ISSN:2476-2377

Review Article

International Journal of Cancer Research & Therapy

Deep Learning in Oncology- A Case Study on Brain Tumor

Prisilla Jayanthi G[1]*, Murali Krishna IV[2], Bakhtair Choudhary[3], Evangelyne Esther Y[4]

[1]*Assistant Professor (Research Scholar), Secunderabad, India*

[2]*Defence Research and Development Organization, Hyderabad, India*

[3]*Hyderabad Spine Clinics, Hyderabad, India*

[4]*ME (VLSI), Manipal University, Karnataka, India*

***Corresponding author**
Prisilla Jayanthi G, Assistant Professor (Research Scholar), Secunderabad, India, E-mail: prisillaj28@gmail.com

Submitted: 04 Dec 2017; **Accepted**: 11 Dec 2017; **Published**: 04 Jan 2018

Abstract

The brain tumor detection continues to be a challenge owing to the complexity of its symptoms. The research era indicates the tumor diagnosis and identification of tumor exact indicators are still uncertain. These tumors can appear anywhere in the brain and have any kind of shape, size, and contrast. The brain tumor exploration with deep learning is a solution for flexible, high capacity and extreme efficiency. The deep learning is an application of the artificial intelligence with multiple layers helping to predict the outcome of the disease early detection. This paper presents an approach to recognize the indicators and show that deep learning drops error rate for brain tumor diagnoses by 80%.

Keywords: Brain tumor, Deep Learning, Matlab, Neural Networks, Pesticides

Introduction

The increasing pace through the continuous improvement of techniques in medicine has scientific progress in the diagnosis and treatment of any disease. The patient management needs the appropriateness of the medical intercessions centered on scientific evidence or else, the health care is at risk, leaving little defense for the involvements and their associated costs. The concept of evidence-based medicine has influenced the patient and helped to recognize the diagnosis and treatment of diseases to be established for both patient and the community as a complete. The risks of surgical procedure, unwanted side effects of immunosuppressive drugs, life expectancy, recurrence of disease and quality of life should be considered by medical management. The patient management involving the prediction of future events is increasing tremendous.

The accurate predicting model takes abundant deal of time to progress. The models that survive are called robust, slight change to the environment in which prognostic models will not affect the performance level. A model-based system for prediction and prognosis promoted in artificial intelligence, through attractive tools for decision support.The structuring of predicting models is usually carried out using information extracted from clinical databases on scientific evidence. Machine learning provides new opportunities for automatically learning research in expressive models. Unfortunately, clinical databases of sufficient size and reliability filled with data from studies with rigorous patient follow-up, have still a certain drawbacks [1].

The wide spread of medical information systems and unpredictable growth of medical databases requires traditional manual data analysis to be coupled with methods for efficient computer-assisted analysis [2]. Considering the issues, artificial intelligence techniques and intelligent systems have found many valuable applications to be upheld [3]. In many biomedical areas, the neural networks have found to be very useful to help diagnosis of disease, for the pathological study and for monitoring the progress of outcome results from various treatments. Artificial neural networks (ANNs) are powerful tools to model the non-linear cause and effect relationships inherent in complex processes, usually for quality control [4,5].

Artificial Neural Networks

The applications of artificial neural networks are used in science and expertise. ANNs, make use of different types of input data that are processed in the prior training history on the defined database to produce a relevant output. The assessment of evaluation of the values, limits, potential trends, and future developments and associations to other branches of human medicine has greater confidence for the use of ANNs in the diagnostics approaches through selected examples, documenting the enormous variety of data that serve as inputs for ANNs.

