

Application of Convolutional Neural Network for Cancer Disease Diagnosis - A Deep Learning based Approach

S. Sivanantham^{1*}, Dr. Hema Kumar M², A. K. Velmurugan³, Dr.K.Deepa⁴,
Akshaya V⁵

¹ Assistant Professor, Sree Vidyanikethan Engineering College, Tirupati, AP, India,

² Assistant Professor, Electronics and Communication Engineering, Sona College of Technology, Salem

³ Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India

⁴ Assistant Professor, Department of Computer Science and Artificial Intelligence SR University, Warangal, Telegana.

⁵ Assistant Professor, Sree Vidyanikethan Engineering College, Tirupati, AP, India,

ABSTRACT

Human are vulnerable to the terrible disease named cancer, which is a major factor in the high mortality rate. There are currently lesser DLTs (Deep learning techniques) or MLTs (machine learning techniques) for identifying cancer, despite advances in cancer treatment approaches. The proposed work performs a comparative study which compares the some significant DLTs like RFs (Random Forests), LSTMs (Long Short Term Memories), CNNs (Convolutional Neural Networks) and BPNNs (Back Propagation Neural Networks). These techniques are used here in this work for classification problem. The techniques are made to classify the medical records into benign and cancerous. Three pathological datasets are used to evaluate the above said techniques. CNNs provide the best performance of 0.97 accuracy and it is even good at its values of precisions, recalls and F1 scores.

Corresponding Author e- mail: sivanantham.s@vidyanikethan.edu

How to cite this article: Sivanantham S, Kumar M H, Velmurugan K A, Deepa K, V A, Application of Convolutional Neural Network for Cancer Disease Diagnosis - A Deep Learning based Approach. Journal of Complementary Medicine Research, Vol. 14, No. 1, 2023 (pp.69-75)

INTRODUCTION

Technology has advanced to the point that radiation now controls the atmosphere. These hazardous radiations result in a number of health problems. Cancer is one of the deadliest illnesses. Every year, there is an increase in the cancer death rate. Although many medications have been created to treat this dangerous condition, taking preventive measures may be the best courses of actions to reduce cancer mortalities. The proposed framework would assist in the development of a more effective strategy for cancer prediction. With the aid of the DLTs, classifications and predictions of cancerous data is possible [1]. To demonstrate the value of the model, each of the four technique is contrasted among themselves as they are the top performing classification methodologies. Examining the datasets of many cancer patients is the primary task to predict and prevent cancer. The datasets should include relevant information, such as the smoothness of the cell and the size of the tumor cell. Occurrence of classes in each dataset is the primary factor used to classify the cells.

KEYWORDS:

Application, Neural,
Deep, Learning

ARTICLE HISTORY:

Received : Nov 15 , 2022

Accepted : Dec 16, 2022

Published: Jan 13, 2023

DOI:

10.5455/jcmr.2023.14.01.14

RELATED WORK

The prediction of tumours has been the subject of numerous studies utilising association rule mining and different classification methodologies. Different classifiers get different results. When evaluating the performance of various classifiers, H. Bouali et al. used certain metrics that were computed from the confusion matrix. DTs (Decision trees), J48, - RFs, SVMs (Support Vector Machines) and NBs (Naive Bayes) are the classifiers that were used [2]. In many research investigations, RFs is regarded as the top categorizing method out of those listed. However, the experimentation carried out in this system refutes the already established RFs [3] techniques by showing data that show this classifier's accuracy to be worse than J48's. The recurrently occurring item set in a dataset is typically found through association rule mining. To determine the sort of cancerous tumours in the brain, T. Vijayakumar et al. discussed the capsule neural networks, well known MLTs that can be trained using fewer datasets than CNNs and are resilient to rotations or affine conversions [4]. In order to speed up the billing process, Subarna et al. developed an artificial intelligence-based image classification system that uses a camera to identify fruits and vegetables. The system's accuracy in comparison to other classification techniques, such as SVMs , KNNs (K-Nearest

Neighbors), RFs, and DA (Discriminant Analysis), is assessed [5]. When compared to existing MLTs like RFs, SVMs , NBs, and DTs in terms of accuracies (AUC values), DRs (detection rates), and FARs (false alarm rates), V. Akshaya et al. suggested ARM (Association rule mining) based cancer diagnosis method performed better [6], [7]. Patch level deep neural networks that learn from pixel-level labels was proposed as having a two-stage design by N. Wu et al. [8]. The authors have marked the beginning of applying DLTs in cancer diagnosis.

