Deep Learning models for Healthcare Prediction

Ms. Ghousiya Farheen Shaikh Mukhtar[1], Professor. M. P. Dhore[2]

[1] Research Scholar, Department of Computer Science,
Shivaji Science College, Nagpur.

[2] Pro-VC, Sant Gadge Baba University,
Sant Gadge Baba University, Amravati.

Abstract: In recent years, healthcare prediction has played a significant role in saving lives. Intelligent technologies for evaluating complex data linkages and turning them into accurate information for use in the prediction process are developing quickly in the field of health care. Artificial intelligence is rapidly changing the healthcare sector. As a result, systems that rely on deep learning and machine learning are becoming more important in the development of steps that diagnose and predict diseases. These systems can be based on images or clinical data and offer significant clinical support by simulating human perception, including the ability to diagnose diseases that are difficult for humans to detect. Challenges exist for the healthcare sector in critical areas such as data integration, computer-aided diagnosis and illness prediction, and electronic record administration. The trend towards more personalized treatment and lower healthcare expenditures are essential. The fast-developing domains of predictive analytics and deep learning have begun to significantly influence the development of large-scale healthcare data practices and research. Deep learning provides a multitude of methods, strategies, and structures to deal with these issues. Predictive analytics for health data is becoming a revolutionary tool that can make treatment choices more pro-active and preventive. The framework for deep learning making is the main emphasis of this research. It shows the study of various deep learning tools and techniques in practice as well as the applications of deep learning in healthcare.

Keywords: Predictive Analytics, Machine learning, Deep learning, Artificial intelligence.

1. Introduction

Healthcare prediction is the process of analyzing historical and current healthcare data to predict future events, trends, or outcomes. It can help healthcare organizations make better decisions, improve patient care, and reduce costs. The process of converting observational data into official disease nomenclature is known as illness diagnosis. Predictive analytics in healthcare plays a major role in improving care delivery and patient outcomes. By leveraging historical data, this type of analytics allows health systems to gauge what is likely to happen in the future, both from an operational and clinical perspective. The healthcare sector is producing enormous amounts of new data sets because of the digitization of healthcare. To mention a few, imaging equipment, doctors' notes, and computerized physician order entry are possible sources of clinical data. Compared to records from other industries, these datasets are especially complicated and dispersed. Compared to other industries, these datasets are especially complicated and fragmented, which creates enormous challenges for diagnosis, treatment, and prevention. Improving these databases would be of immeasurable worth. Predictive analytics uses a variety of methods, including machine learning, artificial intelligence, and deep learning, to help the life sciences in healthcare.

Deep learning is a key technique in big data analysis because it uses massive volumes of data to automatically find patterns and extract features from complicated unstructured data without the need for human intervention. With the help of pertinent clinical questions and deep learning techniques, clinically relevant information can be unearthed from large data sets to support clinical decision-making. This allows doctors to make better medical decisions by accurately diagnosing diseases and tailoring treatments. This paper summarizes the Models and tools of deep learning for predictive analysis in the health sector.

2. Methodology

Deep Learning algorithms

Deep learning algorithms are a subset of machine learning techniques that use neural networks with many layers to model and understand complex patterns in data. These algorithms have achieved significant success in various domains such as image recognition, natural language processing, and game playing. Here's an overview of some prominent deep learning algorithms and architectures:

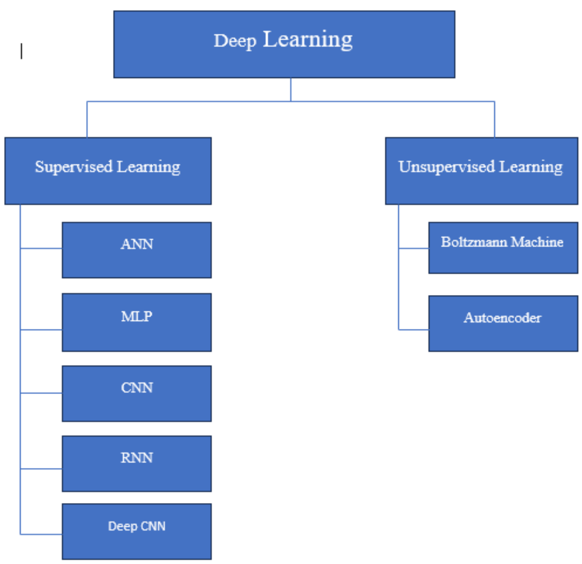

Fig 1: Hierarchical representation of deep learning.

A) Supervised learning Models:

Supervised learning is a fundamental approach in machine learning where a model is trained using a dataset consisting of input-output pairs. The primary goal is for the model to learn the mapping from inputs to outputs so that it can make accurate predictions on new, unseen data.

