

A ANN Based Machine Learning Predictive Techniques to Classify the Level of Cancer Diseases using Recurrent Neural Networks

Nageswara Rao Eluri¹, Ram Mohan Reddy Dondeti², Naresh Malapati³

¹Associate Professor, RVR&JC College of Engineering, Guntur

^{2,3}Associate Professor, Newtons Institute of Engineering and Technology, Macherala

***Corresponding Author:** Nageswara Rao Eluri, Associate Professor, RVR&JC College of Engineering, Guntur

Abstract: Cancer is a disease which plays vital role in the field of medical domain. Machine Learning predictive techniques used by the various researcher in order to determine the level of cancer diseases. This paper used to predictive the level of cancer diseases using Recurrent Neural Network by three process. Initially Trained and Testing dataset given as an input and these data are further processed into the proposed research work. Next Process is predictive techniques using Long Short Term Memory with Recurrent Neural Network which consist of three stages such as Forget Gate Layer, Input Gate Layer and Output Gate layer. These Three Layers will predict the past, present and future level of cancer diseases. Next Process about performance evaluation of the proposed research work using confusion matrix which leads to be an effective result in classification of cancer diseases.

Keywords: RNN, CNN, Cancer Disease, LSTM

1. INTRODUCTION

1.1. Predictive Techniques

A total examination of various sort of profound learning neural networks and their relating working in the event of less number of info information should be finished. There are three kinds of neural organization that are thought about. Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Recursive Neural Network (RNN).

1.2. Convolutional Neural Network (CNN)

Utilizes multi-facet perceptron by interfacing every hub with hubs of the following secret layer. This sort of neural organization has numerous associations which increment the quantity of loads and handling time for every cycle. The handling season of calculation is intensely reliant upon the yield unions. At the point when the quantity of inner association increments over fitting happens. This over fitting prompts expansion in preparing and preparing time which may not be appropriate for ongoing applications. In CNN the quantity of covered up layer neuron is straightforwardly relative to number of information factors. Each secret layer of neural organization will be utilized to characterize an explicit piece of the information. Along these lines utilizing a solitary covered up layer design acknowledgment framework will lessen the precision of CNN.

1.3. Recursive Neural Network (RNN)

Utilizes dull investigation to discover the example in the information. In recursive neural organization the weight between every hub will stay unaltered until intermingling. The neural organization will be shaped on the premise of tree structure and each datum that is identified with the following information will be associated with a similar parent hub. The covered up neurons are partitioned into three hubs initiation hub, parent hub, kin hub. The initiation work is utilized to compute the heaviness of the parent hub for every information type. These parent hubs are then associated with yield hubs to discover the yield likelihood. This sort of approach will be reasonable on the off chance that the quantity of yield and info hubs are two little. This is because of the way that with expansion in

number of yield hubs the likelihood worth will be dispersed and will lessen the time for assembly. This sort of calculation is best appropriate for feeling recognition in word handling where the yield esteems either sure or negative.

1.4. Recurrent Neural Network

Mainly focused with hidden layer used in the research work. The main objective of the Recurrent Neural Network focused on the previous input remains in the network which access the internal memory. Recurrent Neural Network focused on forget gate layer, input gate layer and output gate layer. Long Short term memory is the one of the functionalities of Recurrent Neural Network.

2. LITERATURE REVIEW

2.1. Lei Chen et. Al (2019)

Proposed tissue-precisenatural phenomenon shapes the formation of tissues, while gene expression adjustmentsmirror the reaction of the bod to environmental stimulations or pressure, significantly in wellness conditions, comparable to cancers. Applied the innovativefunctiondesiremethod with a supervised classifier perennial neural network (RNN) (Chen, 2019)to choose out the discriminate alternatives for classifying huge expressed genes from seldom expressed genes partner degreed assemblealmost appropriate RNN classifier. The Youden' indexes generated through the most appropriate RNN classifier and evaluated using a 10-fold pass validation had been 0.739 for conventional tissues and 0.639 for cancers. Furthermore, the underlying mechanisms of the important thing discriminate GO and KEGG alternativeshad beenanalysed. Results will facilitate the identity of the expression panorama of genes and elucidation of butnatural phenomenon shapes tissues and additionally the microenvironment of cancers.

