



Opinion Analysis and Statistics about the Success Rates of Treatments Extracted From Online Healthcare Forums for the Disease Psoriasis using a Pipeline of LDA, NLP and Artificial Neural Network

Mamatha Balipa, Balasubramani R

Abstract: This paper describes how the authors have applied NLP and text processing techniques to mine the messages related to the disease Psoriasis available online and posted by users who have undergone various types of treatments for the disease. The authors discuss the various classification techniques applied to classify the extracted messages as solution for the disease or non-solution. The authors also provide statistics regarding the success rates of the various types of treatments and medications in a consolidated form. A comparison study of the different types of treatments and their success rates is performed and statistics provided to the end users.

Index Terms: Artificial Neural Network, NLP, Psoriasis, LDA.

I. INTRODUCTION

The disease Psoriasis is an auto immune disease. The solution for the disease is very rare. Many people have tried various medications and treatments. The treatments people have undergone are Allopathic, Homeopathy, Ayurveda, etc., People who have undergone different types of treatments have shared their experiences online on the web on various healthcare forums. People who are interested in finding different types of treatments for the disease and their success rates online, have to wade through different types of comments. This work extracts the messages about successful treatments posted on online health forums. This work by no means recommends any medication or treatment. The work only sifts through the different types of treatments undergone by people for the disease and gives a consolidated output containing statistics about the success rates of different types of treatments and medications. The messages or comments posted on healthcare forums are in English, hence to analyse the text, Natural Language Processing techniques are applied.

Revised Manuscript Received on 30 July 2019.

* Correspondence Author

Mamatha Balipa*, Department of MCA, NMAM Institute of Technology, Nitte, Karnataka, India.

Dr. Balasubramani R, Department of Information Science and Engineering, NMAM Institute of Technology, Nitte, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

To classify the comments as solutions or non-solutions, machine learning algorithms like Naive Bayes, Decision Trees, Support Vector Machine, and Logistic Regression were explored. Nave Bayes provided an accuracy of 94%, Decision Tree provided an accuracy of 88%, SVM provided an accuracy of 88%, Artificial Neural Network provided a mean score of 99%. Convolutional Neural Network is also used and an accuracy of 84% was achieved because of the lack of volume of data.

II. LITERATURE

The popularity of the social media is growing and they provide opportunities to study interactions among humans and their experiences. In the last ten years there is an increase in the consumption of health information available online. Online health information act as a reservoir of information for researchers in the field of healthcare too. Pew Internet reports in 2011 in The Social Life of Health Information, that 80% of the people who have used the internet have referred a website for health related information [18] and 59% of people among them have searched for information related to certain medical issues. An online survey conducted by Angus Reid on 1,010 Canadian people who were randomly selected, showed that 89% of the people consulted the web to research health related issues and symptoms [5]. Similarly, it was found by Pew Research Internet Project that in 2013, 59% of Americans searched for health related issues online [5]. Health related information extracted from social network sites like Twitter and Facebook have been used in research by certain organizations[5]. Hariprasad Sampath kumar et al. have extracted drug side-effects from messages posted on the web which help in post-marketing drug surveillance[17]. Leaman et al. [10] in their research work created a dictionary of terms related to adverse effects of certain drugs and used a sliding window technique to identify adverse drug reactions in messages posted by users in online health forums. Li [11] used statistical techniques to analyse messages in sites where reviews on drugs are posted, to identify relations between certain class of drugs and some of the health disorders, which can be backed by existing research literature.

Similarly Wu et al.[20] created UDWarning, that identifies side effects of drugs that were not recognized earlier. They have used co-occurrence statistics to identify the relevance of a web page that has messages posted about the side effects of various medications. Liu et al. [13] proposed a framework called AZDrugMiner that used statistical techniques to retrieve side effects of various medications from messages posted online by patients. Chee et al.[3] used natural language processing to perform sentiment analysis and they used a combination of classifiers to find posts on drugs that are monitored by the FDA. They used messages posted by users on Health and Wellness Yahoo Groups. Bian et al.[1] used Support Vector Machine(SVM) for classification and UMLS meta thesaurus to analyze textual and semantic features for the purpose of mining adverse drug reactions. Yang et al.[23] used Proportional Reporting Ratios and association rule mining to find relations between the ADRs(Adverse Drug Reactions) and the drugs used from user messages posted on health forums. V.G.Vinod Vydiswaran, PhD, et all in their research, proposed a pattern-based approach to mine the text from Wikipedia for identifying relations between commonly used terms for a health issue by laymen and terms for the same used by professionals.[19]. In the current work the authors have used a pipeline of NLP techniques and Artificial Neural Network to classify the messages posted online about various treatments and medications available for the disease Psoriasis. The process of text classification involves the following steps. [4]

- Text pre-processing.
- Feature extraction.
- Developing a model by using various classification techniques.
- Using the data prepared for training, train the model.
- Test the model using test data.
- The common algorithms used for text classification are
 - Naive Bayes
 - Decision Trees
 - K-Nearest Neighbour
 - SVM

