularly input that contains mistakes (as is exceptionally normal for genuine information), and produce more dependable outcomes when incorporated into a bigger framework including various subtasks.

Since the neural turn, measurable strategies in NLP research have been to a great extent supplanted by neural organizations. Be that as it may, they keep on being pertinent for settings in which factual interpretability and straightforwardness is required.

2. Neural organizations:

A significant disadvantage of factual strategies is that they require elaborate component designing. Since the mid 2010s, the field has consequently generally deserted measurable strategies and moved to neural organizations for AI. Mainstream procedures incorporate the utilization of word embeddings to catch semantic properties of words, and an expansion in start to finish learning of a more elevated level errand (e.g., question replying) rather than depending on a pipeline of isolated transitional assignments (e.g., grammatical feature labeling and reliance parsing). In certain spaces, this shift has involved considerable changes in how NLP frameworks are planned, to such an extent that profound neural organization based methodologies might be seen as another worldview particular from factual normal language preparing. For example, the term neural machine interpretation (NMT) stresses the way that profound learning-based ways to deal with machine interpretation straightforwardly learn grouping to-succession changes, hindering the requirement for middle of the road steps, for example, word arrangement and language demonstrating that was utilized in factual machine interpretation (SMT). Most recent works will in general utilize non-specialized design of an offered undertaking to fabricate legitimate neural organization.

Python:

Python is a deciphered, significant level and broadly useful programming language. Python's plan theory underscores code comprehensibility with its remarkable utilization of huge space. Its language develops and object-arranged methodology plan to assist software engineers with composing, legitimate code for little and enormous scope projects.

Python is powerfully composed and trash gathered. It upholds numerous programming ideal models, including organized (especially, procedural), object-situated and useful programming. Python is frequently depicted as a "batteries included" language because of its far reaching standard library.

III. Result and discussion:

The basic target of this exploration subject is to spot out the feelings and assessments of the clients or clients by means of a study premise. Despite the fact that various examination works have been brought placed in this field through different models, feeling investigation is as yet thought to be a difficult issue with such countless contentions to be addressed. The review has been led on coordinated premise with 66 typical individuals for example individuals with no physical and mental issues of various ages. Coming up next are the inquiries that they were posed:

1. As we were all very ill-equipped when the report about COVID'19 broke , what was your underlying response to it?
2. As public spots have been shut for around a couple of months presently how might you depict your versatility to the circumstance, remaining back at home?
3. Covid'19 has given us sufficient opportunity to go through at home with our families, how has this influenced your relationship with the relatives as you will see a greater amount of them now?
4. Staying ceaselessly, understanding news or media communicates and the encompassing climate around any pandemic would influence an individual's condition of care. How would you think this pandemic influenced you intellectually?
5. E-learning is being adjusted all through the world as a medium to spread information and stay up with the latest with their picked courses. What are your musings on E-learning, has it been productive to you? Would you actually pick grounds classes? Will there be a few enhancements to improve online appraisal?
6. With vulnerability approaching up over our heads, how has the present circumstance influenced your future possibilities of school affirmations, aces abroad or occupations?

7. How has the present circumstance where we need to remain back at home or keep up friendly separating when leaving need influenced your social connections? How are you actually staying in contact with your companions and individuals you used to associate with? Has this pandemic influenced these connections? Assuming this is the case, how?
8. Based on the answers obtained of the above questions below are the results obtained:

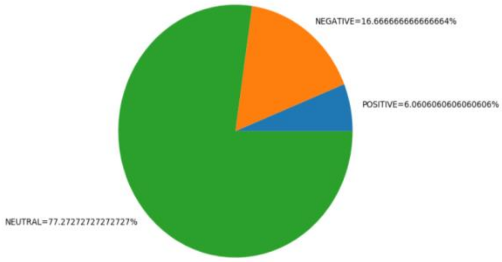

Figure 3.1: Collective response to question 1

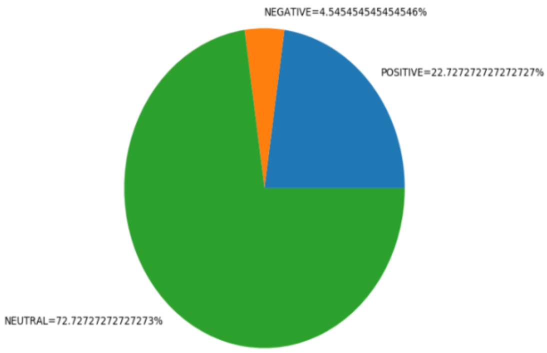

Figure 3.2: Collective response to question 2

Figure 3.1 shows the collective response to question 1 that shows negative response is 16.67%, positive response is 22.72% and neutral response is 77.27%.