2.2. Konstantina Kourou et.al (2015)

The capability of ML equipment to notice key aspects from complicated datasets exhibits their importance. A range of these techniques, which include Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs) and Decision Trees (DTs) have been broadly utilized in most cancers lookup for the improvement of predictive models, ensuing in wonderful and correct choice making. Even although it is evident thatthe use of ML strategies can enhance our grasp of most cancers' progression, a suitable degree of validation isneeded in order for these techniques to be viewed in the daily medical practice. In this work, we current an evaluate of latest ML processes employed in the modelling of most cancers' progression. The predictive fashions discussedhere are based totally on quite a number supervised ML methods as nicely as on unique enter aspects and statistics samples.Given the developing vogue(Kourou, 2018) on the utility of ML strategies in most cancers research, we existing right here the most current publications that rent these strategies as a purpose to mannequin most cancers hazard or affected person results

Abdullah AL Nahid (2018) proposed a cancer disease is a genuine danger and probably the biggest reason for death of ladies all through the world. To identification of malignant growth to a great extent relies upon advanced biomedical photography examination, for example, histopathological pictures by specialists and doctors. Examining histopathological pictures is a nontrivial assignment, and choices from examination of these sorts of pictures consistently require particular information. Be that as it may, PC Supported Determination (computer aided design) methods can help the specialist settle on more solid choices. Tostate of-the-workmanship Profound Neural Organization (DNN)(Nahid, 2018) has been as of late presented for biomedical picture investigation. Regularly each picture contains underlying and factual data. This paper classifies a bunch of biomedical bosom disease pictures (BreakHis dataset) utilizing novel DNN methods directed by primary and factual data got from the pictures. Specifically, a Convolutional Neural Organization (CNN), a Long-Momentary Memory (LSTM), and a mix of CNN and LSTM are proposed for bosom disease picture classification. SoftMax and Backing Vector Machine (SVM) layers have been utilized for the dynamic stage after extricating highlights using the proposed novel DNN models. In this examination the best Exactness worth of 91.00% is accomplished on the 200x dataset, the best Accuracy esteem 96.00% is accomplished on the 40x dataset, and the best F-Measured.

2.3. Priyadharshini et al (2019)

Described in the field of helped malignancy determination, it is normal that the association of AI in illnesses will give specialists a subsequent assessment and assist them with making a quicker/better assurance. There are an immense number of studies around here utilizing customary AI strategies and in different cases, utilizing profound learning for this reason. This article plans to assess the prescient models of AI arrangement in regards to the precision, objectivity, and reproducible of the determination of dangerous neoplasm with fine needle goal. Additionally, Try to add one more class for testing in this data set as suggested in past investigations. Present six distinctive characterization techniques: Multi-facet Perceptron, Choice Tree, Arbitrary Timberland, Backing Vector Machine and Profound Neural Organization for assessment. For this work, we utilized at College of Wisconsin Emergency clinic information base (Priyadharshini, 2019) which is made out of thirty qualities which describe the properties of the core of the bosom mass. As showed in outcome areas, DNN classifier has an extraordinary exhibition in precision level (92%), showing better outcomes comparable to conventional models. Arbitrary woodland 50 and 100 introduced the best outcomes for the ROC bend metric, considered a great forecast when contrasted with other past investigations distributed

3. PROBLEM STATEMENT

A Various ML Application are used to determine the level of cancer diseases implemented in the cyber world. Prediction of cancer diseases are predicted in three different levels. The Problem to determine the level of cancer diseases are as follows

- a) Prediction based on Susceptibility: Prediction of cancer diseases is to monitor the level of cancer disease which in turn to identify or monitor the cancer diseases.
- b) Prediction based on recurrence: This stage is used to monitor the redevelopment of cancer even when they are in medical treatment.
- c) Prediction based on Survival: This stage used to predict the survival of cancer diseases

This paper mainly focused on the predication based on recurrence. The Cancer disease which may be regenerated within short interval of time. Initially cancer disease may be identified and treatment will be in process and after a regular interval of time cancer disease may be regenerated which is unpredictable. This unpredictable, regenerated cancer disease form the data set using in the machine learning application leads the important problem in the medical domain. These Problem are considering to provide an optimal solution to predict level of cancer diseases using Recurrent Neural Network.

