Secured Health Recommendation System With 3D-CNN Approach Towards Analysis of Medical Big Data

Gopesh Kumar Bharti[1], Deshmukh Vaishnavi Jaikumar[2]

[1]Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India.
[2]Research Scholar, Department of CS & IT, Kalinga University, Raipur, India.

All parties participating in the nursing period, especially the older patients' families, will receive healthcare recommendations as part of an integrated patient-based e-health platform that offers healthcare for the elderly in need of care. Clinical information, such as medication recommendations, test results, and therapy recommendations, however, heavily burdens medical professionals when it comes to making patient-centered decisions. A health recommendation algorithm that analyses massive amounts of data and forecasts health risk before it reaches a critical level has been developed recently by a number of researchers. The primary goal of the thesis is to concentrate on an effective approach for recommending healthcare services while protecting patient and user privacy and allowing for system data scalability. In order to accomplish this, we have created a 3D-CNN with a two-way recommendation system that will offer patients or end users safe, efficient, and effective health recommendation services.

Keywords: Health care, Decision, Recommendation system, Deep learning.

1. Introduction

The swift growth of big data has led to an increase in data mining and analytics across various industries, including trade, e-commerce, and healthcare. Healthcare systems are an integrity analysis that has arisen as a potential area associated with big data analytics. Three features of big data in healthcare—variety, volume, and velocity—can be distinguished [1]. Variety refers to the fact that the data may come from a wide range of sources in a variety of formats, including internal and external sources [3]. In Figure 1, the general organisation of HRS is displayed.

In general, when Personal Identity (PID) is concealed by data privacy approaches, the integrity of the medical information must be maintained. Additionally, the common method avoided using the person ID when information was connected by using the De-identification procedure (anonymization). In order to protect sensitive data, the medical information incorporates pseudo-identifiers (ZIP code, date of birth) and masking personal identifiers (unique ID number, name). RS are very application-focused systems with distinct goals and duties. Numerous studies have been proposed to compare current and existing algorithms from the specified range of techniques in order to assess the performance of the RS [15]. Valuation metrics that yield a numerical score or ranking measures can be used to estimate this. Therefore, tasks and objectives that are application-oriented will be the basis for the evaluation of RS. The ensuing subsections [2] provide explanations of the most well-liked and frequently applied evaluation measures.

Big data analytics is a rapidly expanding field that offers a range of options for research and commercial experts nowadays. Many big data tools are not created with security considerations in mind, endangering the privacy and data security aspects of the organizations [9]. A few of the attacks, such as ransomware, DDoS attacks, ransomware, and information threats, significantly disrupt the system's analytics and data storage. Both the internet and offline domains are the source of security assaults. Theft of sensitive data, such as credit card numbers or other financial information, is one of the main sources of the data threat. [13].