1) Multi-layer Perception (MLP)

Multi-layer perception is a neural network (branch of machine learning) extension. It is seen as an initial step towards the mathematical representation of the biological process. The human brain simulates every neurone to analyze information and produce possible actions. Neurons in the biological process change in response to simulation characteristics, develop new neural connections, and undergo self-transformation. Three layers make up the MLP architecture: the input layer is connected to the output nodes, while the intermediate layer is hidden. The main difference between MLP and ANN is that MLP has several hidden layers, and biological processes are carried out with the help of the activation function and a small number of given weights. When this design is used, data usually flows in a single direction because the network can only have a certain number of hidden layers.

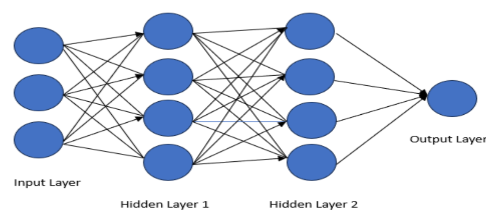

Fig 2: Simple multi-layer perception model.

2) Recurrent neural networks (RNN)

Since a neural network can analyze data from all network states, RNNs are seen to be a suitable option when applying deep learning approaches to sequentially ordered data. Traditionally, RNNs have a directed graph with temporal behavior made up of connections between hidden layers. it is helpful, when the trained data contain strong interdependencies and the output must depend on every previous condition(states). the working of RNNs, hidden state at the time 't' is sequentially updated not only by the initiation of the current input state at the same time. Also, on the shrouded state at time 't-1', which is revised by the activation of input state at 't-1' and the hidden state at the time 't-2' and so on. In summary, RNNs are said to have memory when they use two sources of input, i.e. current and recent back. One major benefit of employing RNNs is that, in contrast to MLP, all the steps in an RNN have the same weights. This means that the total number of parameters the network needs to learn is reduced. Speech, text, and DNA sequences are a few real-world applications of RNNs.

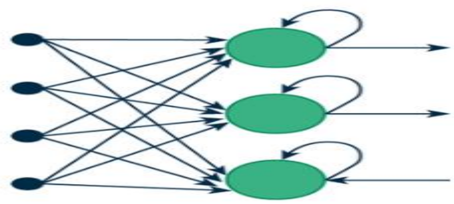

Fig:3: Recurrent neural network

3) Convolutional neural networks (CNN)

It is not possible to apply neural network models that have been explored thus far to multi-dimensional correlated data. For example, the number of nodes and parameters become almost impossible when imaging data is uploaded as training data. A CNN architecture utilizing the convolutional filter mask is suggested. It employs locally linked neurons and data-specific kernels for learning as opposed to predetermined kernels. Since CNN is specifically suggested for the analysis of image data, the filter mask is repeatedly applied to the entire image. The resulting connectivity appears as a series of overlapping fields. Backpropagation is employed in many CNN applications in the field of neurobiology. The architecture must adapt all parameters to a specific instance of the mask, which decreases the links from the architectures.

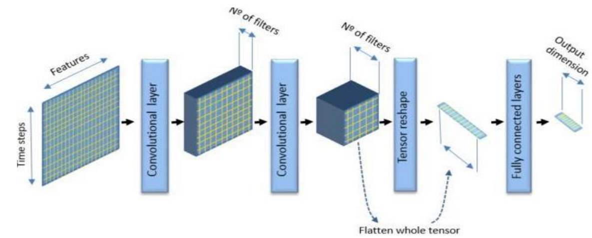

Fig 4: Convolutional neural networks

4) Deep Conventional Neural Network

Deep CNN is a typical feedforward neural network, it uses the Backpropagation (BP) algorithm to modify network parameters (weights and biases) to lower the value of the cost function. There are four innovative ideas that set it apart from the typical BP network: shared weight, local receptive field, grouping, and different layers combinations. Deep CNNs are made to handle data using network known as topology It is often used to classify sensor data, diagnose trends in time series data, and recognize objects in photos. A special kind of artificial neural network (ANN) known as "deep conventional neural networks" substitutes convolution for standard matrix multiplication in at least one of its layers instead of regular matrix multiplication. It has drawn a lot of attention in the field of noise reduction in image,

It has two disadvantages:

1) It is challenging to train deep CNNs for image demonization tasks.
2) The majority of deep CNNs struggle with performance saturation.