4. PROPOSED WORK

The Proposed Research Work to determine the level of cancer disease from the given dataset using Enhanced Recurrent Neural Network. The Proposed Work flow explain in the given figure below.

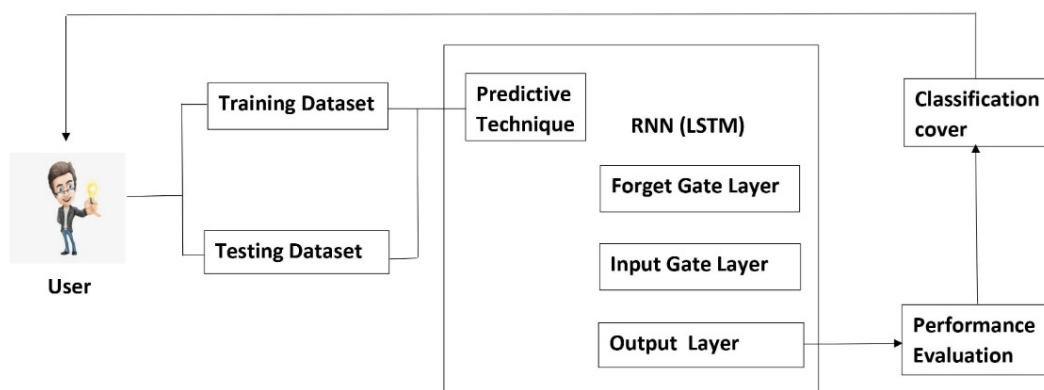


Fig4.1. Proposed Research Work

Initially, the dataset about the patients of the cancer hospital is considered as an input the machine learning application.

Training data set which consist of tuned data set of patients details in terms of Data Pre-processing, Data Transformation and Feature Construction. Testing Data set are determined by the experts which act as input to the proposed Machine learning application.

Predictive Techniques which consist of Feature construction, Feature selection method to determine the next level of cancer disease. Feature Construction is to determine the experimental analysis based on the two-cluster gene in the form of tissues related to cancer. Feature Selection method used to select the appropriate dataset which based on Recurrent Neural Network.

Recurrent Neural Network is an optimized prediction technique to classify the level of cancer disease with given dataset. RNN mainly focused on Neural Network with loop inside while enabling the information to sub sequential outputs with persistence.

Long Short-Term Memory (LSTM) Network is a special type of Recurrent Neural Network which consist of three stages as shown in the following



Figure4.1. Process of Long Short-Term Memory

The Proposed research work, forget gate layer which used to determine the previous information layer which may be disregarded. Input Gate Layer used to identify the information which used to pass the information to the next subsequent layer and update the current value which related to disease level. The Output Gate Layer which predict the classification of cancer diseases with current state value.

Let x be the sequence of x^t , the hidden state represents as h^t , the state of each cell represent by c^t , and output of the dataset represent o^t . Element wise multiplication process are determining in order to calculate the sigmoid based on the Long Short Term Memory.

The research work to predict the level of cancer diseases based Long Short Term Memory based on the Recurrent Neural Network to predict the level of diseases.

5. PERFORMANCE EVALUATION

The Final Stage of this research work used to determine the efficiency of research work with basic confusion matrix as shown in the following figure.