K-Nearest Neighbour and SVM algorithms were utilized by Nihar Ranjan et al[16] to classify text as belonging to the categories politics, sport and art. The authors investigate the utilization of Support Vector Machines (SVMs) for analyzing text data and recognize the suitability of using SVMs for the job. Quoc Le et al, proposed a framework, Paragraph Vector which is an unsupervised framework for learning continuous and distributed vectors that represent chunks of text. Max Entropy algorithm was used by Kaufmann [8] to identify similar sentences between any two languages using limited training data. SVMs were utilized by Li and Li [12] to classify sentiment polarity. They proposed that the subjectivity of opinions and credibility of exresser should also be considered for analyzing the polarity of sentences. Moraes and Valiati [15] provided an evaluation of Artificial Neural Networks and SVM for sentiment analysis performed on entire documents. They have also performed various experiments to evaluate the performance of various

supervised methods and concluded that ANN provided better results. Yan and Bing [22] explored a graph-based approach incorporating a propagation approach to utilize the outside and inside features of a sentence. Ko and Seo [9] put forth a technique that utilizes categorized keywords and sentence similarity measures to categorize sentences in a document. Xianghua and Guo [21] also used unsupervised methods to automatically identify sentiments and aspects discussed in reviews posted on Chinese social media. They used LDA model for topic detection in social reviews and applied sentiment analysis using a sliding window approach on the review text. The data set they used contained reviews extracted from blogs(2000-SINA) and a lexicon (300-SINA Hownet) was aslo used. Victoria Bobicev [2] applied machine learning techniques for classifying sentiments in forum texts that are labelled with different sentiment labels and achieved an F-measure of 0.805.

III. ABOUT THE WORK

In this work, messages extracted from health forums are extracted and classified as treatments that have worked and comments that do not mention successful treatments. So the issue is to group a remark as an answer for the sickness Psoriasis or as not an answer. In this work, Artificial Neural Network algorithm is applied for classification. Two layers of neurons (1 concealed layer) and a "pack of words" way are utilized to deal with arranging the training data. Classification of text can be achieved using various techniques. Some of them are:

- Pattern matching
- Algorithms
- Neural nets.

Though the algorithmic approach like Naive Bayes, SVM and Logistic Regression are effective in classifying text, they have their drawbacks. The algorithms create a score as opposed to probability. We need a probability to overlook predictions under a certain threshold. The algorithms get trained about comments belonging to a class. But they do not learn about the text that do not belong to the class. Sometimes it is important to identify text that do not belong to the class. The algorithms are sometimes forced to adjust the scores relative to the class size due to incorrect classification scores got by large training sets that belong to a particular class. Similarly as with Naive Bayes algorithm, the classifier developed isn't endeavouring to comprehend the meaning of a sentence, it's attempting to classify it.

Steps involved in the work carried out:

- 1) Information retrieval
- 2) Prepare the training data
- 3) Develop the ANN Algorithm
- 4) Test the results
- 5) Tune the model
- 6) Iterate
- 7) Abstract

A. Information retrieval:

Information retrieval was performed using crawlers developed by the authors to extract messages from sources like psoriasis-association.org.uk, healingwell.com, MedHelp.org [7] and HealthBoards.com [7]. The search engine to do the same was developed using JSoup API[6], a Java HTML parser library and Apache Lucene[14]. About 10000 posts were collected

From psoriasis-association.org.uk, healingwell.com, MedHelp.org [7] and HealthBoards.com [7]. The Text processing part of the system was used to extract the text messages from the document collection. The text extracted was pre-processed and then transformed into a form which can then be used by the information extraction part of the system. The information extraction part of the system uses a pipeline of NLP techniques to process the text. Since the text is in English language, the concepts of NLP is utilized in developing the classifier. Python and NLTK APIs have been used in the work. The comments are first divided into training and test text. The training text are manually labelled as belonging to pos and neg classes. A database of the types of treatments and medications are maintained. The comments labelled as pos are comments that specify a treatment that has worked. Comments labelled as neg are those comments that may either be another query regarding the disease or a discussion about a treatment or medicine that has not worked. Healingwell.com, MedHelp.org [7] and HealthBoards.com.com [7] are health forums having multiple threads discussing issues regarding Psoriasis. Users discuss treatments they have undergone that have not worked, treatments that have worked, post questions, food that aggravate the symptoms or are the cause for the disease, food that give relief from the symptoms and all the issues pertaining to the disease. A message may consist of single sentence or multiple sentences. Since the messages are free flow of text in English, the text needs to be transformed into a form which can be processed and since the messages are in English language, techniques of NLP are required to mine meaningful information from the web pages. Since the data used is extracted from online healthcare forums, the system utilizes the features of Big Data. Healthcare message boards available online provide huge volume of latest and raw data which can be used to mine useful information. In the first step, that is information retrieval, a search engine was developed that will search and download all the pages from the web pertaining to the disease Psoriasis. To check the relevance of the page, a threshold value for the count of occurrence of the word Psoriasis in the page is maintained. The search engine was developed using Apache Lucene and Jsoup API. Using JSoup API, individual comments from the online users in the page are extracted and a corpus of text containing the comments is created. Topic detection: It is ensured that the topic of discussion in the page is about Psoriasis and some of the medicines in the database by applying the Latent Dirichlet Allocation (LDA) model. For this, first the text is normalized by eliminating stop words, punctuation symbols and lemmatizing the text. A term dictionary of the text and Document Term Matrix is created. Finally the topic of discussion in the text is arrived at by applying the LDA model

