Exploring Sentimental Analysis of Medicine Reviews through Machine Learning Techniques

Nallamekala Sreedhar

Department of Computer Science Sri Venkateswara University, Tirupati

Abstract— *With the Internet's vast expansion, more individuals are participating in health communities like medical forums to exchange health-related information, share treatment experiences, and interact with others facing similar conditions. Recently, medical natural language processing researchers have been intrigued by the potential of monitoring social media platforms to detect various medical anomalies, including adverse drug reactions. This paper introduces a benchmark setup for analyzing sentiment regarding users' medical conditions, focusing on data from the website 'patient.info'. We've specifically explored domains like depression, anxiety, asthma, and allergies. Our goal is to identify various forms of medical sentiment expressed by users regarding their conditions, treatments, and medications. To achieve this, we've developed a deep Convolutional Neural Network (CNN)-based medical sentiment analysis system for evaluation. The resources used in this study are made accessible to the research community via LRE map for further exploration*.

I. INTRODUCTION

Sentiment analysis has gained momentum with the rise of social media, particularly in health communities like medical forums, where millions seek information, share experiences, and seek support. Medical sentiment analysis finds application in clinical record assessment and automated decision support systems for healthcare professionals. With a large portion of internet users seeking health-related information online, there's a need to extract valuable insights from the vast amount of freely available medical texts on the web. However, sentiment analysis in medical narratives remains underexplored. Machine Learning algorithms, including supervised, unsupervised, and semi-supervised methods, are commonly used for sentiment analysis, with recent advancements in deep learning showing promising results. This project focuses on applying conventional methods like Convolutional Neural Networks (CNN) in dense layers for sentiment analysis.

II. LITERATURE REVIEW

2.1 Medical Imaging Using Machine Learning and Deep Learning Algorithms: A Review

Jahanzaib Latif, Chuangbai Xiao, Azhar Imran, Shanshan Tu

Machine and deep learning algorithms are rapidly advancing in medical imaging research, aiding in disease diagnosis and early symptom prediction. Techniques like convolutional networks are specifically tailored for medical image analysis, utilizing supervised or unsupervised algorithms with standard datasets for predictions. These algorithms enhance image classification, object detection, pattern recognition, and reasoning, improving accuracy and supporting decision-making processes. This survey aims to highlight the use of machine and deep learning in medical imaging, providing an overview of existing techniques, discussing their advantages and drawbacks, and exploring future directions. These methods offer valuable tools for analyzing multi-dimensional medical data, enabling classification and automatic decision-making in disease diagnosis.

2.2 Review of Medical Decision Support and Machine-Learning Methods

Abdullah Awaysheh, Jeffrey Wilcke, François Elvinger

Machine-learning methods can assist with the medical decision-making processes at the both the clinical and diagnostic levels. In this article, we first review historical milestones and specific applications of computer-based medical decision support tools in both veterinary and human medicine. Next, we take a mechanistic look at 3 archetypal networks—commonly learning algorithms—naive Bayes, decision trees, and neural inner workings used to power these medical decision support tools. Last, we focus our discussion on the data sets used to train these algorithms and examine methods for validation, data representation, transformation, and feature selection. From this review, the reader should gain some appreciation for how these decision support tools have and can be used in medicine along with insight on their inner workings.

2.3 A Survey on Medical Diagnosis of Diabetes Using Machine Learning Techniques

Ambika Choudhury, Deepak Gupta

In medical diagnosis software design, disease prediction is a critical task, and machine learning techniques offer effective solutions. These algorithms aid in early disease detection, enhancing speed, reliability, and accuracy of diagnosis. This paper focuses on reviewing diabetes detection using machine learning techniques, employing the PIMA Indian Diabetic dataset and various algorithms such as artificial neural networks, decision trees, random forests, naïve Bayes, k-nearest neighbors, support vector machines, and logistic regression, discussing their outcomes and limitations.