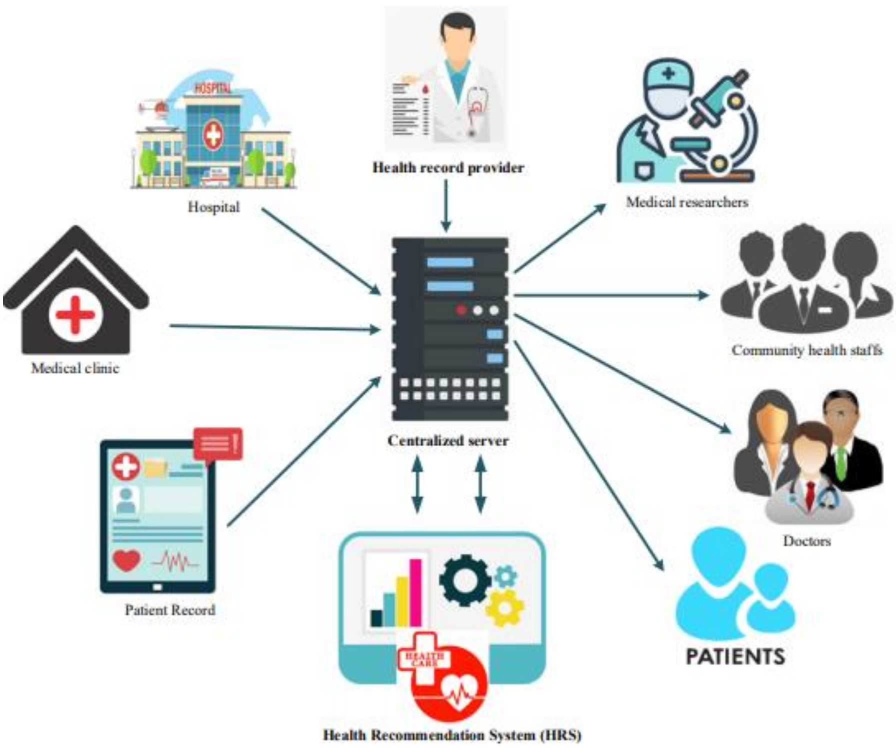

Figure 1 Overall structure of HRS

The rest of the paper is organized as follows: Section 2 provides the classification scheme for the survey; Section 3 provides an overview of proposed architecture. Section 4 provides a summary and comparison of the results of the various papers discussed in this taxonomy. Finally, Section 5 concludes the paper.

2. Related Works

The medical profession regularly generates a vast volume of health record data through software-based applications due to the advancements in technology. The information is used as the basis for several health-related applications that track and evaluate patients' medical conditions. As a result, these datasets are used by researchers and analysts to automatically identify health recommendations for patients [6]. Only by fully comprehending the patient's symptoms and overall health can better medication recommendations be made. Nevertheless, the development of a system that suggests the optimal medication to patients is jeopardised by both good and negative review comments for a single medication [4].

When it comes to managing large datasets and deriving intricate associations from a variety of clinical data sets, deep learning (DL) algorithms are the most innovative method available. Deep learning techniques are rapidly expanding in a number of fields, including computer vision and speech recognition. It is also used in applications based in academics and industry because of its capacity to solve complicated patterns. Deep learning models are currently used in the recommendation system to improve the recommender's performance even further. Deep learning model advancements have drawn a lot of attention since they address issues with traditional methods and produce recommendations of higher quality. [5].

Because the ensemble learning model performs significantly better than the single model, it has recently produced good results and attracted a lot of study attention. Boosting, bagging, and stacking are the three most used ensemble learning strategies. In order to create ensemble models for the prediction of a certain disease, researchers created a variety of models. The ensemble learning model builds an efficient prediction framework by combining different tree-based algorithms. The primary idea behind this kind of learning is that a group of ineffective models are combined to create a stronger learning system, which improves suggestion accuracy. This model lowers the variances and biases as well as other sorts of issues. [14].

The Health Recommendation System reads disease symptoms, recognises the ailment, and slows down the process of medical diagnosis because big data analytics is open source [11]. The reason for this is inaccurate assessments of patients' reluctance and the confidentiality of patient or physician health information [16]. Furthermore, not only do the methods suggest that the patient see a physician for excellent care, but the medical records kept in the database also do not safely exchange data [10]. Data corruption results from unauthorised access because of privacy concerns [12]. In order to promote the privacy-preserving measure of providing the medical record to the service provider and to precisely rate a high-quality physician for a referral, an effective recommendation system that protects privacy is thus put forth.

3. Methodologies

The exponential growth in health data available on multiple platforms over the past few decades has led to the use of recommendation systems in healthcare applications. Big data analytics tools can make patient-centered decisions using this information. Currently, people can get relevant information on various websites by searching through a vast amount of clinical data that is dispersed across the internet. Here, RS is achieved via 3D-CNN. The normalisation of data from a single batch includes the computation of the mean and variance from mini-batch data points as well as the norming of a mean and unit variance of zero. The CNN modifies its output weights and parameters by evaluating the error using a few loss functions, e (also referred to as cost functions, error), and replicating the errors using specific input rules. The loss is calculated using a partial derivative e concerning the outcomes of every neuron in a layer, such as $\partial e/yi,j,k\ \ell$ for its result, $y_{i,j,k}$ of $(i,j,k)\ th$ unit in layer ℓ. The cFhain rule allows each variable's contribution to be written and contributed separately:

1. $$\frac{\partial e}{\partial x^{\ell}_{i,j,k}} = \frac{\partial e}{\partial y^{\ell}_{i,j,k}} \frac{\partial f(y^{l}_{i,j,k})}{\partial y^{\ell}_{i,j,k}} = \frac{\partial e}{\partial y^{\ell}_{i,j,k}} f'(x^{l}_{i,j,k}) \quad (1)$$

2.