on the document term matrix. Extraction of Features: The comments belonging to the entire corpus are categorized as solutions as well as non solutions. An algorithm is developed and implemented to extract features that identifies a particular text as solution or treatment for the disease Psoriasis. The algorithm works as follows. First the corpus of comments is read. The corpus has comments that suggest solutions as well as comments that are not solutions. The comments present in the corpus are manually categorized as solutions and non-solutions. All the text present in the solution category are extracted. Using regular expression all the alphabets, digits and exclamatory marks are extracted from the text and other unnecessary symbols removed. The text extracted using regular expression is converted to lowercase. The text is further tokenized and the list of words present in the text is retrieved. The most commonly occurring words in comments categorized as solution is found using frequency distribution. The most commonly occurring rare words are found using frequency distribution. The rare words are words not found in English dictionary. They are assumed to be medicines or treatments and added to the dictionary. Solution and non-solution messages are labelled with the medicine name or treatment name. In the dictionary the medicines are categorized according to treatment type. The comments are partitioned into data set to train the model and test set. Each comment in the training set is word tokenized and added to documents list labelled with its class name. The stem or root word[18] of each word in the document is found and is converted to lowercase. Duplicate words are eliminated. List of classes are retrieved from the documents and unique classes from the class list is prepared. The training data is finally converted into a bag of words. Each comment in the training set is converted into an array of 0s and 1s against the array of unique words in the corpus. A comment text can belong to multiple classes or none. A two layered Neural Network is developed with the following functions. A sigmoid function is utilized to normalize values and a secondary function to measure the rate of error. Repeating and modifying till the point that the error rate is acceptably low. Also a bag-of-words function is implemented to transform an input sentence into an array of 0s and 1s. This is also to transform the training data. A neural network training function is developed to create synaptic weights. In this function ten hidden neurons are created. Gradient descent(alpha) value is assigned 1, epochs are 50000, dropout is False, Dropout percent is 0.5. Seed is 1 and last mean error is assigned 1. The weights are randomly initialized with mean 0. The synaptic weights are found, each time checking the error rates for every 10000 iterations. Once the synaptic weights are ready the neural network model can be built. The Gradient descent parameter, alpha helps in finding the lowest error rate.

$$synapse_{0+} = alpha * synapse_{0}weight_{update} \quad (1)$$

Once the synaptic weights are ready the neural network model can be built. The Gradient descent parameter, alpha helps in finding the lowest error rate.



$$synapse_0+ = \alpha * synapse_0weight_{update} \quad (2)$$

Twenty neurons are used in the hidden layer which can be adjusted. The parameters will vary depending on the dimensions and shape of the training data. They are tuned down to $\sim 10^{-3}$ as a reasonable error rate. The synaptic weights are saved into a file which is the model created. Once the synapse weights have been calculated, the function to classify the comments as solutions or non-solutions is created. The probability of a comments belonging to either of the classes can now be generated. The error threshold is assigned as 0.2. The synapses stored in the file is loaded and used for classification in the function. The neural net learns from non-matching word also. A low-probability classification is easily shown by providing a sentence where a (common word).

B. Results:

The results are shown in the form of statistics and graphs in a web page. The results are also read out using espeak.