		ACTUAL VALUES	
		POSITIVE	NEGATIVE
PREDICTED VALUES	POSITIVE	TP	FP
	NEGATIVE	FN	TN

Figure5.1. Confusion Matrix

From the above graph, the following formula generated used to predict the performance of proposed research work

- a) Accuracy = $(TP+TN)/\text{No. of Samples} \rightarrow 1$
- b) Sensitivity = $TP/(TP+FN) \rightarrow 2$
- c) Specificity = $TN/(TN+FP) \rightarrow 3$

Accuracy, Sensitivity and Specificity are the three attributes used to determine the performance of the proposed research work.

6. EXPERIMENTAL ANALYSIS

Experimental Analysis are done based on the given data set in term of train data set, testing data set are implemented in MATLAB 2016 and execute in 2.6GHz clock and 8GB RAM in Microsoft Windows 10 environment. The data for analysis are taken from the given dataset. The following tables are generated from the sample of dataset

Table6.1. *Experimental Data*

No.of Samples	TP	FN	TN	FP	Accuracy	Sensitivity	Specificity
30	6106	1524	659	165	80.02	80.03	0.80
75	6097	1254	668	154	82.77	82.94	0.81
65	6037	1044	625	144	84.87	85.26	0.81
45	5971	1421	425	185	79.93	80.78	0.70
24	5468	1477	725	125	79.45	78.73	0.85
98	2458	1456	844	145	67.35	62.80	0.85
74	2458	1435	932	165	67.94	63.14	0.85
58	5833	1421	564	144	80.34	80.41	0.80
44	5750	1147	855	124	83.86	83.37	0.87
32	5479	1021	932	187	84.14	84.29	0.83

From the above table the accuracy and sensitivity are retrieved based on the confusion table represents the following figure 6.1

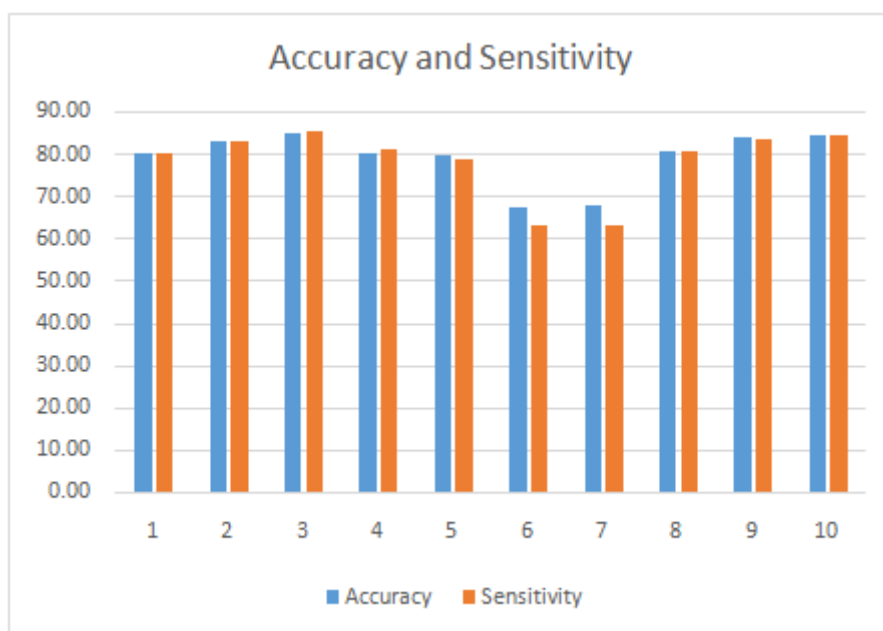


Figure6.1. *Accuracy and Sensitivity*

From the above graph accuracy high range leads to 84% and sensitivity with high range of 85%. Similar low range of accuracy around 67% and sensitivity low range 62%.

Specificity plays an important role in determining the classification of cancer diseases. From the table 6.1, the following graph generated the specificity as shown below.

