International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 8, No. 4, 2021, pp. 24-28.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Multi Disease Prediction Using Deep Learning Framework for Electric Health Record

Vijay Kumar [a], Meera Shah [b]

[a] *Department of Civil Engineering, Structure School of Infrastructure, Bilaspur, India.*
[b] *Department of Civil Engineering, Structure School of Infrastructure, Bilaspur, India.*

Abstract

Information mining for medical services is an interdisciplinary field of study that started in data set measurements and is helpful in looking at the adequacy of clinical treatments. For minor side effects, the trouble is to meet the specialists whenever in the medical clinic. Hence, huge information gives fundamental information with respect to the disease based on the patient's side effects. Thereby, prediction multiple disease with respect to effective energy technology as deep learning. The stages of this paper's effective multi-disease prediction are as follows: a) Gathering data from disease repositories, such as those for diabetes, heart disease, Alzheimer's disease, and pancreatic cancer; b) Preprocessing using grey scaling, histogram equalization, and data normalization; c) Using a convolution autoencoder to extract features; d) using particle swarm optimization (PSO) to select features; and e) using a recurrent neural network (RNN) for prediction. Several state-of-the-art models are evaluated experimentally using a variety of metrics, and the suggested model performs better (Diabetes; precision: 0.95; heart disease; precision: 0.93; Alzheimer's disease; precision: 0.9; pancreatic cancer; precision: 0.93).

Keywords: Alzheimer's, Deep Learning, Deep Neural Network, Multi-disease Prediction, Machine Learning, Particle Swarm Optimization.

DOI:10.71086/IAJSE/V8I4/IAJSE0827
Received: 09 September 2021/Revised: 11 October 2021/Accepted: 08 November 2021/Published: 31 December 2021

Introduction

People's living habits and the surroundings are gradually changing as civilization advances, which gradually raises the hidden risks of many diseases. Globally, major diseases like cancer, diabetes, heart disease, vision impairment, and brain health have a significant influence. In terms of diabetes alone, 422 million individuals worldwide suffer from the disease, with Type 2 diabetes making up over 90% of those affected (Li et al., 2020). The chance of developing heart illnesses rises with age due to cardiac senescence and function loss. Over 30% of deaths worldwide are attributable to heart disease (Landi et al., 2020). These serious illnesses have a negative impact on people's health and productivity. In addition, it will raise medical and health care costs and intensify social pressure. Predicting the likelihood that a person will have a particular disease in the future is the primary goal of disease prediction. When considering various diseases in various populations, a lot of affecting factors must be considered. Comprising a collection of characteristics with incredibly broad dimensions that must be identified, as well as intricate and changing individual and illness differences that must be differentiated. Attempting to complete these duties by hand is not only difficult but also expensive and time-consuming.

Technology has advanced as a result of societal growth. It facilitates the gathering of medical data sets. The amount of health-related data being collected by medical institutions worldwide is continuously growing. Basic patient data, electronic health records (EHR) (Qiu et al., 2021), electronic medical records (EMR) (Komal Kumar & Vigneswari 2019), picture data, medical instrument data, and more are examples of medical data categories. Despite being part of a single category of basic facts, they are very susceptible to influence. For instance, data that is redundant or missing due to timeliness or personal security concerns renders the information false and raises privacy exposure concerns. Furthermore, medical data is more complicated, high-dimensional, heterogeneous, irregular, and comprises more unstructured data than normal data. These differences make it challenging to handle medical data merely by hand. Furthermore, a number of factors, including medical conditions, the diagnostician's skill level, and patient variances, still limit medical diagnosis in practice (Wang et al., 2019).