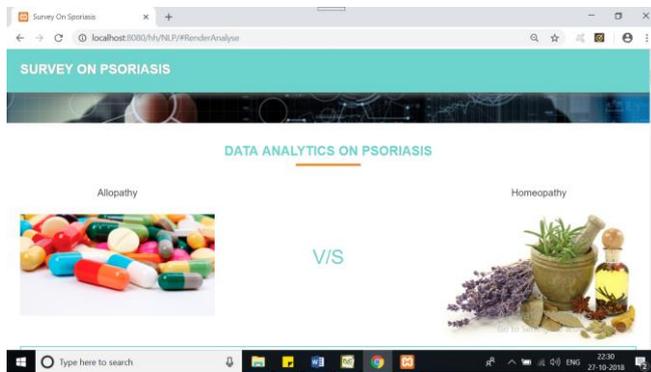


Fig. 1. Data Analytics on Psoriasis: Allopathy V/S Homeopathy

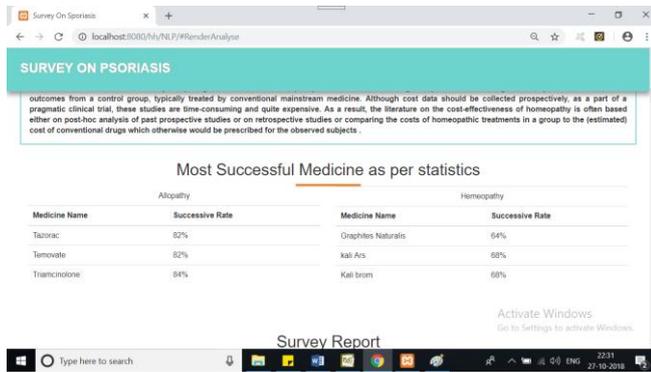


Fig. 2. Most Successful Medicine as per statistics: Allopathy V/S Homeopathy

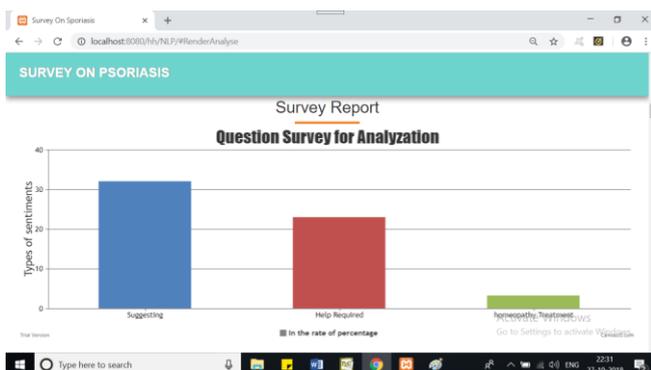


Fig. 3. Survey Report

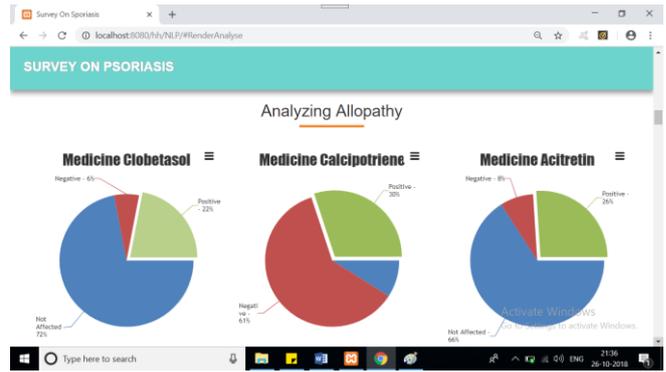


Fig. 4. Analyzing the effect of Allopathy Medicines

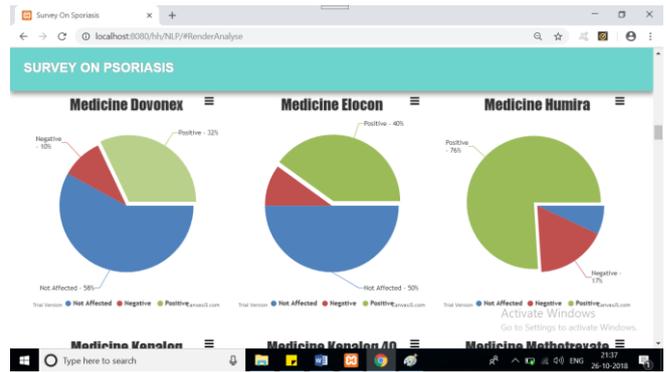


Fig. 5. Analyzing the effect of Allopathy Medicines

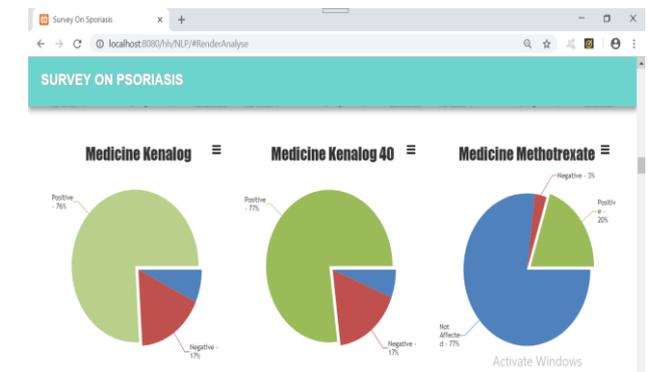


Fig. 6. Analyzing the effect of Allopathy Medicines

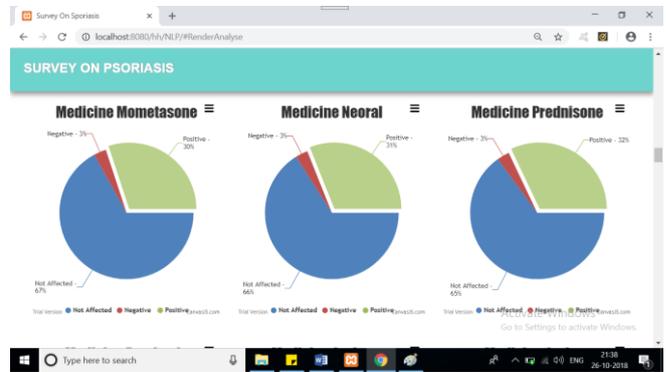


Fig. 7. Analyzing the effect of Allopathy Medicines

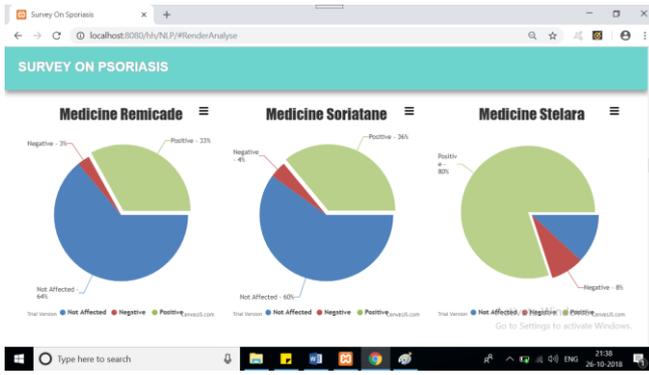


Fig. 8. Analyzing the effect of Allopathy Medicines