The following equation can be used to modify the weights in the previous convolution layers by repeating the error to those layers:

$$\begin{aligned} \frac{\partial e}{\partial y^{\ell-1}_{i,j,k}} &= \sum_{a=0}^{n1-1} \sum_{b=0}^{n2-1} \sum_{c=0}^{n3-1} \frac{\partial e}{\partial x'_{(i-a),(j-b),Uk-c)}} \frac{\partial x^{\ell}_{(i-a)(j-b)(j-b)}}{\partial y^{\ell-1}_{i,j,k}} \\ &= \sum_{a=0}^{n1-1} \sum_{a=0}^{nz-1} \sum_{b=0}^{n3-1} \frac{\partial e}{\partial x^{T}_{(i-a)(j-b),(k-c)}} \omega_{a,b,c^*} \end{aligned} \quad (2)$$

The equation above can be used to calculate the error for the previous layer. Moreover, the equation above is valid for n locations on each side of the data. By adding nulls to each side of the input volume, this problem can be eliminated [7].

4. Results and Discussion

The MATLAB 2020b platform is used to simulate the suggested 3D-CNN model. The recommendation system is constructed using several configuration parameters for performance evaluation, which are derived from the size of the dataset and values found in current literature. Because a proper set of parameters will yield positive outcomes in terms of processing speed, accuracy, and security level, among other things. The parameters that were chosen have specified values in order to assess the suggested work using the sources of each dataset that are now accessible.

We employed big data in our efforts to recommend healthcare. Benchmarked datasets do not contain big data sources. So, for our study, we used datasets that we had acquired ourselves. Metrics including precision, recall, performance efficiency, and F1-measure are used to assess the performance. The dataset is split into two parts for assessment purposes: 75% of it is used for training and 15% is used for testing. The remaining 10% is utilised to assess the RS's accuracy.

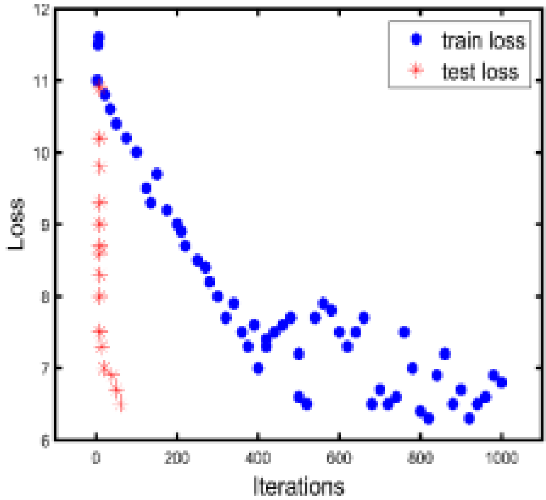

Figure 2 Performance efficiency comparison

When comparing 3D-CNN's performance efficiency to that of the other methods now in use (BIAM, DCNN, PSVM, CNN, and LSTM), Figure 2 illustrates the results [8–12]. Sensitive medical information about the patient or user is insecure with the BIAM technique because of inadequate security. In order to address this problem, the 3D-CNN methodology, a privacy-preserving health recommendation system, was created. It is an effective way to get both high recommendation accuracy and a high degree of security between patients and healthcare professionals. Because of loss, the other approaches do not reach convergence more quickly. However, 3D-CNN converges more quickly, demonstrating a high level of performance.

