Paper Title:

Health Education Based on Natural Language Processing(NLP) for Infectious Disease Outbreak

Paper Link:

https://ieeexplore.ieee.org/document/9534613

1 Summary

1.1 Motivation

The purpose of this paper is to look into how artificial intelligence (AI) and natural language processing (NLP) can be used to study health education connected to cases of infectious diseases. It's the goal to find out what people already know about health education sites and learn more about how to stop and handle outbreaks effectively. For better outbreak prevention and management, the study stresses how important it is to use data-driven methods in health education and how NLP and AI could be used to look at health education sources.

1.2 Contribution

The contribution of this paper is to outline the process of obtaining data from health education sources, coding and analyzing the content using NLP techniques, and constructing measurement statements based on pre-knowledge statements. Therefore, the method allows for the discovery of information that already exists and the discovery of good ways to stop and handle outbreaks. Besides, the paper shows how NLP can be used to assist in identifying, summarizing, and extracting relevant information from a large number of epidemic reports. Moreover, the paper helps validate NLP technology by demonstrating that NLP can be an effective method for obtaining pertinent data for

situational awareness and background knowledge in the field of public health by examining a particular group of epidemiological reports.

1.3 Methodology

The methodology of the paper involves four sequential steps which are data collection, piloting NLP method, coding and construction of measurement statements. Researchers used a variety of methods to look at 100 outbreak case reports from Chinese public media websites between February 1, 2021, and May 15, 2021. They used a qualitative analysis method to find important situations and pieces of information. They then created a Stanford NLP-based program that could quickly handle large amounts of pre-knowledge and situational awareness. In order to map pre-knowledge successfully, the process included keyword searches and manually loading content. A random sample was used to come up with measurement statements, and NLP technology was used to carefully extract prior information. A consensus rate of 93% was reached in the qualitative analysis validation by four inspectors. Sentences were shortened, main codes were analyzed, and secondary codes were put together into key themes for further study. The last step was to use keyword searches and an NLP tool to make measurement statements and map them to 100 reports. The goal of this thorough method was to get useful information from epidemic case reports so that a good health teaching system could be made.

1.4 Conclusion

In conclusion, this study shows how important Natural Language Processing (NLP) technology is for improving the analysis of disease reports so that a good health education system can be set up. Using NLP, the study showed that it was easier to get pre-knowledge and situational awareness from 100 outbreak case reports. This shed light on important factors that affect health education. Even though the structure of COVID-19 case files was different, NLP was very helpful in finding important information, though there are still problems with automated rapid analysis. The study stresses the need for

more research using analysis tools based on artificial intelligence to make analysis methods smarter. Some suggestions include including objective prior knowledge in epidemiological case reports and using NLP technology to make sure a full understanding.

1 Limitations

2.1 First Limitation

The paper's first limitation is that it only uses 100 epidemic case reports, which is a small sample size. This means that it might not fully show the variety and complexity of viral diseases or all the possible outcomes. Also, the paper is mostly about looking at COVID-19 case reports, so the results might not be directly applicable to other infectious illnesses. The unique features of COVID-19 data may not be typical of how diseases are reported in other situations.

2.2 Second Limitation

Another limitation is that the paper talks about the importance of NLP technology, but it also says that complicated semantic mapping needs to be done by hand for it to be perfect. The paper doesn't provide specific details about the NLP techniques used. Additionally, NLP technology cannot automatically summarize complex concepts independently.

3 Synthesis

The paper emphasizes the significance of NLP in addressing challenges in health education by analyzing epidemiological reports efficiently. The synthesis of the paper involves the utilization of NLP techniques to extract pre-knowledge and situational awareness, crucial elements in health education. The paper talks about the problems that infectious diseases cause, especially during the COVID-19 pandemic, and how smart

tools are needed to help with the study of epidemiological reports. The synthesis also involves the integration of both qualitative analysis and natural language processing to find important information in epidemic case reports. The study suggests adding objective prior information to epidemiological case reports and points out that more research could be done to make NLP analysis better.