System Model and Assumptions

The overall architecture of the proposed system is shown in Fig. 2, which includes the following stages: a) Data collection from popular repositories for the diseases this paper will address, such as Alzheimer's disease (Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset), heart disease (heart disease data set UCL), diabetes (Pima Indian diabetes (PID)), and pancreatic cancer (Peking Union Medical College Hospital (PUMCH)) is necessary for any deep learning framework (Xie et al., 2021). Once these images are collected, they are undergone for b) preprocessing in which these raw images will be undergone technique like gray scaling for gray conversion, then histogram equalization for improving the quality and finally Normalization technique as well. Then followed by c) feature extraction using Autoencoder for extracting quintessential features from those inputs, then d) feature selection using PSO for converting the extracted features to vector forms for e) classification stage. This stage will used to predict the disease using RNN. For training deep learning model, sufficient data input is necessary, so this section describes the 4 dataset which required for Multi-disease prediction using deep learning model (Men et al., 2021; Ampavathi & Saradhi 2021).

As of July 1, 2016, there were 7128 longitudinal T1 MR pictures from 1723 participants in the ADNI archive. Since ADNI-2 is a research project, data is constantly expanding. Consequently, only a subset of 1674 subjects (5074 pictures) were processed during a large-scale segmentation analysis (into 138 anatomical sections using MALP-EM; Ledig et al., 2015), from which the subset utilized here was chosen Fig. 1.

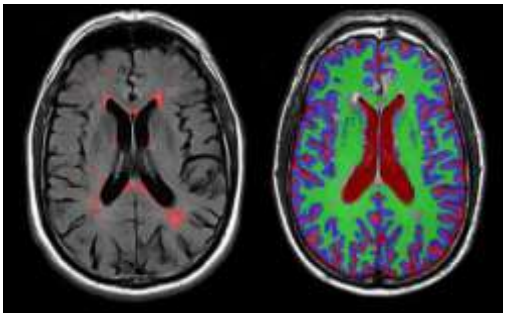


Fig. 1: Alzheimer's Dataset Instance

The model is implemented using hardware specifications such as the Ryzen 5/6 series CPU, 1 TB HDD, NV GPU, and Windows 10 OS, as well as software specifications such as Google Collaboratory, an open source Google environment for developing deep learning models, and Pytorch, an open source Python library for building deep learning models. The four diseases are assessed using a variety of metrics and cutting-edge models.