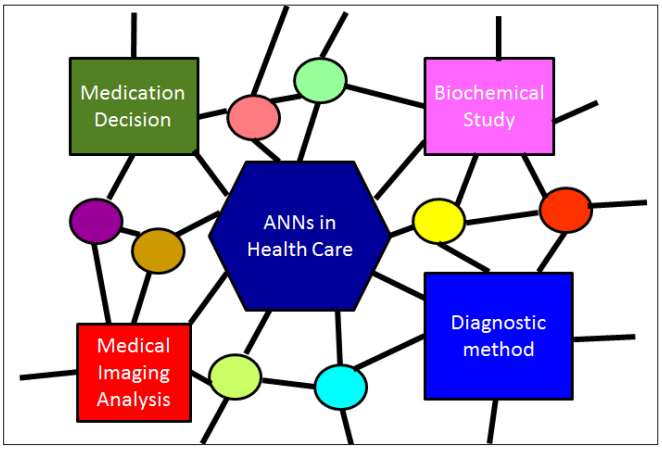

Figure 1: Overview of the main applications of ANNs in Health care

ANN applications are applied in research as they can represent non-linear systems in which the correlation among the variables is unknown. The added advantages of ANNs include self-learning, associative memory, high parallelism strength and high speed and error tolerance against noises which might be in constraints and their cheapness in the reuse of existing solutions is the best decision [6]. ANNs are called connection oriented networks, which include a set of processors act as parallel, taking the arrays of input at a time and produce output based on processing algorithm [7,8].

A neural network is fashioned by the series of neurons, organized in layers and each neuron is linked with another neuron in the next layer through a weighted link. The neural network is formed by an input layer, one or more hidden layers and the output layer. The complexity of the system determines the number of neurons and the number of layers required. The input layer receives the data and transfer to the neurons through the weighted links in the first hidden layer. The data are processed and the result is transferred to the neurons in the next layer, the network's output is provided by the neurons in the last layer [9].

Deep Learning

Deep learning, a technique with its foundation in artificial neural networks is emerging as a powerful tool for machine learning, promising to reshape the future of artificial intelligence. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. In solving problems, deep learning is making advances that resist the best attempts of the artificial intelligence community for many years. Deep learning has explored tremendous outstanding results in image and speech recognition over the other techniques at predicting the activity of potential drug molecules, analyzing particle accelerator data, reconstructing brain circuits and predicting the effects of mutations in non-coding DNA on gene expression and disease [10-14].

For solving complex problems, neural networks create new neural connections with one or more hidden layers of the perceptron. Deep learning has provided new sophisticated approaches to train deep neural networks (DNN) architectures.The two methodologies namely, supervised and unsupervised learning is trained. In supervised learning, labelled data are used to train the DNNs and the weights that minimize the error to predict a target value for classification or regression, but in unsupervised learning, the training is performed without requiring labeling data; it is usually used for clustering, feature extraction or dimensionally reduction.

Medical Informatics

Medical informatics emphasis on the large accumulated of data analysis in health care with the aim to enhance and develop clinical decision support system for quality assurance and accessibility of health care services. Electronic health records are the source of patient information that include medical history details such as diagnosis, diagnostic exams, medications and treatment plans, immunization records, radiology images, electroencephalogram EEG, laboratory and test lab reports. The efficient mining of this detail would provide valuable understanding of disease management for several reasons:

1. The quality of reports varies from one organization and the person owing to varying length, irregular sampling, lab results, lack of structured reporting and missing data.
2. Long and varying delays separate the onset of disease from the appearance of symptoms.

The medical imaging, the cytological follow-ups of a tumor diagnosis may include captivating information like its stage and spread. This information is beneficial to acquire a complete view of a patient condition or disease and then be able to improve the quality of the obtained inference [15].

Deep Learning in Medical and Imaging Informatics

Medicine is an art and a science, but the science commands the art. Medicine, directly or indirectly, is evidence-based and produces hard scientific data to back up the marketing privileges. The challenge of starting the medical company is to focus on the patient pain and suffering without underestimating the size, complexity and regulations of the health care system and scientific consistency. The health care must address the improvement of the health of the populations, minimize the cost of care, improve the patient experience and to improve the provider experience by retaining the best people.

The health care is managed by a large number of highly educated, highly trained, highly cultured individuals with respect to the validated accumulated knowledge and wisdom and culture of unselfishness and compassion; challenging the people to do the best. The health care is fundamentally different from every other industry. People give the highest priority to their health, expect the highest levels of care and service irrespective of cost, emotions and ideological about the industry. Radiology is the domain used to find a technology driven path and apply the tools of the programmer.