BACKGROUND

RFs

RFs can address issues in classifications and regressions. The supervised learning sub-categories of MLTs include RFs which are essentially ensemble learning techniques where outputs of various classifiers are pooled to discover solutions for issues, enhancing model performances. As implied by the name, RFs aggregates or bags data from numerous DTs [9]. Bagging is an ensemble meta-algorithm that, among the existing MLTs, produces reliable classification results. RFs build numerous DTs on distinct sets of provided data and consider averages to improve prediction accuracies in datasets [10].

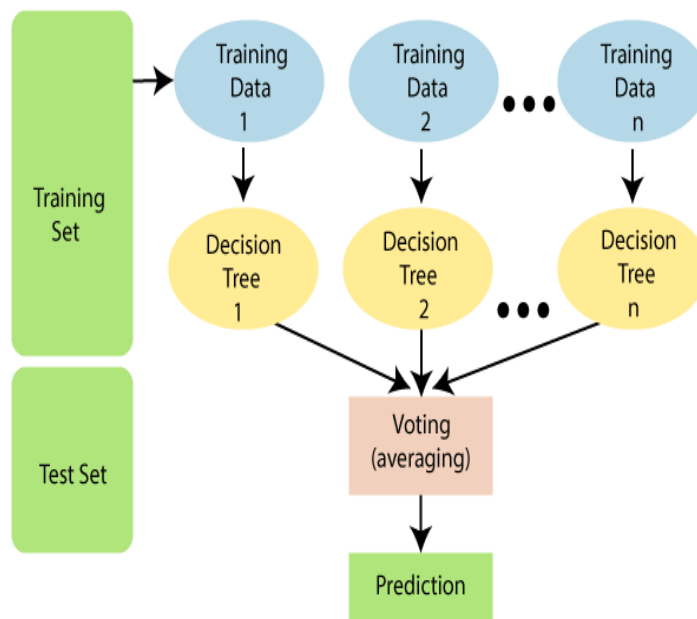


Figure 1 Principle of RFs.

RFs implementation in Python is shown in the following figure.