2.4 Sample-Size Determination Methodologies for Machine Learning in Medical Imaging Research: A Systematic Review

Indranil Balki, Afsaneh Amirabadi, Jacob Levman

The required training sample size for a particular machine learning (ML) model applied to medical imaging data is often unknown. The purpose of this study was to provide a descriptive review of current sample-size determination methodologies in ML applied to medical imaging and to propose recommendations for future work in the field. There were only 4 studies that discussed sample-size determination methodologies, and 18 that tested the effect of sample size on model performance as part of an exploratory analysis. The observed methods could be categorized as pre hoc model-based approaches, which relied on features of the algorithm, or post hoc curve-fitting approaches requiring empirical testing to model and extrapolate algorithm performance as a function of sample size. Between studies, we observed great variability in performance testing procedures used for curve-fitting, model assessment methods, and reporting of confidence in sample sizes.

2.5 Medical Diagnosis Using Machine Learning: A Statistical Review

Kaustubh Arun Bhavsar , Jimmy Singla , Yasser D. Al-Otaibi , Oh-Young Song, , Yousaf Bin Zikria and Ali Kashif Bashir*

Medical diagnosis involves complex decision-making, prone to errors due to human limitations. Machine learning (ML) methods offer assistance in overcoming these challenges, enhancing diagnostic accuracy. This study conducts a systematic review to assess the use of ML in improving diagnosis across medical disciplines. Six databases are analyzed, employing inclusion and exclusion criteria to filter research articles. The selected articles are categorized based on various criteria, including publication year, authors, research objectives, and findings. The study highlights commonly used ML methods and diseases focused on by researchers, demonstrating the increasing adoption of ML in disease diagnosis. These insights aid in identifying neglected areas and exploring diverse applications of ML for improved diagnostic outcomes.

Problem Statement

- The researchers that use pattern recognition of these data mining methods help in predicting models based on the medical reviews.
- The experiments that were carried out using these classification-based algorithms such as Dense layer, Sequential, SoftMax.
- These results have proven to be that of dense layer technique that have performed better than the others when utilized by the techniques

Disadvantages

- The existing system used different algorithm to predict the disease, but accuracy is low comparison of our model.
- Feature Extraction is very complex.
- Training and testing the model is used same algorithm, but we provide different method.

III. PROPOSED WORK

- Our proposed system involves Dense Layer in Convolutional Neural Network (CNN) Algorithm in Deep Learning concept used to train the dataset.
- In **Dense Layer**, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers.
- In Dense Layer uses features of all complexity levels. It tends to give more smooth decision boundaries.

Advantages

- Easy detection of the medical reviews with the concluded technique.
- Time consuming.
- Best accuracy Model helps in better treatment as early.
- Detection of best Model will quick the treatment which is life saving

3.1 Sentiment analysis

Sentiment analysis (or opinion mining) is a natural language processing technique used to determine whether data is positive, negative or neutral. Sentiment analysis is often performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs.

3.1.1 Types of Sentiment Analysis:

Sentiment analysis models focus on polarity (*positive, negative, neutral*) but also on feelings and emotions (*angry, happy, sad),* urgency *(urgent, not urgent) and* even intentions (*interested v. not interested*). Depending on how you want to interpret customer feedback and queries, you can define and tailor your categories to meet your sentiment analysis needs. In the meantime, here are some of the most popular types of sentiment analysis:

3.1.2 Fine-grained Sentiment Analysis:

If polarity precision is important to your business, you might consider expanding your polarity categories to include:

- Very positive
- Positive
- Neutral
- Negative
- Very negative

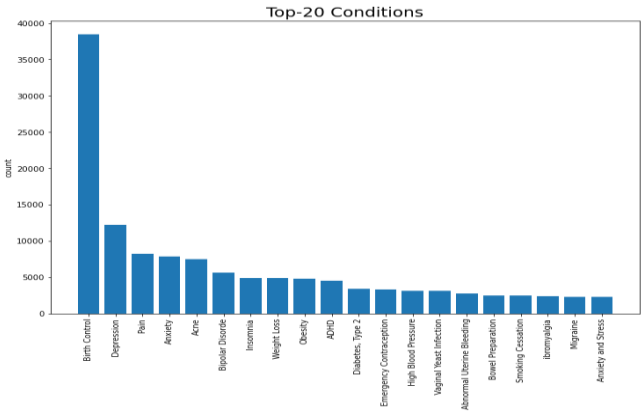
This is usually referred to as fine-grained sentiment analysis, and could be used to interpret 5-star ratings in a review, for example:

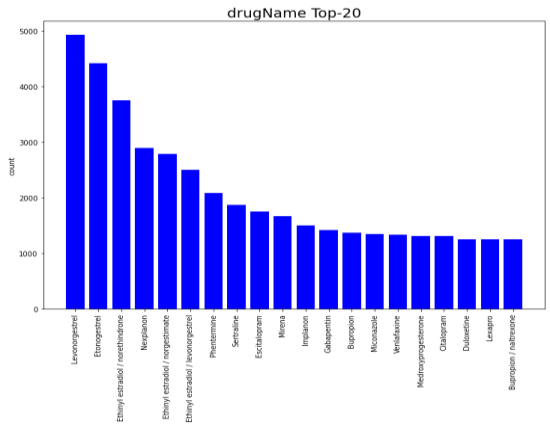
- Very Positive = 5 stars
- Very Negative = 1 star

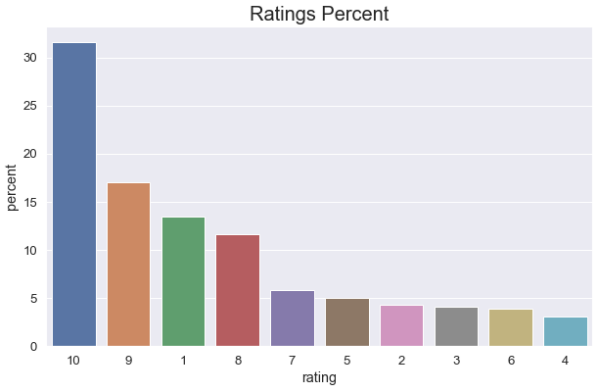
3.2 Emotion detection

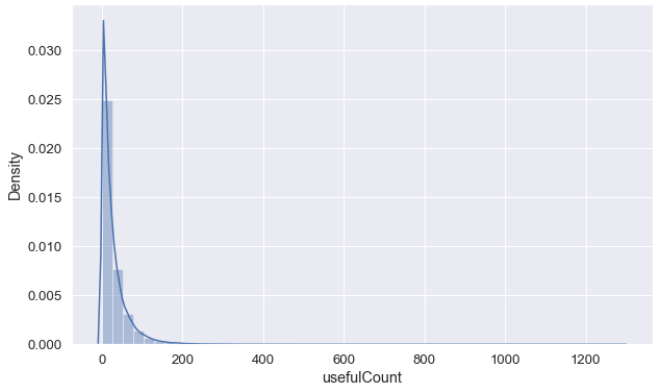
This type of sentiment analysis aims to detect emotions, like happiness, frustration, anger, sadness, and so on. Many emotion detection systems use lexicons (i.e., lists of words and the emotions they convey) or complex machine learning algorithms.