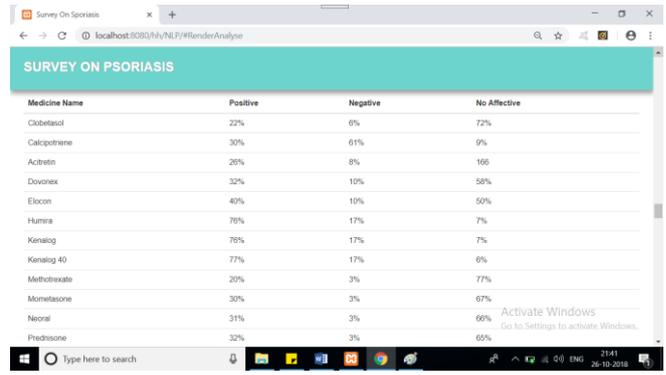


Fig. 12. Allopathy Medicines' Statistics

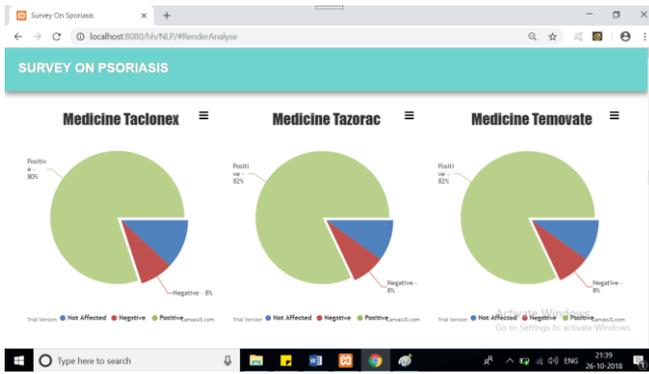


Fig. 9. Analyzing the effect of Allopathy Medicines

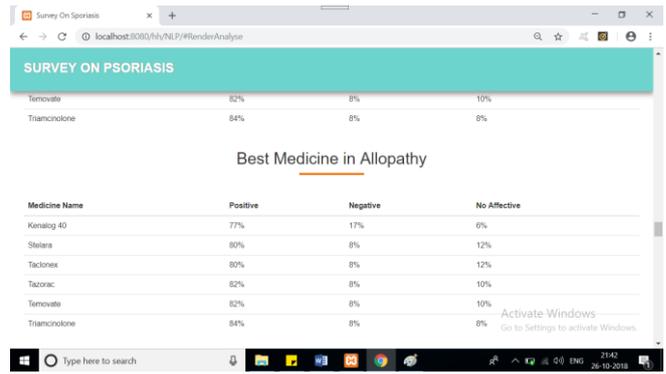


Fig. 13. Best Medicine in Allopathy

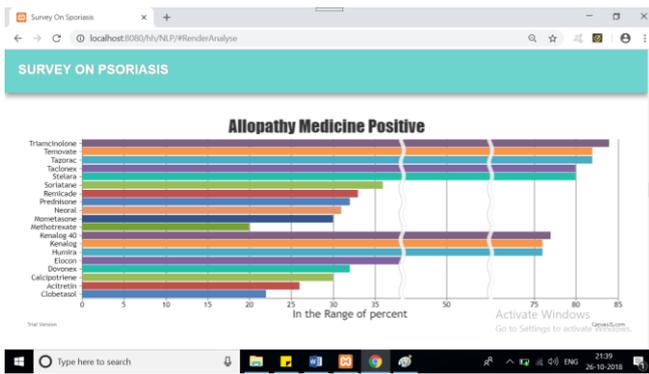


Fig. 10. Allopathy Medicine Positive

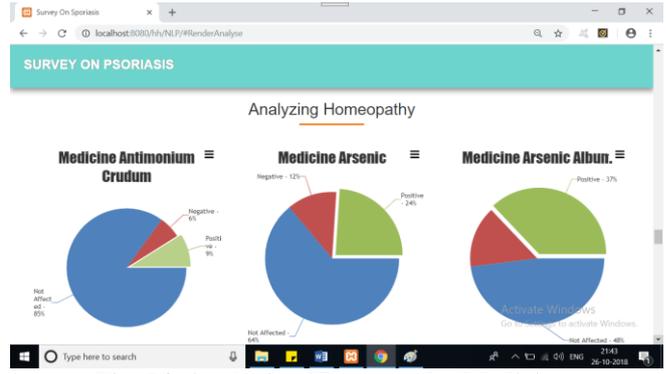


Fig. 14. Analyzing Homeopathy Medicines

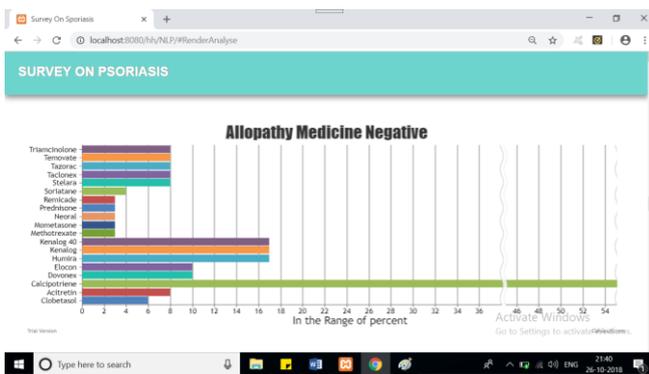


Fig. 11. Allopathy Medicine Negative

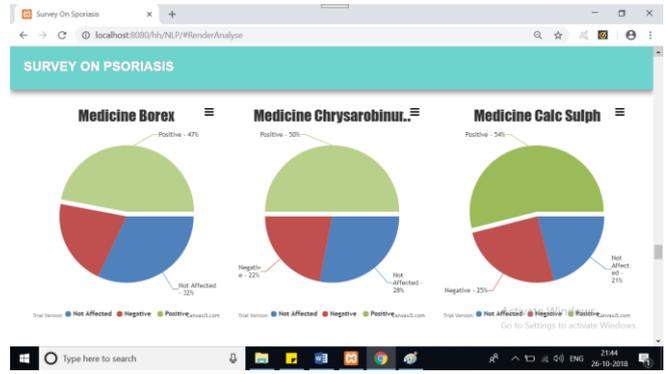


Fig. 15. Analyzing Homeopathy Medicines

Opinion Analysis and Statistics about the Success Rates of Treatments Extracted From Online Healthcare Forums for the Disease Psoriasis using a Pipeline of LDA, NLP and Artificial Neural Network

