A novel technique of fuzzy logic based association rule mining for COVID-19

A. Adhiselvam¹ and S. Agalya²

https://doi.org/10.56343/STET.116.014.001.002

http://stetjournals.com

Article History Received: 19-01-2020 Revised and Accepted : 05-05-2020 Published: 26-09-2020

Abstract

Analyzing COVID-19-related comments to discover mining and semantic ideas concerning to COVID-19 based on the public opinions of people on Reddit. Specifically extraction of COVID-19-related discussions from social media and Internet forums. It has also mobilized researchers from different sciences and different countries in the search for a way to fight this potentially fatal disease. This study analyses the abstracts of papers related to COVID-19 and coronavirus-related-research using association rule text mining in order to find the most interestingness words, on the one hand, and relationships between them on the other. Formerly, a process called information cartography, was useful for extracting structured knowledge from a huge volume of association rules. On the basis of these procedures, the purpose of our study was to show how researchers have responded in similar epidemic/pandemic situations throughout history. The role of fuzzy logic in data analysis and related fields, highlight existing contributions of fuzzy sets in these fields, and outline interesting directions for future work

Key words: association rule, information cartography, epidemic, pandemic, covid-19

INTRODUCTION

The goal of this article is to tangled on the role of fuzzy logic in data analysis, and the contributions it has made in this field. The current study associates the

🖄 A. Adhiselvam

email: adhiselvam.a@gmail.com

¹Department of Computer Applications, Sengamala Thayaar Educational Trust Women's College, Sundarakkottai, Mannargudi - 614 016, Tamil Nadu, India. Association Rule Text Mining (ARTM) (Iztok Fister Jr *et al.*,2020) method with information cartography (Iztok Fister Jr and Iztok Fister, 2020). The existing data mining method used to search for questioning and their mutual relations in the form of association rules. This method loads the describing of text documents and highlights the words that are esteemed according to the proper measures. The proposed method consists of the following steps: text preprocessing, generation of ARTM database, association rule simplification, word graph generation, metro map construction, and the exploration of extracted knowledge (Karin Fister *et al.*, 2020).

COVID 19 is an infectious illness caused by the recently identified coronavirus. It was not understood until the outbreak in Wuhan, China, started in December 2019 (Feng Shiet et al., 2020). Initially, a few cases were testified in the European Union in countries such as France and Germany, which later escalated at an alarming rate. The outbreak has spread to several cruise ships and cruise operators have activated either to cancel or to change their routes as countries around the world have introduced travel restrictions to control the spread of disease (Lu,H et al., 2020). As of 14 April 2020, 128,000 people died of COVID-19, while 1.99 million cases in 210 countries and territories were reported in 219.747 cases. The most common signs of COVID-19 include fever, tiredness, and dry cough. Most individuals (about 80percent) are healing from the disease without demanding extra treatment. Nearly 1 of every 6 people that have COVID-19 are seriously ill and experience trouble breathing. Senior citizens and those with chronic medical issues such as high blood pressure, cardiac disorders and diabetes are more prone to experience severe illnesses.

Nowadays X-ray and CT systems are equipped with cameras for patient monitoring purposes (Feng Shi *et al.*, 2020). During the outbreak of COVID-19, those

²Department of Computer Science, Sengamala Thayaar Educational Trust Women's College, Sundarakkottai, Mannargudi - 614 016, Tamil Nadu, India.

devices enable the implementation of a contactless scanning workflow. since the overhead view of the camera, it is still interesting for the technician to discover the scanning parameters such as scan range. Moreover publicly available datasets are also introduced. Finally, we discuss several open problems and challenges. We assume to provide assistance

In another solution, the author address the problem of automatically identifying new, unexpected, and interesting patterns in surveillance data. As a result, we propose a novel data mining surveillance process that is not reserved to observing for outbreaks within user-defined outcomes. Expectation of subjectively complex outcomes are characterized by association rules, and their occurrences are captured in the confidences of those rules over time. To clarify the need for automated pattern discovery and data mining in hospital infection control and public health surveillance. Next, they describe association rules, explain and how to applications can be used in surveillance, and present a novel process and system the Data Mining Surveillance(DMSS)-that utilize

Table 1. Clinical symptoms of patients with

 COVID - 19 infection

Study	Huang et al.	Chen et al.	Chung et al.
	2020	2020	2020
Fever	99%	85%	73%
Chest pain	-	5%	3%
Sore throat	-	15%	20%
Cough	77%	75%	55%
Diarrhoea	7%	25%	-
Rhinorrhoea	-	10%	30%
Confusion	-	19%	60%
Shortness of breath	75%	41%	-
Sputum production	62%	-	12%
Myalgia	8%	3%	5%
Haemoptysis	15%	-	-



Fig. 1. Total percentage of covid-19 cases against the timestamp in linear representation

patterns in surveillance data. Now the system are efficient and effective in identifying new, unexpected, and interesting patterns in surveillance data. Thus the data can be clinically significant and resourceful for the future process.