The only distinction between CNNs and deep CNNs is the quantity of layers. Conventional neural networks are the primary distinction between CNNs and Deep CNNs. It uses hierarchical patch-based networks, which lowers costs and abstracts the images on a many of feature levels. The deep CNNs is a neural network with multiple layers

B. Unsupervised Learning Models:

The deep learning models that have been covered so far are the supervised learning models. However, there might be instances in which unstructured and raw data is offered. A few unsupervised learning algorithms are suggested, such as the autoencoder and Boltzmann machine, which are briefly discussed in the sections that follow.

1) Auto-encoders:

When a hidden layer reconstructs the input layer in auto-encoders, the unsupervised mode is employed. Dimensions are assigned, in contrast to RNN, which incorporates weights and bias from the input state to the hidden state. The non-linear transformation function is used to calculate the hidden layer's stimulation since it has smaller dimensions than the input layer. Meanwhile, a dominant structure emerges in the input as the dimensions of the hidden layer are reduce To learn about the identity function, the dimensions of the input and hidden layers should remain unchanged, and no non-linearity function should be included. There are two subcategories of autoencoders. The first the denoising auto-encoder, and the second model is proposed to prevent the trivial solution of learning. This type uses the corrupted version of noise to rebuild the input. Stacked auto-encoders are a different type that are constructed by stacking the auto-encoder layers on top of each other. In healthcare applications, supervised learning techniques are used to fine-tune the entire network after each layer is trained separately to anticipate the output.

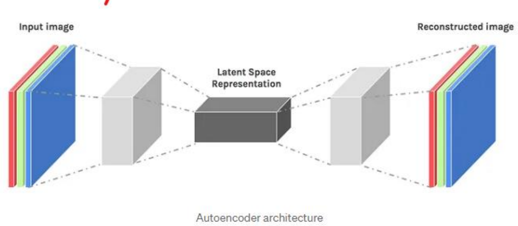

Fig 5: Auto-encoders

2) Boltzmann Machine and Restricted Boltzmann Machine

The Boltzmann machine's version, the Restricted Boltzmann machine, is a sort of Markov Random Field that is used in unsupervised learning input data models. It is regarded as a type of stochastic neural network that is modelled with a particular Gaussian distribution and stochastic nodes. The main goal is to minimize the reconstruction error by allowing the model to learn by adjusting the weights through sampling or Gibbs sampling. These networks are known as acyclic directed graphs, helpful when a probabilistic distribution between the input data is needed. The composition of undirected nodes in RBMs is another crucial feature, implying that nodes propagate in all directions. The RBM is trained using contrastive divergence, i.e. an unsupervised learning approach. The deep Boltzmann machine is another variation in the Boltzmann family. Its concept resembles stacked autoencoders, the primary distinction is the restricted Boltzmann machine which replaces the auto encoder's layer. First, training for the individual layers is performed in an unsupervised manner, and then a linear classifier is included in the topmost layer, leading the training in a supervised direction.

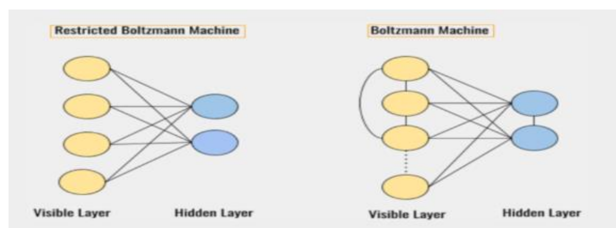

Fig 4: Boltzmann Machine and Restricted Boltzmann Machine

3.Deep Learning Tools:

Deep learning provides an array of tools that leverage computers to transform data into useful knowledge. Machine learning relies heavily on tools, and choosing the appropriate tool can be just as crucial as using the best algorithms. Deep learning tools are mentioned below.

1. TensorFlow: An open-source artificial intelligence library, using data flow graphs to build models and allows developers to create

large-scale neural networks with many layers.

2. MXNet: An open-source software library for numerical computation using data flow graphs and supports DL architectures CNN and RNN.
3. Caffe: A cross-platform support C++, MATLAB, and Python programming interfaces.
4. Theano: Provides capabilities like symbolic API supports looping control (scan), which makes implementing RNNs easy and efficient.
5. Keras: Theano based deep learning library.
6. ConvNet: MATLAB based convolutional neural network toolbox.
7. Deeplearning4j: An open-source, Apache 2.0-licensed with distributed neural net library written in Java and Scala.
8. Apache Singa: Open-source library for deep learning.
9. H20.ai: Open source used by leading healthcare companies to deliver AI solutions that are changing the industry.