The total number of trees created by default is 100. It depends on how many attributes there are in the dataset.

```
[18] # Import Random Forest Classifier
from sklearn.ensemble import RandomForestClassifier

# Instantiate the classifier
rfc = RandomForestClassifier(random_state=0)

# fit the model
rfc.fit(X_train, y_train)

# Predict the Test set results
y_pred = rfc.predict(X_test)

# Check accuracy score
from sklearn.metrics import accuracy_score
print('Model accuracy score with 10 decision-trees : {0:0.4f}'.format(accuracy_score(y_test, y_pred)))
Model accuracy score with 10 decision-trees : 0.9416

[19] from sklearn.metrics import classification_report
print(classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
0	0.97	0.96	0.96	68964
1	0.83	0.85	0.84	15129
accuracy			0.94	84093
macro avg	0.90	0.91	0.90	84093
weighted avg	0.94	0.94	0.94	84093

Figure 2 Implementation of RFs

Convolutional Neural Network

CNNs are powerful DLTs for self - directed learning. CNNs are majorly used in image classification. By exchanging weights, it may reconstitute the original image's high-level semantic properties and enhance training efficiency. The convoluted neural networks use cascaded structures of simulated simple/complex cells for higher level feature extractions in visual cortexes. They consist of alternating layers of convoluted and pooling layers. The convolution layer's neurons react to a piece of the preceding layer's region and draw out the input's more complex properties. A portion of the input from the preceding layer is averaged or maximized by the neurons of the pooling layer, which prevents them from being somewhat deformed or displaced [11]. The complex neural network's final layers typically consist of several fully linked layers and classifiers. CNNs have been effectively used in recent years for voice recognitions, age estimates, human posture estimations, facial expression identifications, and face recognitions.

Feature-based approaches used for physical recognitions are CNNs. They differs from conventional artificial feature extraction techniques and feature-based high-performance classifier designs. It has the advantage that features are extracted using layer-by-layer convolution and dimensionality [12]. Through the use of multilayer nonlinear mapping, the network can then automatically learn to produce the identification task for the feature extractor and classifier from the training sample. The method reduces the requirement for the training sample, and the more network layers there are,

the more broadly applicable the learning feature is. The best DLTs to utilize for classification problems are CNNs. In this paper, CNNs are employed to distinguish between normal and malignant data in cancer identifications.

BPNNs (Back Propagation Neural Networks)

One of the key neural network methods is back propagation. Gradient descent is employed in the linear regression method to maximize the parameter. The gradient descent algorithm is utilized in a similar way here with back propagation. Back propagation techniques determine error function gradients in singular training. Back propagation can be expressed as a neural network function. They are used for effectively training ANNs (artificial neural networks) using gradient descents that take advantage of chain rules. Back propagations