Consequently, the association rule high-order pattern discovery, states the real world data and circumstances have organized scientists from different domains around the world to attempt to find a way to control the coronavirus. These events are not only the domain of researchers in medical labs, where they are searching for a new vaccine, but all the other mass of researchers from various scientific disciplines being indirectly affected. In this, data scientists also play an key role. Recently, this methodology has been applied to understanding information in many areas.

Therefore, in-depth research and conversation on the role of "animation elements" in the COVID-19 epidemic response propaganda is of great implication.



Fig.2. Total percentage of covid-19 cases against the timestamp in logarithmic representation



Fig. 3. Total percentage of covid-19 cases against the age

In the process of image processing, computer will involve a lot of technology, such as physical modeling processing and image hardware processing technology, these technologies are also additional important in the process of computer use, the sensible use of these technologies can clearly expand the graphic design and manufacturing of the truth and aesthetic sense, to provide users with better visual feelings.

LITERATURE SURVEY

The authors in (Huang, C.L. et al., 2020) delivered another arrangement pressure calculation called Expert Model (XM) utilizing measurable strategies. In [Marr, David., 1982] the author created GenBit Compress, a calculation that packs clinical highlights and potential medicines for the COVID-19. Authors in (Huang, C.L. et al., 2020) and (Ronneberger O., Fischer P. and Brox T., 2015) give a short outline of the COVID-19 episode in addition to its clinical highlights, avoidance, finding, and treatment. The essential issue with both of these works is that they audit a subset of a lot more extensive subject. Despite the fact that these studies shed some light on the current situation of the COVID-19 flare-up, they give an extremely short and constrained thoughts regarding its specific circumstances from the technical perspectives. Authors in (Rajpurkar, P., 2017) considered ninety-nine cases of the COVID-19, forty-nine of whom had an immediate connection to the Wuhan seafood market, known to be the COVID-19 focal point. Their The authors have considered points of interest, in socioeconomics, signs and indications, and clinical history of the considerable number of patients to survey their cases cautiously. The creators have likewise introduced the research facility and discoveries of these patients to show the impacts of the SARS-CoV-2 infection on various fundamental organs of the body. In their discoveries, they report that among all the patients that were considered, 17% and 11% of cases showed intense respiratory trouble conditions and kicked the bucket of various organ brokenness disorders respectively (Dong, D., 2020). In spite of the robustness of exploration in the area of COVID-19 advancement, apparently, at the hour of this composition, there is no study that gives a thorough survey of the COVID-19 flare-up and its expected ramifications. This is one of the most utilized procedures in the health system and it is commonly entrenched. This learning method utilizes information in making precise desires and learns the mapping between the yield and its relating input while gathering analysis through the learning strategy in recognizing things reliant on near characteristics. Various ways currently used in envisioning an outcome or the future outcome or gathering to a great deal of needed states (Feng Shi, 2020). The ordinary procedures finished accurately currently is the SVM, NN, and RC. In setting up these estimations, generally, the limit is being described in the medical system. It can induce the best in association between the yield and data. By then, the cost work tells the practitioner how far it can be from the accurate output, so this goes about as a data signal. The model had the choice to arrange medical data with an exactness of up to 90%. All things considered with significant learning; it has become a jump forward framework right now in all locales. Supervised learning can assemble an astute stage for programmed observing and forecasting of COVID-19. A neural system can likewise be created to extricate the visual highlights of this ailment, and this would help in legitimate checking and treatment of the influenced people. It can give the most recent updates of the patients and answer associated queries with COVID-19 (Xu, X. et al., 2020).

K-means and also the auto encoder is the most regularly known unsupervised technique. One of the most broadly perceived jobs of this learning technique in the medical line is the quirk acknowledgment (Liu, S. et al., 2019). The data made from the affiliation will start from equivalent scattering, if there is any sort of interference or any counterfeit read from the affiliation data, this data will be hailed an exemption and can without a lot of a stretch be hailed or seen. The K-means is a common and generally utilized grouping mechanism in image processing (Jelodar, H. et al., 2019). This calculation requires the client to indicate the quantity of cluster k to be produced. The utilization of deep learning can be valuable to perceive, analyse, anticipate, and clarify the COVID-19 infection, and help in keeping up financial effects. Since the flareup of the pandemic, there has been a scramble to utilize and investigate AI, and other information-analytical instruments, for these reasons (Plangprasopchok, A. and Lerman, K., 2010). The study of single neuron receptive fields that appeared in 1959 was one of the most important papers in Computer Vision describing the central reaction properties of visual cortical neurons as well as how the sensory experience of a cat affects its cortical architecture. Lawrence Roberts described the method used to extract 3D data on solid objects from 2D images in 1963. Essentially, the external world has been distilled into flat geometric shapes. It was established in 1982 that the vision was hierarchical (Chen, J., 2019). The primary function of the vision system is to create 3D world models so that we can interact with them. A perception system was developed where low-level algorithms that could detect lines, curves, and corners were used as stepping stones to a high-level understanding of visual information. At the same time, a self-organizing