The present concern in deep learning in healthcare stems from two factors, firstly the flowering of machine learning techniques and secondly, the explosion of available healthcare data. This effectively transformed medical record from carbon paper to silicon chips is available in structured and unstructured data. Medical images are much more complex. The machines producing images of hundreds of anatomical structure and pathophysiologic process results in thousands of observable imaging features.The worth of deep learning in healthcare systems comes with improving accuracy and increasing efficiency. Present health care is a human-machine collaboration that ultimately becomes inter dependence and both humans and machines performing both tasks at which they are suboptimal. Deep learning systems develop and evolve by assisting humans with those tasks at which humans are not good. For instance, humans

are very good at processing information from their senses including vision. They are very good at recognizing human emotions. But at certain measures, humans cannot be reliably at remembering things, searching for and organizing data and not good at correlating and reasoning about that data. The deep learning systems will make doctors and other providers faster and smarter in their diagnoses and reduce uncertainty in their decisions, thereby avoiding the costs and hazards and saving time [16].

Patient Statistics Inputs

The predicted values for the brain tumor of 101 patients are disclosed in table1.

Table 1: Predicted values for the brain tumor of 101 patients

Observation	*Predicted brain Tumor*	*Residuals*	*Standard Residuals*
1	1.039921052	-0.039921052	-0.119690046
2	1.315733843	-0.315733843	-0.946623297
3	0.891355556	0.108644444	0.325734361
4	0.571949314	-0.571949314	-1.714800476
5	1.071756737	-0.071756737	-0.215138796
6	1.183516327	-0.183516327	-0.550212891
7	0.810524572	0.189475428	0.568079279
8	0.948181416	0.051818584	0.155360851
9	1.017776105	-0.017776105	-0.053295761
10	0.718262012	-0.718262012	-2.153470614
11	0.797368616	0.202631384	0.607523053
12	1.025535593	-0.025535593	-0.076560013
13	0.985127725	0.014872275	0.044589588
14	1.205525637	-0.205525637	-0.616200512
15	0.924418352	0.075581648	0.226606526
16	0.589630755	-0.589630755	-1.767812418
17	1.187523897	-0.187523897	-0.562228261
.....			
88	0.699938275	0.300061725	0.899635643
89	0.732357194	0.267642806	0.802438259
90	0.578996905	-0.578996905	-1.735930342
91	0.709804903	0.290195097	0.87005383
92	0.615644378	0.384355622	1.152362959
93	0.70416683	0.29583317	0.886957723
94	1.138702022	-0.138702022	-0.415852049
95	1.012425939	-0.012425939	-0.037255061
96	0.740814303	0.259185697	0.777082419
97	0.533892322	0.466107678	1.397469406
98	0.856630444	0.143369556	0.4298461
99	0.547987504	-0.547987504	-1.642958931
100	0.925921773	0.074078227	0.222099013
101	0.560673168	-0.560673168	-1.68099269

Table 2: The significance F's value

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	8	3.667389775	0.458423722	3.791115489	0.000695804
Residual	92	11.12468943	0.120920537		
Total	100	14.79207921			

Table 2 shows the significance F value has 0.000695804 which is less than 0.05, if the value is greater than 0.05, it's perhaps better to stop using the set of independent variables.

After analyzing 101 reports of various patients suffering from brain tumor of different ages, the various symptoms are trained through neural network, using Leven-berg Marquardt backpropagation algorithm adjusting the number of hidden layers and modifying weights. Using matlab R2012b, the outputs of the network fitting model and performance plot are shown in figure 2 and 3.