keep improving until there is no more scope for improvement mainly due to their iterative, recursive, and efficient methods for determining updated weights. Activation function's derivatives must also be known when networks are built for back propagations and are crucial to training of neural networks. They modify weights of neural networks in accordance with observed previous epoch's error rates (iterations). By properly tweaking the weights, we may lower error rates and improve the model's reliability by broadening its applicability [13].

The term "backward propagation of errors" has become "back propagation" in neural networks and is commonly used in developing ANNs. The weights of networks assist in computing gradient losses.

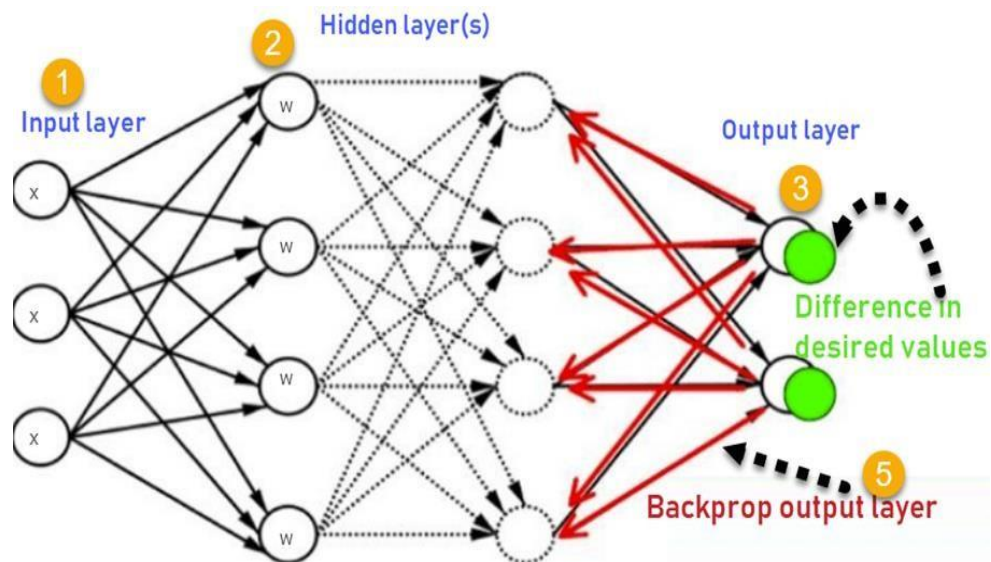


Figure 3 Illustration of Backpropagation NN

BPNNs are implemented using the following steps:

1. Input X comes through the associated channel.
2. The input is simulated using real weights W. Weights are often selected at random.
3. Identify the output of each neuron from the input layer to the output layer via the hidden layers.
4. the output error, please.

Actual Output: Error B; Desired Output:

5. To adjust the weights and lower the error, go back from the hidden layer to the output layer. Till the desired outcome is attained, repeat the process.

LSTMs

DLTs use LSTMs and are a type of RNNs (recurrent neural networks), particularly in sequence prediction issues and are capable of learning long-term dependencies. LSTMs have feedback connections, with the exception of single data points like images, which enable them to process data sequences [14]. These approaches can be used in speech recognitions, machine translations, and other complex tasks. LSTMs are peculiar variants of RNNs with exceptional performances. Models based on LSTMs rely on memory cells called "cell states" that is maintained across time spans. Cell states are depicted by horizontal lines that run across top in the image below. They are comparable to conveyor belts where data just travels without being altered. Information in LSTMs cell states can be added to or withdrawn through gates. Information can be optionally moved in and out of cells through these gates [15]. They facilitate the mechanism with layers of sigmoid neural networks and dot multiplication functions. Sigmoid layers output numbers between 0 and 1, where 0 implies "nothing should be let through" while 1 specifies "everything should be let through." [16]. To distinguish between recurrent nets,

training sequences are presented in both forward and backward directions in bidirectional LSTMs [17]. The same outputs are shared by both the sequences. For data points in given sequences, bidirectional LSTMs have complete information i.e. information before and after executions.

RESULT AND DISCUSSION

In the proposed research study, RFs, CNNs, Back Propagation NN and LSTMs are implemented, and the classification outcomes are evaluated with respect to accuracy. Data import, Pre-processing, and feature selection are carried out as general procedures in the application of any classification technique. Jupyter Notebook is used and python is the platform. A training set and a testing set are created from the entire dataset in an 8:2 ratio of the available instances. The first step is always to clean the data. The nulls are eliminated from the dataset in this step. This facilitates the mining of valuable information.