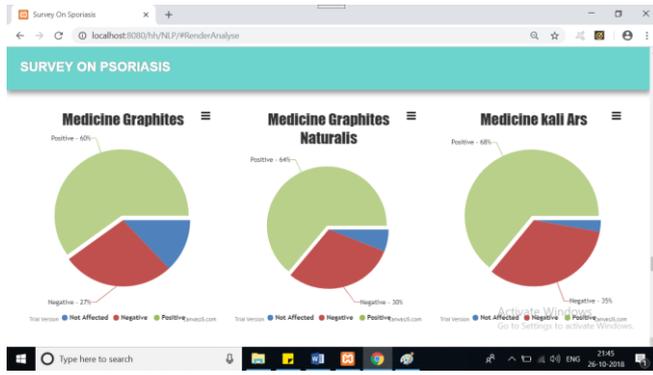


Fig. 16. Analyzing Homeopathy Medicines

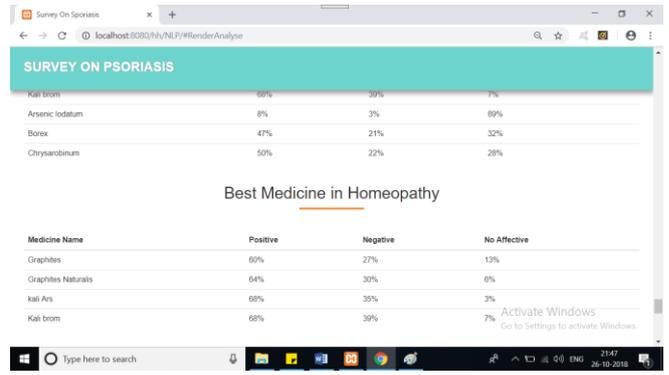


Fig. 20. Best Medicine in Homeopathy

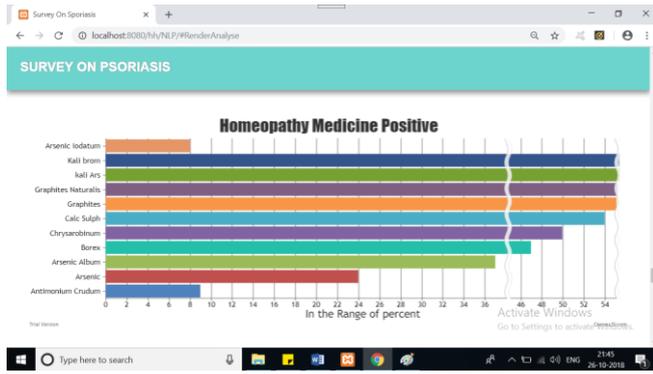


Fig. 17. Homeopathy Medicine Positive

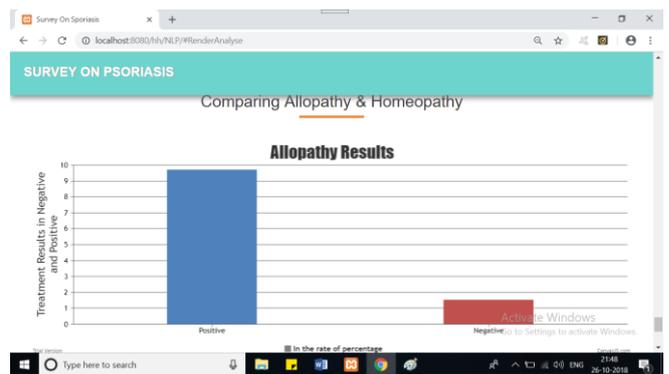


Fig. 21. Allopathy Results

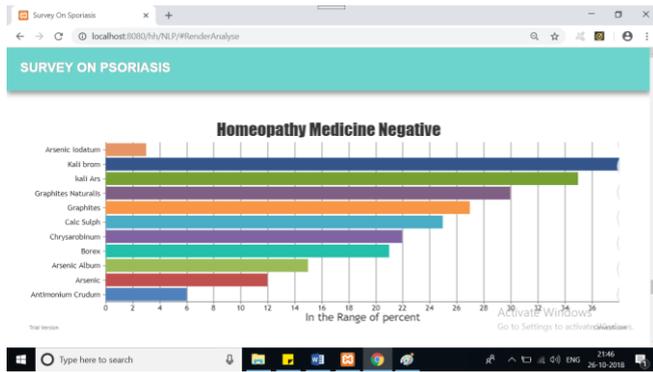


Fig. 18. Homeopathy Medicine Negative

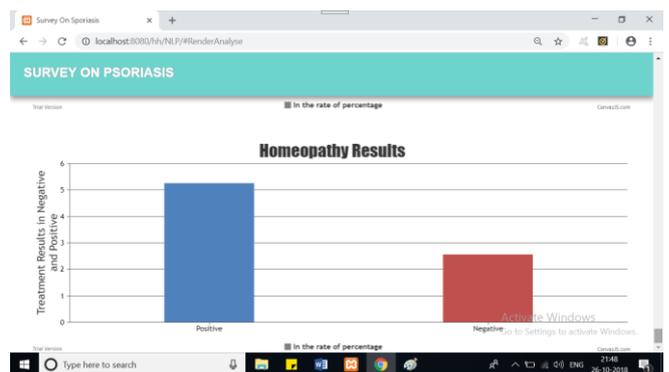


Fig. 22. Homeopathy Results

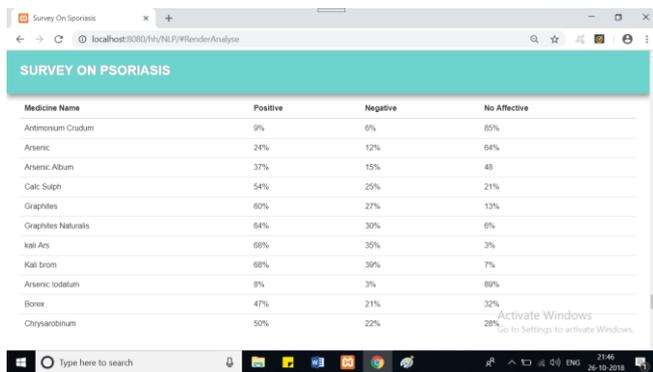


Fig. 19. Homeopathy Medicine Statistics

IV. CONCLUSION

In the work the authors have implemented a pipeline of LDA algorithm, Natural language processing, Artificial Neural Network and Frequency distribution to classify comments posted on health forums pertaining to the disease Psoriasis as comments that depict solution for the disease or not a solution. The work also computes the statistics about the success rates of the various treatments for the disease gathered by analysing the comments posted on health forums regarding the treatments for the disease. After analysing the comments on the treatments for the disease Psoriasis posted on online health forums, it is found that Allopathic medicines are found to be most effective. This may also be due to more number of people using Allopathic medicines and also discussing their effects on online health forums. The statistics are displayed in the form of text as well as graphs. This type of work is not found in the existing literature.



REFERENCES

- Jiang Bian, Umit Topaloglu, and Fan Yu. Towards large-scale twitter mining for drug-related adverse events. In Proceedings of the 2012 international workshop on Smart health and wellbeing, ACM, 2012, pages 25–32.
- Victoria Bobicev. Text classification: the case of multiple labels. In Communications (COMM), 2016 International Conference on, IEEE, 2016, pages 39–42..
- Brant W Chee, Richard Berlin, and Bruce Schatz. Predicting adverse drug events from personal health messages. In AMIA Annual Symposium Proceedings, volume 2011, American Medical Informatics Association, 2011, page 217.
- Mita K Dalal and Mukesh A Zaveri. Automatic text classification: a technical review. International Journal of Computer Applications, 28(2):37–40, 2011.
- Diana L Gustafson and Claire F Woodworth. Methodological and ethical issues in research using social media: a metamodel of human papillomavirus vaccine studies. BMC medical research methodology, , 2014, 14(1):127.
- Jonathan Hedley et al. jsoup: Java html parser, 2015. Website (<https://jsoup.org/>).
- Hyeju Jang, Sa-Kwang Song, and Sung-Hyon Myaeng. Text mining for medical documents using a hidden markov model. In AIRS, . Springer, 2006, pages 553–559.
- Max Kaufmann. Jmaxalign: A maximum entropy parallel sentence alignment tool. Proceedings of COLING 2012: Demonstration Papers, 2012pages 277–288.