artificial network of simple and complex cells, which could recognise patterns and not be influenced by position changes, was developed (Ravneet Punia, Lucky Kumar, Mohd. Mujahid., 2020). This included several convolutional layers whose receptive fields. had weight vectors known as filters. The aim of these filters was to slide through 2D image pixel arrays and, following accurate calculations, to generate activation events that would be used as inputs for subsequent layers of the network. Text recognition and commercial zip code decoding applications have been released . In fact, this culminated in the development of the MNIST data collection with handwritten digits.

PROPOSED MODEL

DATA ANALYSIS

Data analysis is the method of learning and inspecting as well as cleansing and transforming of data to recover

useful information or recommend a solution and this method helps in making decisions for business or other processes. Other than that data scientists and analysts search for patterns in the various observations in data. In light of its viability in managing monstrous data, it turns out to be globally the most mainstream point at present.

CLUSTERING

Cluster Analysis in Data Mining means that to find out the group of objects which are similar to each other in the group but are different from the object in other groups.. The aim is to identify the objects in one group which are similar to each other (related) and different groups of objects which are different (not related). The greater similarity within the group, and the greater difference between the groups provides the better clustering. Clustering is one of the main methods of data mining, which has been widely used in pattern recognition, trend analysis, similarity search and other areas.

FUZZY CLUSTERING

Under the ontic perspective independence refers to the pairs of random fuzzy sets defined over the same probability space, and considered as random objects. For example, a constant random fuzzy set will be stochastically independent of any other random fuzzy set defined over the same probability space. The COVID-19-related statements from an online healthcare-oriented group can be measured theoretically useful for mining meaningful topics to better understand the opinions and highlight discussions of people/users and improve health strategies. The proposed system helps to recover the way and to detect the uncover meaningful topics that are being conferred on COVID-19-related problems on reddit, as major research. We propose a method for radiology to detect the novel corona, a process called information cartography, was useful for extracting structured knowledge from a huge volume of association rules. It plays an important role in delivering the best results in a shorter period of time. Here the Table 1 represents symptom percentage level of covid-19 infected patient.

The Algorithm 1 describes a general process as part of our framework for extracting and mining the data. The input data consists of the number of COVID-19–related comments as the context of the document then compute the probability of the word distribution from Topic K[x]. Compute the probability of the topic distribution from the COVID-19-Content-Document m[x].

The Algorithm 2 states that the LDA handles topics as multinomial distributions in documents and words as a probabilistic mixture of a pre-determined number from. Because the Gibbs sampling method is used in this step, the time requested for model inference can be specified as the sum of the time for inferring LDA. Therefore, the time complexity for LDA is O(NK), where N denotes the total size of the corpus (COVID-19– related comments) and K is the topic number.

RESULT AND DISCUSSIONS

COVID-19 is obviously a serious disease of international concern and more number of people have been testified to have been infected or died from it. Due to expeditious spread, countries around the world should increment attention into disease surveillance systems and scale up country readiness and replication operations including establishing expeditious replication teams and ameliorating the capacity of the national laboratory system. Perpetuated research into the virus is critical to trace the source of the outbreak and corroborate future outbreak.COVID-19 epidemic situation predicated on the computer vision availed systems is analyzed in this paper Judging from online reviews and surveys, this form of expression has been popular with the public, but there are additionally a sizably voluminous number of netizens who verbally express they do not understand the content of information and illustrations. This shows that China lacks general information visualization or information illustration. The following figures Fig. 1, Fig. 2, and Fig. 3 represent the total percentage of Covid-19 cases in linear mode, total percentage of Covid-19 cases in logarithmic mode and total percentage of cases against age respectively.

Accurate image-based disease diagnosis requires highquality image data. CT images sometimes have low contrast that may hamper the visualization of critical structures. Firstly, enhancing the visual quality of input data improves visualization of significant pathologists; secondly, it improves the performance of feature extraction and segmentation algorithms.