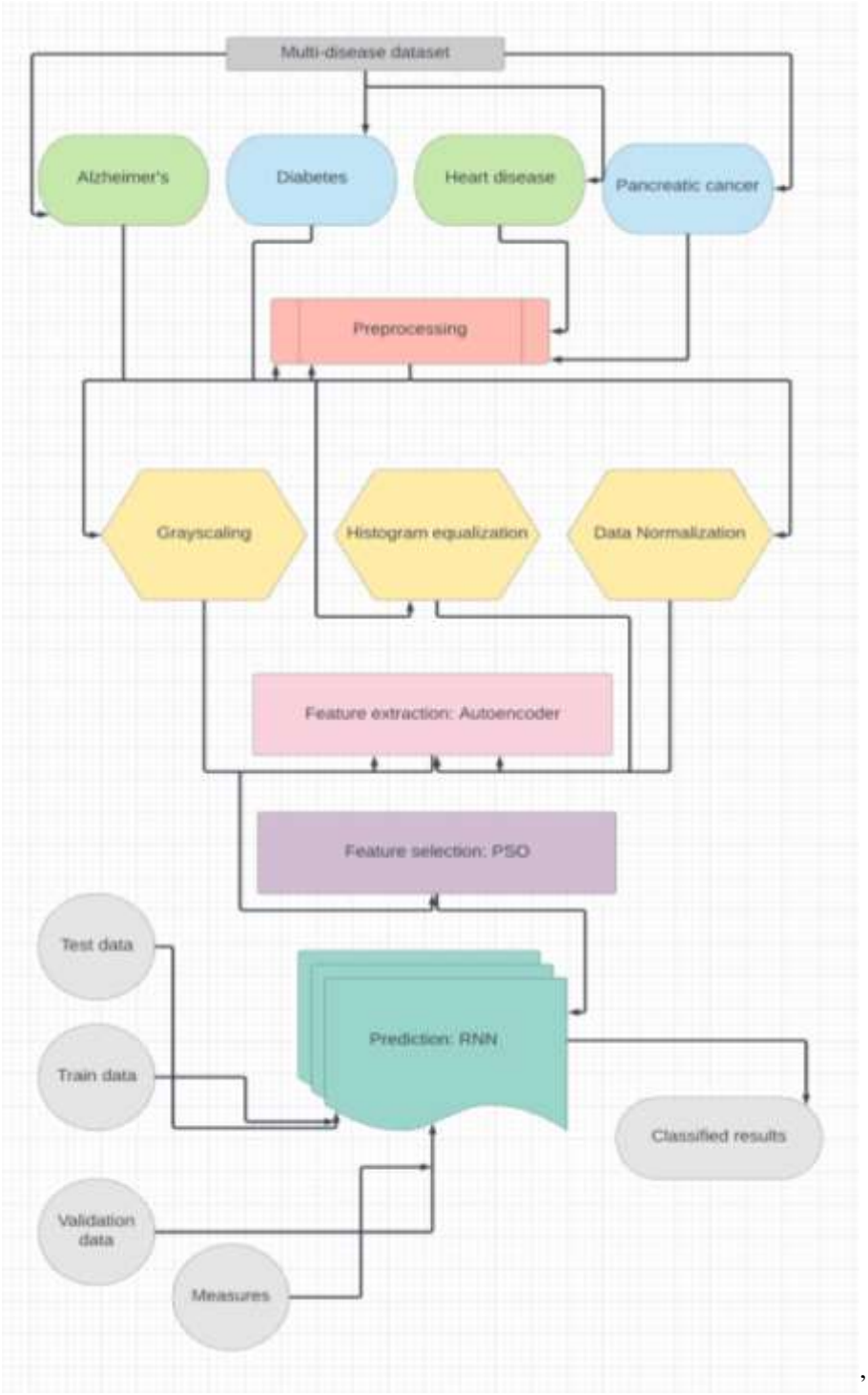


Fig. 2: Overall Architecture of Proposed Framework

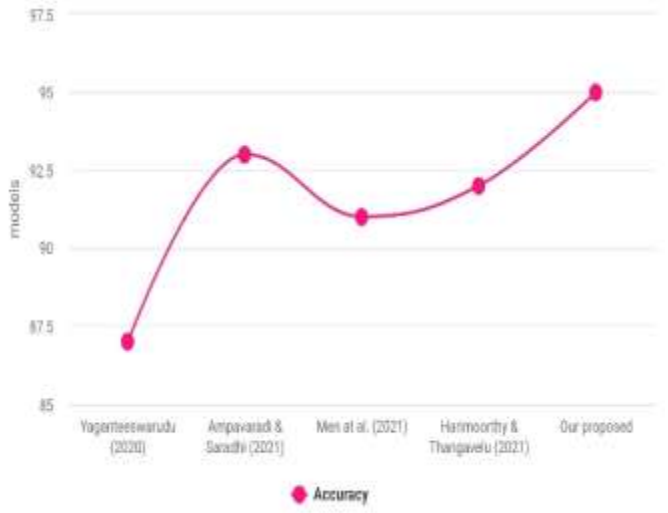
.