- Youngjoong Ko and Jungyun Seo. Automatic text categorization by unsupervised learning. In Proceedings of the 18th conference on Computational linguistics-Volume 1, Association for Computational Linguistics, 2000, pages 453–459.
- Robert Leaman, Laura Wojtulewicz, Ryan Sullivan, Annie Skariah, Jian Yang, and Graciela Gonzalez. Towards internet-age pharmacovigilance:extracting adverse drug reactions from user posts to health-related social networks. In Proceedings of the 2010 workshop on biomedical natural language processing, Association for Computational Linguistics, 2010, pages 117–125.
- Yueyang Alice Li. Medical data mining: Improving information accessibility using online patient drug reviews. PhD thesis, Massachusetts Institute of Technology, 2011.
- Yung-Ming Li and Tsung-Ying Li. Deriving market intelligence from microblogs. Decision Support Systems, 2013, 55(1):206–217.
- Xiao Liu and Hsinchun Chen. Azdrugminer: an information extraction system for mining patient-reported adverse drug events in online patient forums. In International Conference on Smart Health, Springer, 2013, pages 134–150.
- Apache Lucene. Apache lucene core, 2013.
- Rodrigo Moraes, Jo'aO Francisco Valiati, and Wilson P Gavi'aO Neto. Document-level sentiment classification: An empirical comparison between svm and ann. Expert Systems with Applications, 2013, 40(2):621–633.
- Nihar Ranjan, Abhishek Gupta, Ishwari Dhumale, Payal Gogawale, and Rugved Gramopadhye. Full length review article. 2015.
- Hariprasad Sampathkumar, Xue-wen Chen, and Bo Luo. Mining adverse drug reactions from online healthcare forums using hidden markov model. BMC medical informatics and decision making, 201414(1):91.
- V. G Vinod Vydiswaran, Qiaozhu Mei, David A. Hanauer, and Kai Zheng. Mining consumer health vocabulary from community-generated text. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium, , 2014, 2014:1150–1159.
- VG Vinod Vydiswaran, Yang Liu, Kai Zheng, David A Hanauer, and Qiaozhu Mei. User-created groups in health forums: What makes them special? In ICWSM, 2014.
- Hao Wu, Hui Fang, and Steven J Stanhope. An early warning system for unrecognized drug side effects discovery. In Proceedings of the 21st International Conference on World Wide Web, ACM, 2012, pages 437–440.
- Fu Xianghua, Liu Guo, Guo Yanyan, and Wang Zhiqiang. Multi-aspect sentiment analysis for chinese online social reviews based on topic modeling and hownet lexicon. Knowledge-Based Systems, 2013, 37:186–195.
- Zhao Yan-Yan, Qin Bing, and Liu Ting. Integrating intra-and interdocument evidences for improving sentence sentiment classification. Acta Automatica Sinica, , 2010, 36(10):1417–1425.
- Christopher C Yang, Haodong Yang, Ling Jiang, and Mi Zhang. Social media mining for drug safety signal detection. In Proceedings of the