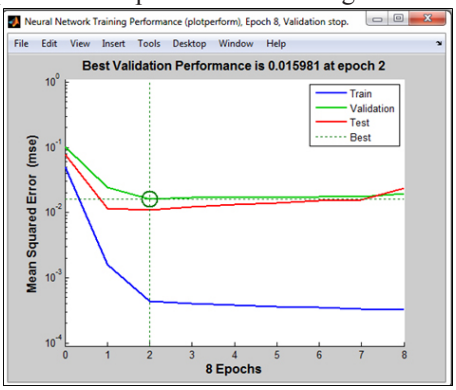

Figure 2: Performance plot for best validation performance is 0.015981 at epoch 2

The performance plot shows how the network mean squared error drops drastically; the blue line shows decreasing error on the training data, the green line shows the validation error, the training stops when the validation error stops decreasing and the red line shows the error on test data indicating how well the network will generalize to the new data. The standard network that is used for function fitting is n-layer feedforward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. Let's say the number of hidden neurons is set to 44. With the input vectors and target vectors randomly divided into three sets as follows: for training, 65% will be used; 20% will be used to validate that the network is generalizing and to stop training before overfitting and the last 15% will be used as a completely independent test of network generalization. The purpose of each training set is used to fit the models; the validation set is used to estimate the prediction error for model selection, the test set is used for assessment of the generalization error of the final chosen model.

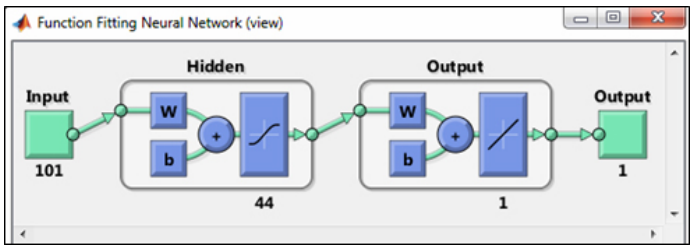

Figure 3: Neural network training with 101 inputs, 44 hidden layers and 1 output

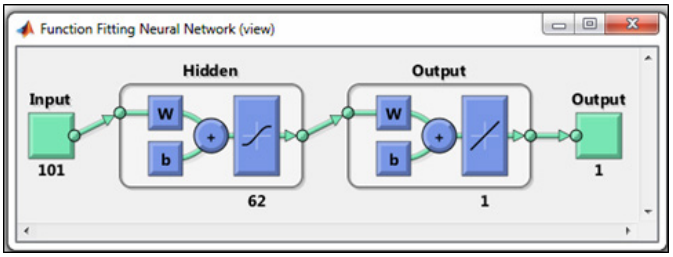

Figure 4: Neural network training with 101 inputs, 62 hidden layers and 1 output

The advantages of neural networks are their ability to generalize. This means that a trained net would classify data from the same class as the learning data that it has never seen before. The learning should be stopped in the minimum of the validation set error. At this point the net generalizes best. When learning is not stopped, overtraining occurs and the performance of the net on the whole data decreases, the error on the training data is reduced. After finishing the learning phase, the net should be finally checked with the third data set: the test set. The key reasons for using neural network:

1. The dependence between input and output data is nonlinear and the neural networks have ability to model non-linear patterns.
2. The ANN can be more reliable at predicting. When huge numbers of data available, the neural network, trains this data, adjusts the weights and predicts output with small error when working on new data with similar characteristics of the input data.

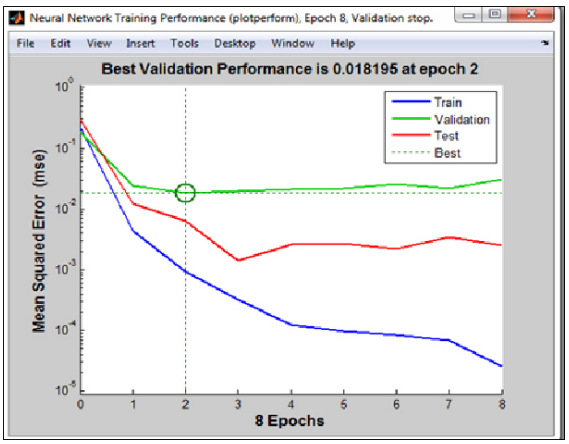

Figure 5: Performance plot for best validation performance is 0.018195 at epoch 2

Discussions

Occupation of the patient is the key component of health hazards. The widespread use of pesticides in the agricultural industry to enhance the crop production and to control the insects, pests and fungus are identified as a severe chemical health hazard for the agricultural workers, residents and children by the direct contact and by polluting the airborne, soil and water environment [17]. Tons of pesticides, insecticides, chemical sprayers and fungicides (chemicals like chlorpyriphos, mancozeb, captan, dimethoate, gasoline, etc.) are being used by the farmers to spray the plants, fruits and the leaves at different stages of growth to avoid the infestations and destruction of the fruits. etc., and indirectly like children playing in and around crops, houses near the crops are affected by meningioma, one type of brain tumor [18].