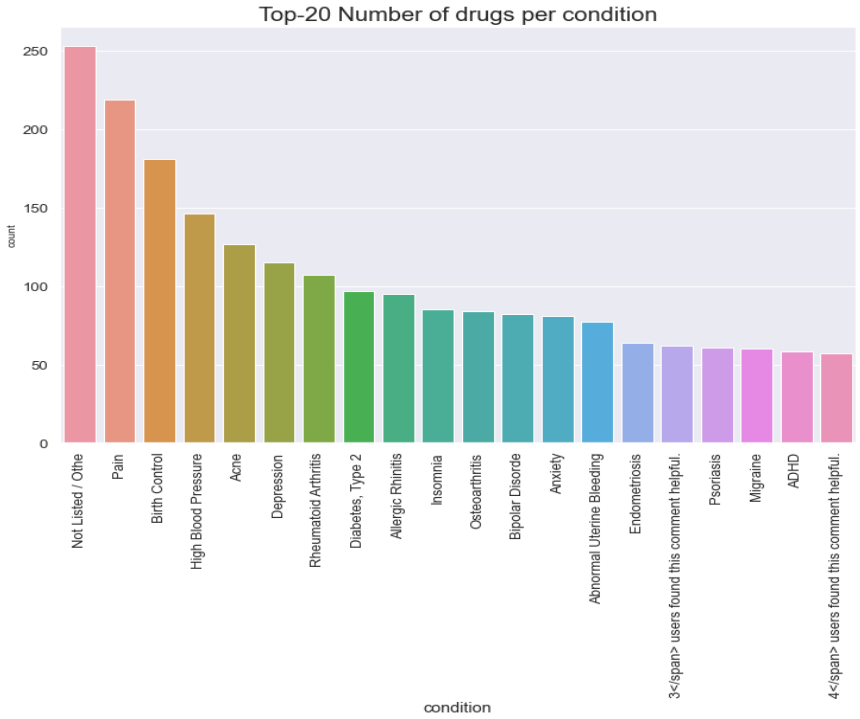
IV. IMPLEMENTATION

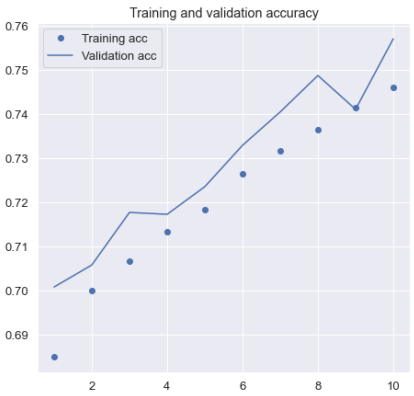












V. CONCLUSION AND FUTURE WORK

This paper conducts sentiment analysis on a drug review dataset using conventional and deep learning models, finding that deep learning models outperform conventional ones. Vector representation improves accuracy across methods. Unequal class distribution poses a challenge, leading to lower accuracy. Limited dataset size may also contribute to reduced model performance. Binary classification performs as expected, yielding higher accuracy compared to 3-class classification.

REFERENCES

[1] Adrover, Cosme, Todd Bodnar, Zhuojie Huang, Amalio Telenti, and Marcel Salathé. "Identifying adverse effects of HIV drug treatment and associated sentiments using Twitter." JMIR public health and surveillance 1, no. 2 (2015): e7.

[2] Sinnenberg, Lauren, Alison M. Buttenheim, Kevin Padrez, Christina Mancheno, Lyle Ungar, and Raina M. Merchant. "Twitter as a tool for health research: a systematic review." American journal of public health 107, no. 1 (2017): e1-e8.

[3] Latif, Jahanzaib, Chuangbai Xiao, Azhar Imran, and Shanshan Tu. "Medical imaging using machine learning and deep learning algorithms: a review." In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pp. 1-5. IEEE, 2019.

[4] Choudhury, Ambika, and Deepak Gupta. "A survey on medical diagnosis of diabetes using machine learning techniques." In Recent developments in machine learning and data analytics, pp. 67-78. Springer, Singapore, 2019.

[5] Kaur, Prabhpreet, Gurvinder Singh, and Parminder Kaur. "A review of denoising medical images using machine learning approaches." Current medical imaging 14, no. 5 (2018): 675-685.

[6] Houssein, Essam H., Marwa M. Emam, Abdelmgeid A. Ali, and Ponnuthurai Nagaratnam Suganthan. "Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review." Expert Systems with Applications (2020): 114161.

[7] Samant, Piyush, and Ravinder Agarwal. "Machine learning techniques for medical diagnosis of diabetes using iris images." Computer methods and programs in biomedicine 157 (2018): 121-128.

[8] Pillai, Rohan, Parita Oza, and Priyanka Sharma. "Review of machine learning techniques in health care." In Proceedings of ICRIC 2019, pp. 103-111. Springer, Cham, 2020.

[9] William, Wasswa, Andrew Ware, Annabella Habinka Basaza-Ejiri, and Johnes Obungoloch. "A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images." Computer methods and programs in biomedicine 164 (2018): 15-22.

[10] Islam, Md Aminul, and Nusrat Jahan. "Prediction of onset diabetes using machine learning techniques." International Journal of Computer Applications 975 (2017): 8887.

[11] Rahman, Atiqur, and Md Sharif Hossen. "Sentiment analysis on movie review data using machine learning approach." In 2019 International Conference on Bangla Speech and Language Processing (ICBSLP), pp. 1-4. IEEE, 2019.