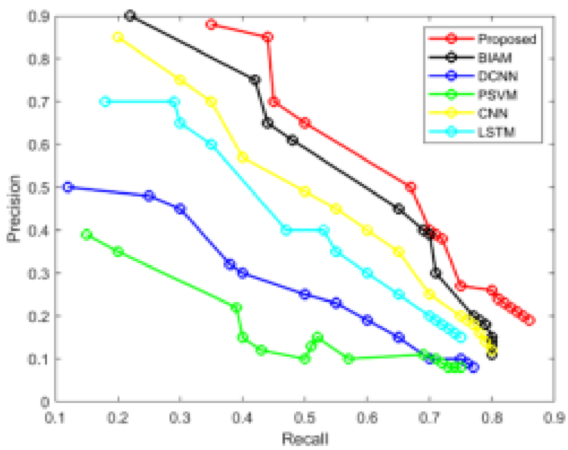

Figure 3 Precision-Recall comparison curve for different existing methods

Figure 3 makes it evident that the 3D-CNN outperforms the other techniques (BIAM, DCNN, PSVM, CNN, and LSTM) in terms of precision and recall values. Greater precision means the RS provides better recommendations relating to health. Higher recall scores also indicate that the recommendations made to the user are predicated on their needs. The 3D-CNN methodology outperforms the previous methods in terms of precision and recall by utilising self-collected datasets.

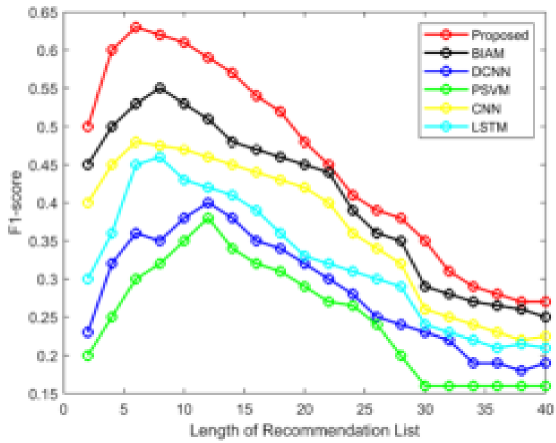

Figure 4 F1-score comparison for various methods

Figure 4 makes it evident that the 3D-CNN outperforms the other techniques now in use (BIAM, DCNN, PSVM, CNN, and LSTM) in terms of F1-score. The user's optimal recommendation list is constructed, as indicated by the high f1-score. The precision and recall numbers are combined to get the F1 score. On a wide number of recommendation lists, the 3D-CNN approach produces a higher F1-score with variable numbers of iterations, ranging from 0 to 40.

5. Conclusions

The user receives health-based recommendations from this work. Based on their request, the system produces the health recommendation. The user's prior input is gathered, and the recommendations are derived from this information. Recommendations for health are based on both explicit and implicit data. The 3D-CNN learns the features for offering recommendations for medical care using this data. In this case, the data is analysed twice before being combined to produce suggestions. In the future Using the updated method, the user's data is extremely secure during this process. The user's privacy is therefore protected. Additionally, the system offers precise and pertinent recommendations.

References

1. Sahu, Hemlata, Ramgopal Kashyap, Surbhi Bhatia, Bhupesh Kumar Dewangan, Nora A. Alkhaldi, Samson Anosh Babu, and Senthilkumar Mohan. "Analysis of Deep Learning Methods for Healthcare Sector-Medical Imaging Disease Detection." Contemporary Mathematics (2023): 830-852.
2. Yar, Hikmat, Naveed Abbas, Tariq Sadad, and Sajid Iqbal. "Lung nodule detection and classification using 2D and 3D convolution neural networks (CNNs)." Artificial Intelligence