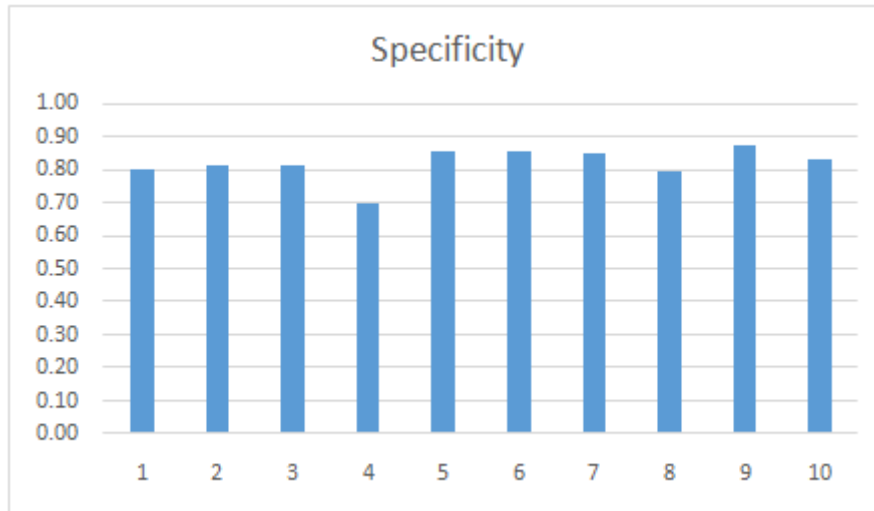


Figure6.2. *Specificity*

The Specificity of the proposed research work ranges from 0.80 to 0.87 ranges which plays vital role in the classification of cancer diseases using recurrent neural network.

7. CONCLUSION

This research work concluded to predictive cancer diseases using Recurrent Neural Network with Long Short-Term Memory. This Long Short-Term Memory which used to predict the cancer disease using Machine Learning application with three stages such as forget gate layer, input gate layer and output gate layer. These Gate will provide an optimized performance in predictive the classification level of cancer diseases. The Performance evaluation of this research work used to determine the accuracy, sensitivity and specificity leads to be an effective and efficient.

REFERENCES

- [1] Ali, N. A.-A. (2021). Enhancing the performance of cancer text classification model based on cancer hallmarks. *IAES International Journal of Artificial Intelligence*, 316.
- [2] Bavani, B. S. (2021). Classification of Arrhythmia Disease using Enhanced RNN Model. *Design Engineering*.
- [3] Calp, M. H. (Springer, Singapore). Use of Deep Learning Approaches in Cancer Diagnosis. In *Deep Learning for Cancer Diagnosis*, 2021.
- [4] Chen, L. X.-H.-D. (2019). Classification of widely and rarely expressed genes with recurrent neural network. *Computational and Structural Biotechnology*, 49-60.
- [5] Dutta, S. J. (2020). Breast Cancer Prediction Using Stacked GRU-LSTM-BRNN. *Applied Computer Systems*.
- [6] El-Halees, A. a. (2020). Breast Cancer Severity Degree Predication Using Deep Learning Techniques. *Jordanian Journal of Computers and Information Technology*.
- [7] Kourou, K. T. (2018). Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology*, 8-17.
- [8] Nahid, A.-A. M. (2018). Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *BioMed research international*.
- [9] Priyadarshini, Z. (2019). Feedback Based Adaptive Recurrent Neural Network for Cancer Detection using Gene Data Pattern. *International Journal of Engineering and Advanced Technology*.

- [10] Sen, A. S. (2021). Deep Learning in the Diagnosis of Lung Cancer. International Journal of Research in Engineering, Science and Management.
- [11] Shamrat, F. J. (2020). An analysis on breast disease prediction using machine learning approaches. International Journal of Scientific & Technology Research.
- [12] Shiva Shankar, R. K. (2020). Breast Cancer Disease Prediction With Recurrent Neural Networks (RNN). International Journal of Industrial Engineering & Production Research.

Citation: Nageswara Rao Eluri et al. "A ANN Based Machine Learning Predictive Techniques to Classify the Level of Cancer Diseases using Recurrent Neural Networks" *International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)*, vol 9, no. 1, 2023, pp. 32-38. DOI: <https://doi.org/10.20431/2349-4859.0901005>.

Copyright: © 2023 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.