Figure 3.2 shows the collective response to question 2 that shows negative response is 4.45%, positive response is 6.06% and neutral response is 77.27%.

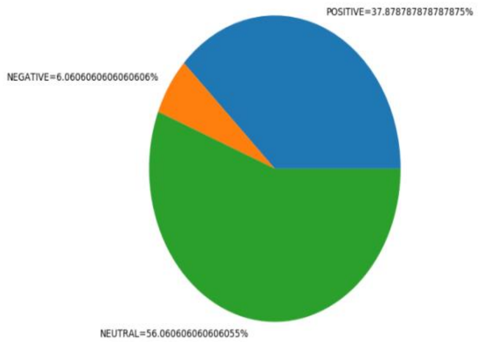

Figure 3.3: Collective response to question 3

Figure 3.3 shows the collective response to question 3 that shows negative response is 6.06%, positive response is 37.88% and neutral response is 56.06%.

Figure 3.4 shows the collective response to question 4 that shows negative response is 16.67%, positive response is 16.67% and neutral response is 56.06%.

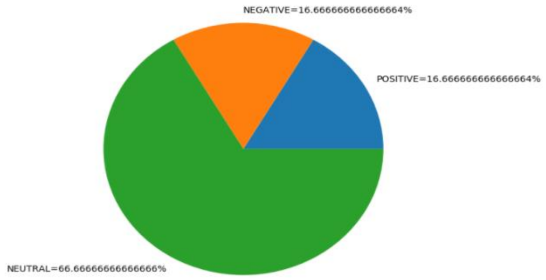

Figure 3.4: Collective response to question 4

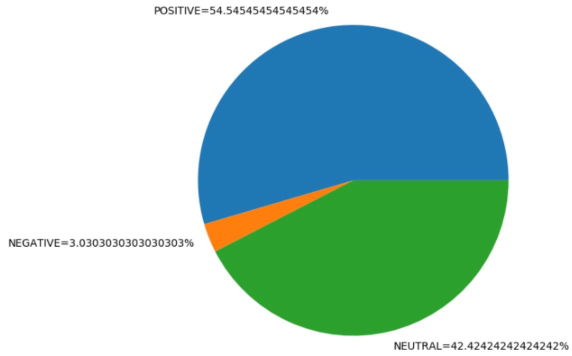

Figure 3.5: Collective response to question 5

Figure 3.5 shows the collective response to question 5 that shows negative response is 16.67%, positive response is 16.67% and neutral response is 56.06%.

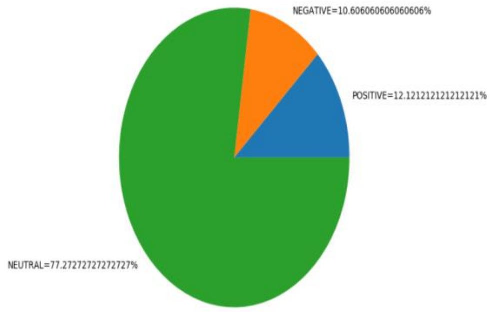

Figure 3.6: Collective response to question 6

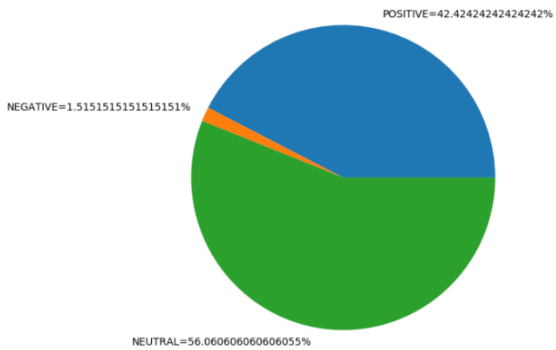

Figure 3.7: Collective response to question 7

Figure 3.6 shows the collective response to question 6 that shows negative response is 10.67%, positive response is 12.12% and neutral response is 77.27%.

Figure 3.7 shows the collective response to question 7 that shows negative response is 1.51%, positive response is 42.42% and neutral response is 56.06%.