2012 international workshop on Smart health and wellbeing, ACM, 2012, pages 33–40.

AUTHORS PROFILE



Ms. Mamatha Balipa, Associate Professor, Department of MCA, NMAM Institute of Technology, Nitte, MCA, (Ph.D), ACM, Publications: MamathaBalipa and Balasubramani R., "Search Engine using Apache Lucene", IJCA (International Journal of Computer Applications), ISBN : 973-93-80889-58-8, DOI: 10.5120/ijca2015906476, pp 27-30, Volume 127, Number 9, 2015, Foundation of Computer Science (FCS), NY, USA., Balipa, Mamatha, and R. Balasubramani. "Search engine using apache lucene." International Journal of Computer Applications 127.9 (2015): 27-30. Balipa, Mamatha. "OPINION ANALYSIS OF TREATMENTS FROM ONLINE HEALTHCARE FORUMS FOR THE DISEASE PSORIASIS USING NLP AND ARTIFICIAL NEURAL NETWORK." Development 5.07 (2018). Balipa, Ms Mamatha, and R. Balasubramani. "Disease-Treatment Relationship Extraction for Psoriasis from Online Healthcare Forums using NLP and Classification Techniques." International Journal of Applied Engineering Research 13.6 (2018): 3568-3573. Mamatha Balipa, Dr Balasubramani R, Harolin Vaz, Christina Shilpa Jathanna, Text Summarization for Psoriasis of Text Extracted from Online Health Forums using TextRank Algorithm, International Journal of Engineering & Technology, 7 (3.34) (2018) 872-873 Mamatha Balipa, Dr. Balasubramani R, Retrieval of Disease-Treatment Information from Online Forums Using NLP and Convolved Neural Network Pipeline, International Journal of Research in Advent Technology, Vol.7 , No.1, January 2019, E-ISSN: 2321-963



Dr. Balasubramani R, Professor, EDC Head, Department of Information Science and Engineering, NMAM Institute of Technology, Nitte, Ph.D. in Content Based Imaged Retrieval for Stock Photography using Semantic and Visual Descriptors from Vinayaka Missions Research Foundation (Deemed-to-be-University) Salem in 2011, M.Tech. in Information Technology from Allahabad Agricultural Institute (Deemed-to-be-University), Allahabad in 2005. Qualified in the AICTE Deemed to be University Examination 2018. B.E. in Electronics and Communication Engineering from Madurai Kamaraj University, Madurai in 1990, Publications: "Attainment of Programme Outcomes through Course Outcomes in Outcome Based Education: A Case Study", Journal of Engineering Education Transformations (JEET) | Online ISSN: 2394-1707, Print ISSN: 2349-2473 | DOI: 10.16920/jeet/2017/v31i2/119555 |pp 26-30 | Volume: 31 | No: 02 | October 2017. personal profile which contains their education details, their publications, research work, membership, achievements, with photo that will be maximum 200-400 words. Balasubramani R., Devidas, "Enhancing Employability Skills through Infosys Campus Connect Programme: A Case Study", JEET (Journal of Engineering Education Transformations), Online ISSN: 2394-1707, Print ISSN: 2349-2473, DOI: 10.16920/jeet/2015/v29i1/77113, pp 60-66, Volume 29, Issue 1, July 2015. Balasubramani R, Manasa I P, "Multi-Level Network Scheduling Scheme for Wireless Sensor Network", IJSR-International Journal of Scientific Research, ISSN No. 2277 – 8179, pp 282-285, Volume: 3 , Issue: 5, May 2014. Balasubramani R, Kannan V, "Semantic Annotation of Stock Photography for CBIR using MPEG-7 standards", Global Journal of Computer Science & Technology 2009(GJCST 2009), Publisher: Global Journals Inc.,(USA), ISSN: 0975-4172(Online), 0975-4350(Print), pp 7-11 , Vol. 9 , Issue 5 , September 2009. Balasubramani R, Kannan V, "Semantic Search and Retrieval of Stock Photography based on MPEG-7 Descriptors", Global Journal of Computer Science & Technology 2009(GJCST 2009), Publisher: Global Journals Inc.,(USA), ISSN: 0975-4172(Online), 0975-4350(Print), pp 146-156 , Vol. 9 , Issue 5 , September 2009. Balasubramani R, Kannan V, "Efficient use of MPEG-7 Color Layout and Edge Histogram Descriptors in CBIR Systems", Global Journal of Computer Science & Technology 2009(GJCST 2009), Publisher: Global Journals Inc.,(USA), ISSN: 0975-4172(Online), 0975-4350(Print), pp 157-163 , Vol. 9 , Issue 5 , September 2009.