and Internet of Things (2021): 365-386.
3. S. Neelima, Manoj Govindaraj, Dr.K. Subramani, Ahmed ALkhayyat, & Dr. Chippy Mohan. (2024). Factors Influencing Data Utilization and Performance of Health Management Information Systems: A Case Study. Indian Journal of Information Sources and Services, 14(2), 146–152. https://doi.org/10.51983/ijiss-2024.14.2.21
4. Faruqui, Nuruzzaman, Mohammad Abu Yousuf, Faris A. Kateb, Md Abdul Hamid, and Muhammad Mostafa Monowar. "Healthcare As a Service (HAAS): CNN-based cloud computing model for ubiquitous access to lung cancer diagnosis." Heliyon 9, no. 11 (2023).
5. Siripurapu, Sridhar, Naresh K. Darimireddy, Abdellah Chehri, B. Sridhar, and A. V. Paramkusam. "Technological advancements and elucidation gadgets for Healthcare applications: An exhaustive methodological review-part-I (AI, big data, block chain, open-source technologies, and cloud Computing)." Electronics 12, no. 3 (2023): 750.
6. Sonya, A., & Kavitha, G. (2022). A Data Integrity and Security Approach for Health Care Data in Cloud Environment. Journal of Internet Services and Information Security, 12(4), 246-256.
7. Gupta, Meenu, Akash Gupta, and Gaganjot Kaur. "Deep learning for medical healthcare: Issues, challenges, and opportunities." Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications (2021): 361-392.
8. Manimegalai, P., R. Suresh Kumar, Prajoona Valsalan, R. Dhanagopal, P. T. Vasanth Raj, and Jerome Christhudass. "[Retracted] 3D Convolutional Neural Network Framework with Deep Learning for Nuclear Medicine." Scanning 2022, no. 1 (2022): 9640177.
9. Mohamed, K.N.R., Nijaguna, G.S., Pushpa, Dayanand, L.N., Naga, R.M., & Zameer, AA. (2024). A Comprehensive Approach to a Hybrid Blockchain Framework for Multimedia Data Processing and Analysis in IoT-Healthcare. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 15(2), 94-108. https://doi.org/10.58346/JOWUA.2024.I2.007
10. Sun, Hanqing, Zheng Liu, Guizhi Wang, Weimin Lian, and Jun Ma. "Intelligent analysis of medical big data based on deep learning." IEEE Access 7 (2019): 142022-142037.
11. Singh, Satya P., Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan, and Balázs Gulyás. "3D deep learning on medical images: a review." Sensors 20, no. 18 (2020): 5097.
12. Ludhiani, Simran, Rajesh Kumar Maheshwari, and . 2022. NOVEL APPLICATION OF MIXED SOLVENCY CONCEPT TO DEVELOP AND FORMULATE LIQUISOLID SYSTEM OF A POORLY WATER SOLUBLE DRUG, FUROSEMIDE AND THEIR EVALUATIONS. International Journal of Pharmacy Research & Technology, 12 (1), 28-57. doi:10.31838/ijprt/12.01.05
13. Bobir, A.O., Askariy, M., Otabek, Y.Y., Nodir, R.K., Rakhima, A., Zukhra, Z.Y., Sherzod, A.A. (2024). Utilizing Deep Learning and the Internet of Things to Monitor the Health of Aquatic Ecosystems to Conserve Biodiversity. Natural and Engineering Sciences, 9(1), 72-83.
14. Saba Raoof, Syed, and MA Saleem Durai. "A comprehensive review on smart health care: applications, paradigms, and challenges with case studies." Contrast Media & Molecular Imaging 2022, no. 1 (2022): 4822235.
15. Yan, Fei, Nianqiao Li, Abdullah M. Iliyasu, Ahmed S. Salama, and Kaoru Hirota. "Insights into security and privacy issues in smart healthcare systems based on medical images." Journal of Information Security and Applications 78 (2023): 103621.
16. Pham, Chi-Hieu, Thien Huynh-The, Ehsan Sedgh-Gooya, Marwa El-Bouz, and Ayman Alfalou. "Extension of physical activity recognition with 3D CNN using encrypted multiple sensory data to federated learning based on multi-key homomorphic encryption." Computer Methods and Programs in Biomedicine 243 (2024): 107854.
17. Xhafa, Fatos. Machine Learning, Big Data, and IoT for Medical Informatics. Academic Press, 2021.