4. Conclusion:

In this publication, we have presented a Brief overview of deep learning studies concerning predictive analysis of healthcare data. This study's main goal is to suggest a methodology for applying DL and predictive analysis to monitor healthcare data. Thus, the suggested model can be put into practice to validate the empirical findings that demonstrate its advantages. By examining recent developments in deep learning, we can pinpoint a crucial region that is gaining traction in unsupervised learning and has great potential for application to medical imaging. Additionally, we anticipate that deep learning techniques will be applied to the health industry's variable domain. Thus, deep learning will have a significant influence on the analysis of healthcare data.

5. References

1. Dur-e-Maknoon nisar, Rashid amin, Noor-ul-huda shah , Mohammed a. al Ghamdi ,Sultan h. Almotiri , Meshrif alruily, “Healthcare Techniques Through Deep Learning: Issues, Challenges and Opportunities”, DOI 10.1109/ACCESS.2021.3095312, IEEE Access

2. Saroj Kumar Pandey, Rekh Ram Janghel, “Recent Deep Learning Techniques, Challenges and Its Applications for Medical Healthcare System: A Review”, published online: 09 January 2009, https://doi.org/10.1007/s11063-018-09976-2

3. S. Sikdar and S. Guha, “Advancements of healthcare technologies: Paradigm towards smart healthcare systems,” in Recent Trends in Image and Signal Processing in Computer Vision. Springer, 2020, pp. 113–132

4. Jafar Abdollahia, Amir Jalili Iranib, Babak Nouri-Moghaddama, “Modeling and forecasting Spread of COVID-19 epidemic in Iran until Sep 22, 2021, based on deep learning”,

5. K. B. DeSalvo, A. N. Dinkler, and L. Stevens, “The us office of the national coordinator for health information technology: progress and promise for the future at the 10-year mark,” Annals of emergency medicine, vol. 66, no. 5, pp. 507–510, 2015.

6. B. Norgeot, B. S. Glicksberg, and A. J. Butte, “A call for deep-learning healthcare,” Nature medicine, vol. 25, no. 1, pp. 14–15, 2019.

7. D. Mahapatra, P. K. Roy, S. Sedai, and R. Garnavi, “Retinal image quality classification using saliency maps and cnns,” in International Workshop on Machine Learning in Medical Imaging. Springer, 2016, pp. 172–179.

8. T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering

algorithm: analysis and implementation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 881–892, 2002.

9. M. A. Carreira-Perpinan and G. E. Hinton, "On contrastive divergence learning." in Aistats, vol. 10. Citeseer, 2005, pp. 33–40.
10. D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.-Z. Yang, "Deep learning for health informatics," IEEE journal of biomedical and health informatics, vol. 21, no. 1, pp. 4–21, 2016.
11. H. Chen, W. Wu, H. Xia, J. Du, M. Yang, and B. Ma, "Classification of pulmonary nodules using neural network ensemble," in International Symposium on Neural Networks. Springer, 2011, pp. 460–466.
12. R. Kapoor, S. P. Walters, and L. A. Al-Aswad, "The current state of artificial intelligence in ophthalmology," Survey of ophthalmology,vol. 64, no. 2, pp. 233–240, 2019.
13. B. A. Landman, I. Lyu, Y. Huo, and A. J. Asman, "Multiatlas segmentation," Handbook of Medical Image Computing and Computer Assisted Intervention, pp. 137–164, 2020.
14. R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," Neural computation, vol. 1, no. 2, pp. 270–280, 1989.
15. Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," Mechanical Systems and Signal Processing, vol. 138, p. 106587, 2020.
16. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.
17. M. Fatima, M. Pasha et al., "Survey of machine learning algorithms for disease diagnostic," Journal of Intelligent Learning Systems and Applications, vol. 9, no. 01, p. 1, 2017.
18. M. De Onis, "World health organization reference curves," The ECOG's eBook on Child and Adolescent Obesity, p. 19, 2015.
19. W. Xu, J. He, and Y. Shu, "Deephealth: Deep representation learning with autoencoders for healthcare prediction," in 2020 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2020, pp. 42–49.
20. O. Aouedi, M. A. B. Tobji, and A. Abraham, "An ensemble of deep auto-encoders for healthcare monitoring," in International Conference on Hybrid Intelligent Systems. Springer, 2018, pp. 96–105.
21. I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, Deep learning. MIT press Cambridge, 2016, vol. 1, no. 2.
22. M. Baucum, A. Khojandi, and R. Vasudevan, "Improving deep reinforcement learning with transitional variational autoencoders: A healthcare application," IEEE Journal of Biomedical and Health Informatics, 2020.
23. Li D (2014) A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Trans Signal Inf Process 3:e2
24. D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," The Journal of physiology, vol. 160, no. 1, pp. 106–154, 1962.