CONCLUSION

To investigate the analysis report, we conclude that the association deep learning classification rules to detect Association rule deep learning classification algorithm states the effort on initial management of COVID-19 must be addressed to the early recognition of the suspect and contain the disease spread by immediate isolation and infection control measures. and also classification on COVID-19-related issues from healthcare forums, such as sub-reddits. Furthermore, results aid in improving practical strategies for public health services and involvements related to COVID-19. The results of social distancing those that completely halt the spread .We believe that our paper will motivate other researchers to search out new methods for detecting potential patients infected with the virus without the specific use of medical COVID-19 test kits.

REFERENCES

- Chen, J., Wu, L., Zhang, J., Zhang, L., Gong, D., Zhao, Y., Chen, Q., Huang, S., Yang, M., Yang, X. and Hu, S., 2020. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Scientific reports, 10(1): 1-11. https://doi.org/10.1038/s41598-020-76282-0
- Dong, D., Tang, Z., Wang, S., Hui, H., Gong, L., Lu, Y., Xue, Z., Liao, H., Chen, F., Yang, F. and Jin, R., 2020. The role of imaging in the detection and management of COVID-19: a review. IEEE Reviews in Biomedical Engineering.

https://doi.org/10.1109/RBME.2020.2990959

Feng Shi, Jun Wang, Jun Shi, Ziyan Wu, Qian Wang, Zhenyu Tang, Kelei He, Yinghuan Shi, DinggangShenh,2020. Review of artificial Intelligence techniques in imaging data acquisition, segmentation and Diagnosis for

https://doi.org/10.1109/RBME.2020.2987975

COVID-19. IEEE reviews in biomedical engineering. 14:.4-15.

- Karin Fister, FisterJr, and IztokFister, 2020.Discovering associations in COVID-19 related research papers.arXiv preprint arXiv: 2004.0339.
- Huang, C.L., Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, W., Si, H.R., Zhu, Y., Li, B., and Chen, H.D., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 579 (7798): 270-273. https://doi.org/10.1038/s41586-020-2012-7
- Iztok Fister Jr and Iztok Fister, 2020. Information cartography in association rule mining. arXiv preprint arXiv: 2003.00348.
- IztokFisterJr, Suash Deb, and IztokFister, 2000. Populationbased metaheuristics for association rule text mining. arXiv preprint arXiv:2001.06517.
- Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y. and Zhao, L., 2019. Latent Dirichlet allocation (LDA) and topic modelling: models, applications, a survey. Multimedia Tools and Applications, 78(11):15169-15211. https://doi.org/10.1007/s11042-018-6894-4
- Kim, E.S., Chin, B.S., Kang, C.K., Kim, N.J., Kang, Y.M., Choi, J.P., Oh, D.H., Kim, J.H., Koh, B., Kim, S.E. and Yun, N.R., 2020. Clinical course and outcomes of patients with severe acute respiratory syndrome coronavirus 2 infection: a preliminary report of the first 28 patients from the Korean cohort study on COVID-19. Journal of Korean medical science, 35(13).
 - https://doi.org/10.3346/jkms.2020.35.e142

- Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S.X., Ni, D. and Wang, T., 2019. Deep learning in medical ultrasound analysis: a review. 5(2):261-275. https://doi.org/10.1016/i.eng.2018.11.020
- Lu, H., Stratton, C.W. and Tang, Y.W., 2020. Outbreak of Pneumonia of Unknown Etiology in Wuhan China: the Mystery and the Miracle. Journal of Medical Virology. 92(4): 401-402. https://doi.org/10.1002/jrnv.25678
- Marr, David., 1982. Vision: A computational investigation into the human representation and processing of visual information. W.H.Freeman and Company, New York.
- Plangprasopchok, A. and Lerman, K., 2010. Modeling social annotation: a bayesian approach. ACM Transactions on Knowledge Discovery from Data (TKDD), 5(1):1-32. https://doi.org/10.1145/1870096.1870100
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Meha, H., Duan, T., Ding, D., Bagul A., Langlotz, C., Shpanskaya, K., and Lungren, M.P., 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.
- Ravneet Punia, J., Lucky Kumar, Mohd. Mujahid., 2020. Computer Vision and Radiology for COVID-19 Detection.IEEE Conference for Emerging Technology. P. 1-5. https://doi.org/10.1109/INCET49848.2020.9154088
- Ronneberger O., Fischer P., and Brox T., 2015. U-net: Convolutional networks for biomedical image segmentation. International conference on medical computing and computer-assisted image intervention. Springer, Cham. P. 234-241. https://doi.org/10 1007/978-3-319-24574-4 2
- Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Ni, Q., Chen, Y., Su, J. and Lang, G., 2020.A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering, 6(10): 1122-1129. https://doi.org/10.1016/j.eng.2020.04.010