IV. Conclusion:

Coronavirus has changed our lives for eternity. The world we knew as of not long ago has been changed and these days we live in a totally new situation in an interminable rebuilding progress, in which the way we live, relate, and speak with others has been adjusted forever. Inside this specific circumstance, hazard correspondence is assuming a definitive job when illuminating, sending, and directing the progression of data in the public eye. Coronavirus has represented a genuine pandemic danger the executives challenge as far as effect, readiness, reaction, and alleviation by governments, wellbeing associations, non-legislative associations (NGOs), broad communications, and partners. This investigation inspects the pandemic influenced hazard correspondence in questionable settings and its effect on the feelings and estimations got from the semantic examination in the public eye during the COVID-19 pandemic.The flare-up of Covid sickness (COVID-19) overturned individuals' lives around the world. Coronavirus is brought about by serious intense respiratory disorder Covid 2 (SARS-CoV-2), a novel human microbe that virologists accept rose up out of bats and ultimately leaped to people through a delegate have. This flare-up has made colossal pulverization in day lives. Conclusion investigation is a field which is getting up to speed in the new years and its applications are liable to increment to a more extensive territory in not so distant future.

References:

[1] A. D. Dubey, “Decoding the Twitter sentiments towards the leadership in the times of COVID-19: A case of USA and india,” SSRN Electron. J., Apr. 2009, doi: 10.2139/ssrn.3588623.

[2] Anvar Shathik J., Krishna Prasad K., “A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques”, International Journal of Applied Engineering and Management Letters (IJAEML), ISSN: 2581-7000, Vol. 4, No. 2, August 2020.

[3] Ashmira Khan, Goshiya Sheikh, Ruchi Agrawal, Raima Rai, Nikita Soni, Mukesh Roy, Sanjay Kalamdhad, “Implementation of Twitter Sentimental Analysis According to Hash Tag”, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 07 Issue: 03 | Mar 2020.

[4] Augustyniak, Łukasz; Szymański, Piotr; Kajdanowicz, Tomasz; Tuligłowicz, Włodzimierz (2015-12-25). "Comprehensive Study on Lexicon-based Ensemble Classification Sentiment Analysis". Entropy. 18 (1): 4. Bibcode:2015Entrp..18....4A. doi:10.3390/e18010004.

[5] C. J. H. E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text,” in Eighth International Conference on Weblogs and Social Media (ICWSM-14). Available at (20/04/16) http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf, 2014.

[6] Chaturvedi, Iti; Cambria, Erik; Welsch, Roy E.; Herrera, Francisco (November 2018). "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges" (PDF). Information Fusion. 65–77. doi:10.1016/j.inffus.2017.12.006 – via Elsevier Science Direct.

[7] Dimple Chehal , Parul Gupta and Payal Gulati, “COVID-19 pandemic lockdown: An emotional health perspective of Indians on Twitter”, International Journal of Social Psychiatry, 2020.

[8] Dr K B Priya Iyer, Dr Sakthi Kumaresh, “Twitter Sentiment Analysis On Coronavirus Outbreak Using Machine Learning Algorithms”, European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 07, Issue 03, 2020.

[9] G. Dharani Devi1 and Dr. S .Kamalakkannan, “Literature Review on Sentiment Analysis in Social Media: Open Challenges toward Applications”, International Journal of Advanced Science and Technology Vol. 29, No. 7, (2020), pp. 1462-1471.

[10] Jim Samuel, G. G. Md. Nawaz Ali, Md. Mokhlesur Rahman, Ek Esawi, Yana Samuel, “COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification”, Information 2020, 11, 314; doi:10.3390/info11060314.

[11] Koppel, Moshe; Schler, Jonathan (2006). "The Importance of Neutral Examples for Learning Sentiment". Computational Intelligence 22. pp. 100–109. CiteSeerX 10.1.1.84.9735.

[12] Liu, Bing, "Sentiment Analysis and Subjectivity" (PDF). Handbook of Natural Language Processing, 2010.

[13] L. Chen, H. Lyu, T. Yang, Y. Wang, and J. Luo, “In the eyes of the beholder: Analyzing social media use of neutral and controversial terms for COVID-19,” 2020, arXiv:2004.10225. [Online]. Available: http://arxiv.org/abs/2004.10225.

[14] Li, Sijia, et al. "The impact of COVID-19 epidemic declaration on psychological consequences: a study on active Weibo users." International journal of environmental research and public health 17.6 (2020): 2032.

[15] Man Hung, Evelyn Lauren, Eric S, Wendy C Birmingham, Julie Xu, Sharon Su, Shirley D, Jungweon Park, Peter Dang, Martin S Lipsky, “Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence”, Journal of Medical Internet Research, 2020 | vol. 22 | iss. 8 | e22590.