Experiments

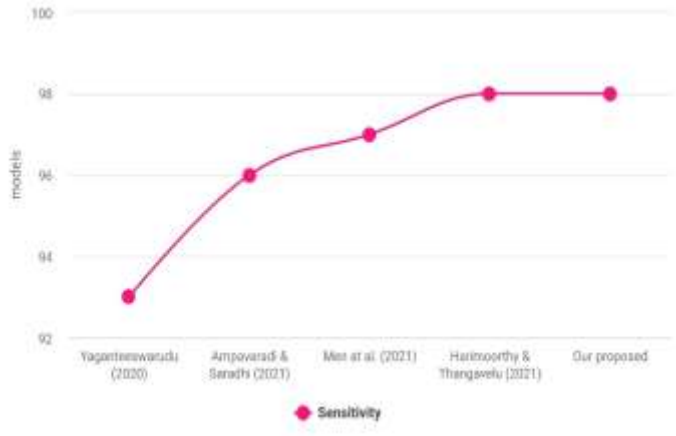
Table 1 depicts the performance analysis under accuracy, sensitivity, specificity. Fig. 3(a,b,c) depict overall analysis of state-of-the-art models and proposed model.

Table 1: Overall Analysis under Alzheimer's

Models	**Accuracy**	**Sensitivity**	**Specificity**
Yaganteeswarudu (2020)	87	93	96
Ampavaradi & Saradhi (2021)	93	96	95
Men at al. (2021)	91	97	96
Harimoorthy & Thangavelu (2021)	92	98	96
Our proposed	95	98	97



(a)



(b)



(c)

Fig. 3: Models' vs a) Accuracy, b) Sensitivity, c) Specificity.

Conclusion

An efficient deep learning framework for multi-disease prediction using RNN is presented in this research, along with a brief overview of each stage. Deep learning's potential as a new technology is obviously growing, and the shape that researchers are taking to act upon it is exciting. Mostly, using this technology in the medical field will be helpful. This study presents a multi-disease prediction system that can forecast conditions such as diabetes, Alzheimer's, heart disease, and pancreatic disease. It also uses effective feature extraction and selection to improve model performance. Additionally, other research specialists will employ this study to delve deeply, comprehend different techniques, and attempt to introduce new integrated models.

References

Li, Y., Rao, S., Solares, J. R. A., Hassaine, A., Ramakrishnan, R., Canoy, D., & Salimi-Khorshidi, G. (2020). BEHRT: transformer for electronic health records. *Scientific reports*, *10*(1), 7155. https://doi.org/10.1038/s41598-020-62922-y

Landi, I., Glicksberg, B. S., Lee, H. C., Cherng, S., Landi, G., Danieletto, M., & Miotto, R. (2020). Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ digital medicine*, *3*(1), 96. https://doi.org/10.1038/s41746-020-0301-z

Qiu, L., Gorantla, S., Rajan, V., & Tan, B. C. (2021). Multi-disease predictive analytics: A clinical knowledge-aware approach. *ACM Transactions on Management Information Systems (TMIS)*, *12*(3), 1-34. https://doi.org/10.1145/3447942

Komal Kumar, N., & Vigneswari, D. (2019, December). A drug recommendation system for multi-disease in health care using machine learning. In *International Conference on Advanced Communication and Computational Technology* (pp. 1-12). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-15-5341-7_1

Wang, T., Tian, Y., & Qiu, R. G. (2019). Long short-term memory recurrent neural networks for multiple diseases risk prediction by leveraging longitudinal medical records. *IEEE journal of biomedical and health informatics*, *24*(8), 2337-2346. https://doi.org/10.1109/JBHI.2019.2962366

Xie, S., Yu, Z., & Lv, Z. (2021). Multi-disease prediction based on deep learning: a survey. *Computer Modeling in Engineering & Sciences*, *128*(2), 489-522. https://doi.org/10.32604/cmes.2021.016728

Men, L., Ilk, N., Tang, X., & Liu, Y. (2021). Multi-disease prediction using LSTM recurrent neural networks. *Expert Systems with Applications*, *177*, 114905. https://doi.org/10.1016/j.eswa.2021.114905

Ampavathi, A., & Saradhi, T. V. (2021). Multi disease-prediction framework using hybrid deep learning: an optimal prediction model. *Computer Methods in Biomechanics and Biomedical Engineering*, *24*(10), 1146-1168. https://doi.org/10.1080/10255842.2020.1869726