A link between polychlorinated biphenyls (PCB) exposure and brain tumor has been strongly proven in the epidemiologic research [19]. The early childhood development and adult neurologic functions may be impaired by PCB exposure [20,21]. The epidemiologic studies support an association between PCB exposure and central nervous system disease are persistent environmental toxicants associated with numerous health effects, including Parkinson disease, amyotrophic lateral sclerosis, non-Alzheimer related dementia, and brain cancer in adults [22,23]. The principal mechanism of PCB-induced brain metastases and involvement of neurotoxic and neurodevelopmental effects of these environmental toxicants. Environmental exposure to PCBs is ongoing as a result of continued use and disposal of products containing these toxicants widespread bioaccumulation of PCBs in the biosphere, and bio-concentration in the food chain [24].

Table 3: Summary output of P-value of patient's indicators

	Coefficients	*P-value*
Intercept	0.522616176	1.91503E-05
Age	0.001409518	0.503870376
Headache	0.135549396	0.002187604
Vomiting	0.073004669	0.047823148
Seizure	0.142084144	0.007944785
Altered behavior	-0.011918441	0.905271455
Decreased vision	0.04522345	0.357974994
Difficulty in speech	0.067787907	0.496593341
Loss of consciousness	0.158662493	0.039570227

The p-value is displayed in the Table 3 for all the independent variables and its coefficient values. The indicators like headache, vomiting, seizure and loss of consciousness have p-value less than 0.05; are measured to be the most common symptoms of 101 patients.

Conclusion

The study of the patient reports reveals that the age has no boundary for the brain tumor to occur. The majority of the patients suffering from Meningioma shows that their occupation or likelihood is agriculture.The patients work on different insecticides, pesticides to yield a good crop; the pesticides used are harmful and hazardous to their health. It affects neurotoxins and cause neuronal disorder and by further research on their disease through questionnaires meningioma has a clear link with pesticide exposure in females rather than males. The extensive use of herbicide increases greatly the risk of meningioma. With the deep learning technique one can reduce the error in diagnosing the disease, predict the result faster and provide indispensable treatment for the patient's life-saving.

Acknowledgement

I express my gratitude to my research guide Dr. I.V Murali Krishna for his constant support, guidance and his efforts to ensure that the study was carried out with care and with due consideration. A special thanks to my medical research guide Dr. Bakhtair Choudary for his advice, support and time. I thank Dr. Mohana Vamsy Ch (Chief Surgical Oncologist, Omega Hospital, Hyderabad, India) for assistance in data collection and for general support.