Datasets

The following section discusses the dataset description used in the proposed research work

GBSG2

The German Breast Cancer Study Group 2 in short called as GBSG2 is a collection of data comprising of several German cancer patients' factors. There are 686 occurrences and 10 attributes in it.

Biopsy

It is a dataset that was collected from cancer patients. With 699 cases and 11 attributes, it aids in determining if a tumour is benign or malignant [6].

WPBC

Wisconsin Prognostic Breast Cancer dataset abbreviated as WPBC is menostat, which has the pre and post classes. There are 198 instances and 34 characteristics in it. In terms of data, equal half of the instances are used for training and then the remaining other half are used as a testing set.

Confusion matrices are obtained for the classifiers during our experimentation that have been run across the three different cancer datasets, and the TP, TN, FP, and FN values are then derived from the matrices.

True positive (TP):- Instances of data that were accurately detected as cancerous or malignant.

True negative (TN):- Correctly sampled benign or Non-cancerous data instances.

False positive (FP):- Right Data samples that were identified as malignant.

False negative (FN):- Instances of data assessed incorrectly as malignant.

$$\text{Accuracy} = \frac{\text{TP_Value} + \text{TN_Value}}{\text{TP_Value} + \text{TN_Value} + \text{FP_Value} + \text{FN_Value}}$$

$$\text{Recall} = \frac{\text{TP_Value}}{\text{TP_Value} + \text{FN_Value}}$$

$$\text{Precision} = \frac{\text{TP_Value}}{\text{TP_Value} + \text{FP_Value}}$$

$$\text{F-Measure} = \frac{2 \times \text{Precision value} \times \text{Recall value}}{\text{Precision value} + \text{Recall value}}$$

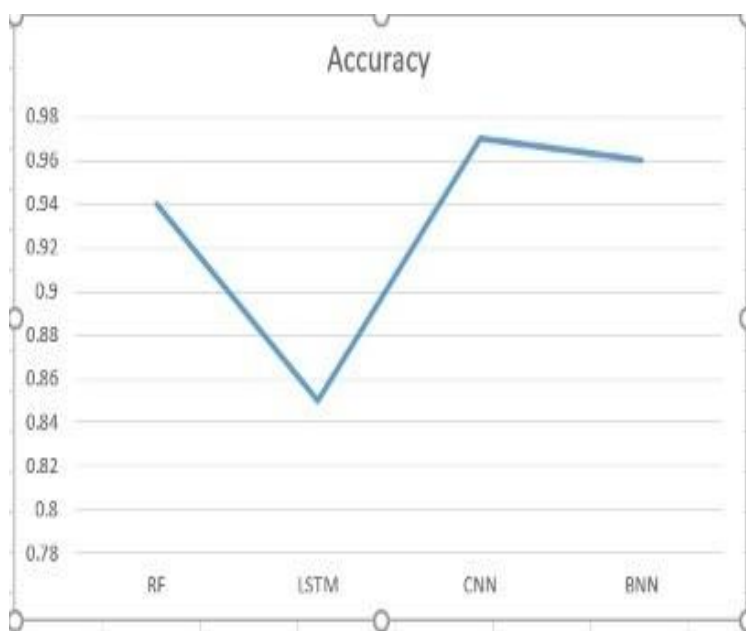


Figure 4 Comparison of Accuracy

The accuracy and F1 score comparison of the various approaches is shown in Figures 4 and 5.

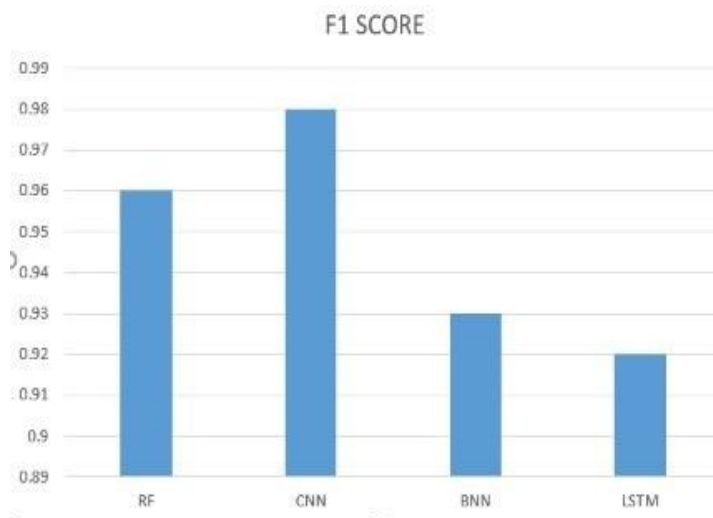


Figure 5 F1 score comparison

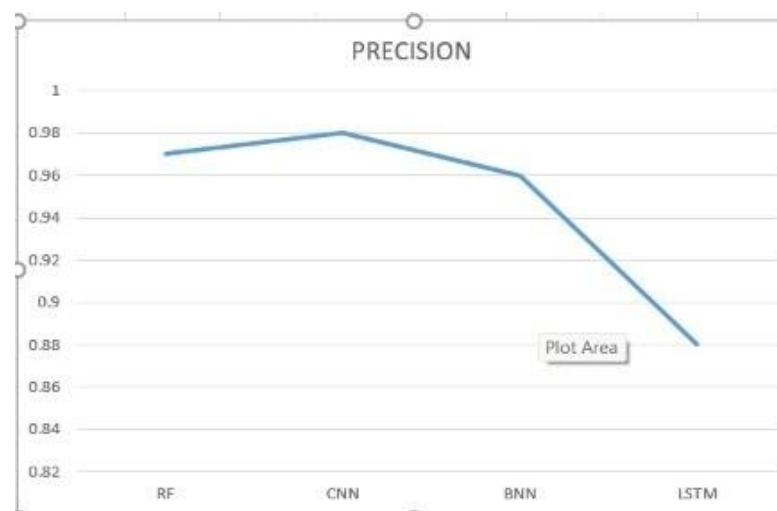


Figure 6 Precision Comparison

The precision and recall comparison of the discussed approaches is shown in Figures 6 and 7.

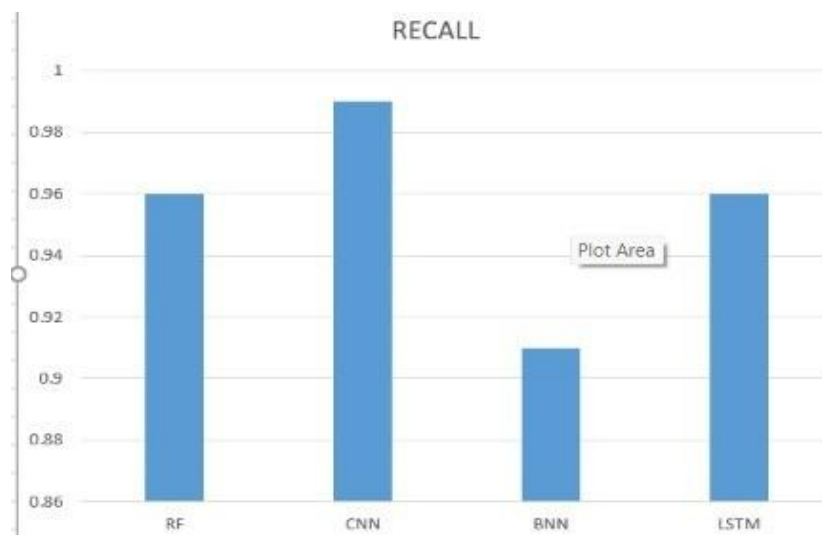


Figure 7 Recall Comparison

It is clear that RFs, CNNs, and BPNNs are all equally good in terms of precision and accuracy. All four methods—RFs, CNNs, LSTMs, and BPNNs deliver the best results in terms of recall. It is evident that the CNNs achieve best resultant values for accuracies, precisions, recalls and F1-scores when compared to the other classifiers.

CONCLUSION

Humans are prone to the lethal disease cancer, which contributes to the high death rate. Using three different cancer datasets, this research primarily analyses the effectiveness of Deep Learning classifiers such as RFs, CNNs, LSTMs, and BPNNs in the diagnosing of cancer. According to the findings, CNNs offer a maximum Accuracy and precision of 0.97 and 0.98 on average, respectively. Recall results 0.99 and F1 score is 0.98 for CNNs. Thus CNNs perform better than the other well-known classifiers in every way.

REFERENCE

1. R. Hajjo (2018). The Ethical Challenges of Applying Machine Learning and Artificial Intelligence in Cancer Care. 2018 1st International Conference on Cancer Care Informatics (CCI). 2018 : 231-231, 2018. doi: 10.1109/CANCERCARE.2018.8618186.