[16] M. Munikar, S. Shakya and A. Shrestha, "Fine-grained Sentiment Classification using BERT," ArXiv, 2019.

[17] Mehmood, Yasir; Balakrishnan, Vimala "An enhanced lexicon-based approach for sentiment analysis: a case study on illegal immigration". Online Information Review. 1097–1117. doi:10.1108/OIR-10-2018-0295 ISSN 1468-4527, 2020-01-01.

[18] Mihalcea, Rada; Banea, Carmen; Wiebe, Janyce, "Learning Multilingual Subjective Language via Cross-Lingual Projections" (PDF). Proceedings of the Association for Computational Linguistics (ACL). pp. 976–983. Archived from the original (PDF) on 2010-07-08.

[19] Pang, Bo; Lee, Lillian, "4.1.2 Subjectivity Detection and Opinion Identification". Opinion Mining and Sentiment Analysis. Now Publishers Inc.2008.

[20] Pang, Bo; Lee, Lillian, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts". Proceedings of the Association for Computational Linguistics (ACL). pp. 271–278, 2004.

[21] Priyanka Tyagi, Sudeshna Chakraborty, R.C Tripathi, Tanupriya Choudhury, "Literature Review of Sentiment Analysis Techniques for Microblogging Site", ICAESMT-19.
[22] Quirk, Randolph; Greenbaum, Sidney; Geoffrey, Leech; Jan, Svartvik, A Comprehensive Grammar of the English Language (General Grammar). Longman. pp. 175–239. ISBN 1933108312. Pang, Bo; Lee, Lillian, "Opinion Mining and Sentiment Analysis". Foundations and Trends in Information Retrieval. 2 (1–2): 1–135. doi:10.1561/1500000011. ISSN 1554-0669, 2008-07-06.
[23] Rajput, Nikhil Kumar, Bhavya Ahuja Grover, and Vipin Kumar Rathi. "Word frequency and sentiment analysis of twitter messages during Coronavirus pandemic." arXiv preprint arXiv:2004.03925 (2020).
[24] Riloff, Ellen; Wiebe, Janyce, "Learning extraction patterns for subjective expressions". Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing. EMNLP '03. USA: Association for Computational Linguistics. 10: 105–112. doi:10.3115/1119355.1119369. S2CID 6541910, 2003-07-11.
[25] Ribeiro, Filipe Nunes; Araujo, Matheus, "A Benchmark Comparison of State-of-the-Practice Sentiment Analysis Methods". Transactions on Embedded Computing Systems, 2010.Samuel, Jim, et al. "Covid-19 public sentiment insights and machine learning for tweets classification." Information 11.6 (2020): 314.
[26] Sohini Sengupta, Sareeta Mugde, Garima Sharma, "An Exploration of Impact of COVID 19 on mental health -Analysis of tweets using Natural Language Processing techniques", doi: https://doi.org/10.1101/2020.07.30.20165571.
[27] Su, Fangzhong; Markert, Katja, "From Words to Senses: a Case Study in Subjectivity Recognition" (PDF). Proceedings of Coling 2008, Manchester, UK.
[28] Thelwall, Mike; Buckley, Kevan; Paltoglou, Georgios; Cai, Di; Kappas, Arvid, "Sentiment strength detection in short informal text". Journal of the American Society for Information Science and Technology. 61 (12): 2544–2558. CiteSeerX 10.1.1.278.3863. doi:10.1002/asi.21416, 2010.
[29] Thelen, Michael; Riloff, Ellen, "A bootstrapping method for learning semantic lexicons using extraction pattern contexts". Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10. EMNLP '02. USA: Association for Computational Linguistics. 214–221. doi:10.3115/1118693.1118721. S2CID 137155.
[30] Venkata Satya Sai, Abhishikth Tholana, "Literature Review on Sentiment Analysis", International Journal of Advance Research, Ideas and Innovations in Technology, Volume3, Issue5.
[31] Wiebe, Janyce; Riloff, Ellen, Gelbukh, Alexander, "Creating Subjective and Objective Sentence Classifiers from Unannotated Texts". Computational Linguistics and Intelligent Text Processing. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. 3406: 486–497. doi:10.1007/978-3-540-30586-6_53. ISBN 978-3-540-30586-6, 2005.
[32] Wiebe, Janyce; Riloff, Ellen, "Finding Mutual Benefit between Subjectivity Analysis and Information Extraction". IEEE Transactions on Affective Computing. (4): 175–191. doi:10.1109/T-AFFC.2011.19. ISSN 1949-3045. S2CID 16820846, July, 2011.