References

1. Peter JF Lucas, Ameen Abu-Hanna (1999) "Prognostic methods in medicine", Artificial Intelligence in Medicine 15: 105-119.
2. Lavrac N (1999) "Selected techniques for data mining in medicine", Journal of Artificial Intelligence in Medicine 16: 3-23.
3. Teodorrescu T, Kandel A, Jain LC (1998) "Fuzzy and neuro-fuzzy systems in medicine", Boca Raton, FL: CRC press.
4. Shi X, Schillings P, Boyd D (2004) " Applying artificial neural networks and virtual experimental design to quality improvement of two industrial processes", International Journal of Production Research 42: 101-118.
5. Fatemeh Hoda Moghimi and Nilmini Wickramasinghe (2014), "Artificial Neural Network Excellence to Facilitate Lean Thinking Adoption in Healthcare Contexts", DOI 10.1007/978-1-4614-8036-5_2, Springer Science + Business Media New York.
6. Lijuan Liu, Mingrong Deng (2010), "An Evolutionary Artificial Neural Network Approach for Breast Cancer Diagnosis", Third International Conference on Knowledge Discovery and Data Mining (WKDD), IEE Phuket, Thailand 593-596.
7. Sarasvathi V, Santhakumaran A (2011) "Towards Artificial Neural Network Model To Diagnose Thyroid Problems", Global Journal of Computer Science and Technology 11: 53-55.
8. Farhad Soleimanian Gharehchopogh, Maryam Molany and Freshte Dabaghchi Mokri, (August 2013), "Using Artificial Neural Network in Diagnosis of Thyroid Disease: A Case Study", International Journal on Computational Sciences and Applications (IJCSA) 3.
9. Filippo Amato, Alberto López, Eladia María Peña-Méndez, Petr Vaňhara, Aleš Hamp, Josef Havel1 (2013), "Artificial neural networks in medical diagnosis", Journal of APPLIED BIOMEDICINE Appl. Biomed 11: 47-58.
10. Ma J, Sheridan RP, Liaw A, Dahl GE, Svetnik V (2015) "Deep neural nets as a method for quantitative structure-activity relationships". J Chem Inf Model 55: 263-274.
11. Ciodaro T, Deva D, de Seixas J, Damazio D (2012) "Online particle detection with neural networks based on topological calorimetry information". J Phys Conf Series 368: 012030.
12. Helmstaedter M, Kevin L Briggman, Srinivas C Turaga, Viren Jain, H Sebastian Seung, et al. (2013) "Connectomic reconstruction of the inner plexiform layer in the mouse retina". Nature 500: 168-174.
13. Leung MK, Xiong HY, Lee LJ, Frey BJ (2014)"Deep learning of the tissue-regulated splicing code". Bioinformatics 30: i121-i129.
14. Yann LeCun, Yoshua Bengio, Geoffrey Hinton, (May 2015) Deep Learning", Article in Nature.
15. Daniele Ravi, Charence Wong, Fani Deligianni, Melissa Berthelot, Jainer Andreu- Perez, Beeny Lo and Guang-Zhong Yang, "Deep Learning for Health Informatics", IEEE Journal of Biomedical and Health Informatics.
16. https://medium.com/the-mission/deep-learning-in-healthcare-challenges-and-opportunities-d2eee7e2545.
17. Brown SK, Ames RG, Mengle DC (1989) "Occupational illness from cholinesterase inhibiting pesticides among agriculture pesticide applicators in California. Arch Environ Health" 44: 34-39.
18. Abdul Rashid Bhat, Muhammed Afzal Wani, AR Kirmani, TH Raina (Oct- Dec 2010) "Pesticides and brain cancer linked in orchard farmers of Kashmir", Indian Journal of Medical and Paediatric Oncology 31.
19. Hennig B, Reiterer G, Majkova Z, Oesterling E, Meerarani P, et al. (2005) "Modification of environmental toxicity by nutrients: implications in atherosclerosis". Cardiovasc Toxicol 5: 153-160.
20. Grandjean P, Weihe P, Burse VW, Needham LL, Storr-Hansen E, et al (2001) Neurotoxicol Teratol 23: 305-317.
21. Faroon O, Jones D, de Rosa C (2001) "Effects of polychlorinated biphenyls on the nervous system", Toxicol Ind Health 16: 305-333.
22. Steenland K, Hein MJ, Cassinelli RT II, Prince MM, Nilsen NB, et al (2006) "Polychlorinated biphenyls and neurodegenerative disease mortality in an occupational cohort". Epidemiology 17: 8-13.
23. Petersen MS, Halling J, Bech S, Wermuth L, Weihe P, et al (2008) "Impact of dietary exposure to food contaminants on the risk of Parkinson´s disease. Neurotoxicology". 29: 584-590.
24. Melissa Seelbach, Lei Chen, Anita Powell, Yean Jung Choi, Bei Zhang, et al. (2010) "Polychlorinated Biphenyls Disrupt Blood-Brain Barrier Integrity and Promote Brain Metastasis Formation", Environmental Health Perspectives 118.