2. H. Bouali and J (2014). Akaichi. Comparative Study of Different Classification Techniques: Heart Disease Use Case. 2014 13th International Conference on Machine Learning and Applications, Detroit, MI: 482-486.
3. Z. Kouzani, S. Nahavandi and K. Khoshmanesh (2007). Face classification by a random forest," TENCON 2007 - IEEE Region 10 Conference, Taipei : 1-4.
4. T. Vijaykumar(2019). Classification of Brain Cancer type using Machine Learning. Journal of Artificial Intelligence. 1(2):105 - 113.
5. Subarna Shakya (2020). Analysis of Artificial Intelligence based Image Classification Techniques. Journal of Innovative Image Processing, 2(1) : 44 - 54.
6. V. Akshaya, V. Krishnamoorthy and S. Sivanantham (2020). Performance Evaluation of Association Rule based Classification in Cancer Diagnosis. 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA) :14-17. doi:10.1109/ICIRCA48905.2020.9182947.

7. Sivanantham S, Mohanraj V, Suresh Y and Senthilkumar J (2023) Association Rule Mining Frequent-Pattern- Based Intrusion Detection in Network. *Computer Systems Science And Engineering*;44(2): 617-31.
8. N. Wu et al (2020). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1184-1194,2020, doi: 10.1109/TMI.2019.2945514.
9. Sanjukta Mohanty, Arup Abhinna Acharya, Laki Sahu, and Sunil Kumar Mohapatra (2020). Hazard Identification and Detection using Machine Learning Approach. <https://ieeexplore.ieee.org/document/9121048>,13-15 May 2020.
10. L. Dong et al.,.Very High Resolution Remote Sensing Imagery Classification Using a Fusion of Random Forest and Deep Learning Technique–Subtropical Area for Example. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 : 113-128. doi: 10.1109/JSTARS.2019.2953234.
11. Er-Yang Huan, Gui-Hua Wen, Shi-Jun Zhang, Dan-Yang Li, Yang Hu, Tian-Yuan Chang, Qing Wang, Bing-Lin Huang (2017). Deep Convolutional Neural Networks for Classifying Body Constitution Based on Face Image. *Computational and Mathematical Methods in Medicine*; vol. 2017. <https://doi.org/10.1155/2017/9846707>
12. Sakthivel M , Sivanantham S, Kamalraj R and Krishnamoorthy V(2022). An Analysis of Machine Learning Depend on Q-MIND for Defencing The Distributed Denial of Service Attack on Software Defined Network. *International Journal Of Early Childhood SpecialEducation*; 14(5) : 3769-3776.
13. B. Leng, K. Yu, and J. Qin (2017). Data augmentation for unbalanced face recognition training sets. *Neuro computing*; . 235 : 10-14.
14. P. Rodriguez, G. Cucurull, J. González et al (2017). Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification. *IEEE Transactions on Cybernetics*. 2017.
15. Rahimzad, M, Moghaddam Nia, A., Zolfonoon, H. et al (2021). Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting, *Water Resour Manage* ; 35 : 4167-4187.
16. Sakthivel, M., Sivanantham, S., Akshaya, V., Sivakumar, D., & Karthikeyan, H. (2022). A Malicious Botnet Traffic Detection Using Machine Learning. *Journal of Pharmaceutical Negative Results*, 13(4), 968-977.
17. Sivanantham, S., Krishnamoorthy, V., Karthikeyan, D., Sakthivel, M., Mohanraj, V., & Akshaya, V. (2022). Web Hazard Identification and Detection Using Deep Learning-A Comparative Study. *Journal of Pharmaceutical Negative Results*, 13(4), 1052-1060.