NEURO FUZZY CLASSIFICATION FOR DATA MINING TASKS

Guided by,
Prof.Himansu Das
(Assistant Professor)

Priyanka Mahato (1505049) & Shalini Jaiswal (1505062)
School of Computer Engineering

Abstract

The ANN have output that fluctuates to a great extent, in order to increase accuracy fuzzy logic is used which takes into consideration contribution of each feature in a class. But it is very time consuming to overcome this problem reduction concept is used (PCA, LDA, ICA) and comparative study is performed.Secondly feature selection algorithm such as DEFS is used and results are presented.

Introduction

Neuro-Fuzzy System is a hybrid between Neural Network and fuzzy logic, it has the capability of fuzzy systems to adapt to problems in a way humans perceive it and the learning ability of Neural Networks.

If the diamensions in a dataset is too high then the processing time will also be large, so PCA which is a dimensionality reduction algorithm plays a key role by only considering the attributes which have high impact on the results thus increasing efficiency.

The goal of LDA is to reduce the dimensions by removing the reduntant and dependent features by transforming the features from higher dimensional space to a space with lower dimensions.

Independent Component Analysis (ICA) is a statistical technique for decomposing a complex dataset into independent sub-parts.

Differential evolution (DE) is a method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality.

Aim & Objectives

Aim of the project is to analyse a group of algorithms/techniques and find which performed better in terms of accuracy and execution time.This can help in development of critical real life application(ex. biomedical field) where high accuracy and precision is required.

Methodology & Implementation

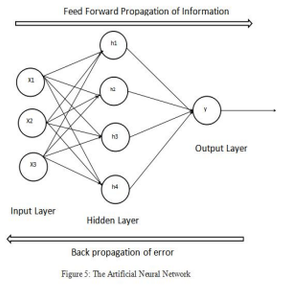


Fig 1: Working of ANN

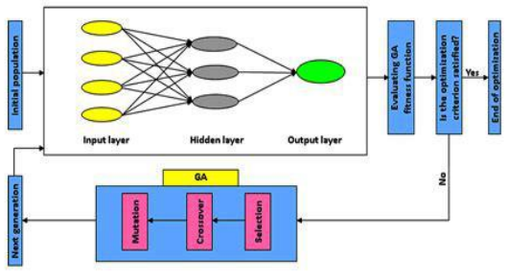


Fig 2: Differential evolution (DE) Algorithm

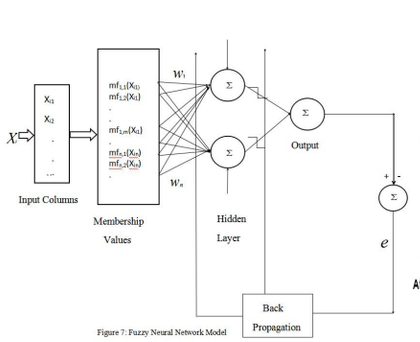


Figure 7: Fuzzy Neural Network Model

Fig 5: NF - Neuro-Fuzzy systems are a hybrid between Neural and fuzzy logic

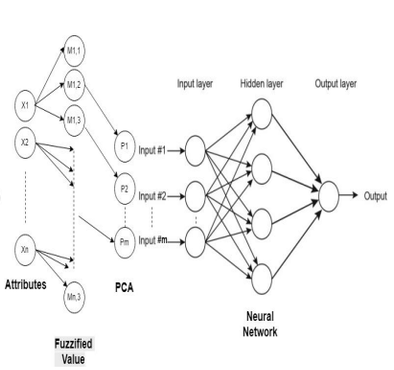


Fig 4: PCA - dimensionality reduction in a model which reduces the noise and redundancy and increases the accuracy of the model.

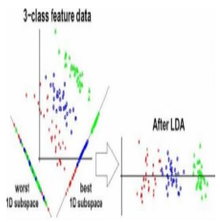


Fig 5. LDA uses both attributes to create a new axis to project the data into axis with max seperation

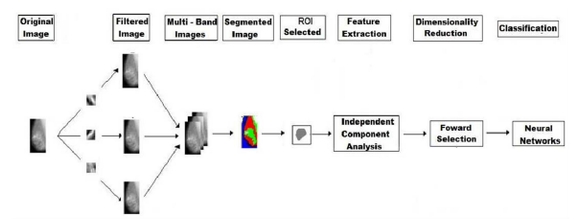


Fig 6 : ICA - type of dimensionality reduction algorithm that carries out the transformation of a set of variables to a new set of components, it is performed by maximizing the statistical independence between the new components

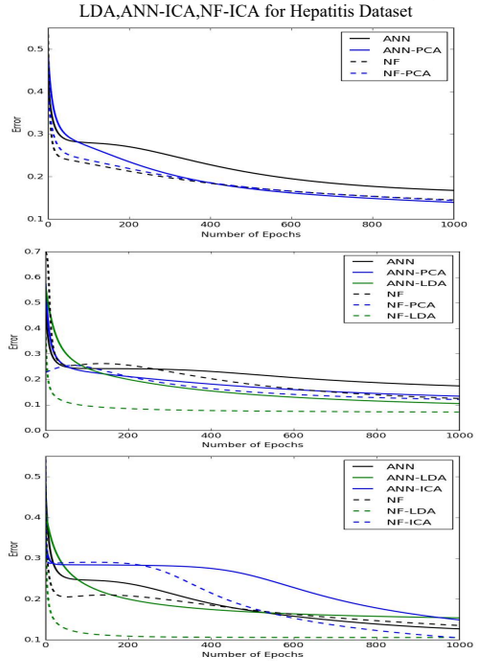
Result Analysis

The values of accuracy,precision,recall,fmeasure of datasets were found.The results of some of the datasets are listed below.

Table:Average Accuracy plot (in %)

	ANN	NF	ANN-PCA	NF-PCA	ANN-LDA	NF-LDA	ANN-ICA	NF-ICA
BCD	90.6	93.9	93.6	95.6	97.2	97.8	95.2	94.6
Wine	91.2	93.1	93.9	95.5	96.5	96.8	93.1	93.2
Hepa	79.1	81.5	81.4	85.6	84.9	88.0	84.2	88.6
Thy	86.7	92.3	92.5	93.1	96.2	96.3	92.8	95.5
Liv	69.2	73.7	68.6	73.9	71.3	73.9	66.2	64.8

Error Plot for ANN,NF,ANN-PCA,NF-PCA,ANN-LDA,NF-LDA,ANN-ICA,NF-ICA for Hepatitis Dataset



Conclusion

From this research it was concluded that

1. NF model performed better than ANN model (Accuracy).

2. Among the feature reduction techniques (PCA,LDA,ICA) the LDA and ICA model had better perfomance than PCA both in terms of accuracy and execution time.

3. The DE optimization algorithm (feature selection technique) perforns better than feature reduction technique for large dataset.

Future works

The use of other optimization algorithms like TLBO, CRO, CTO of feature selection technique and perform their comparative study and find out which will work better for data mining tasks.

Acknowledgements :

We are profoundly grateful to Prof. Himansu Das for his expert guidance and continuous encouragement throughout to see that this project reaches its target since its commencement to its completion.

References

[1] Ghosh, Ashish, B. Uma Shankar, and Saroj K. Meher. "A novel approach to neuro-fuzzy classification." Neural Networks 22, no. 1 (2009): 100-109.

[2] Li, Jia, and Ridge Regression Linear. "Principal component analysis." (2014).

[3] Balakrishnama, Suresh, and Aravind Ganapathiraju. "Linear discriminant analysis-a brief tutorial." Institute for Signal and information Processing 18 (1998): 1-8.

[4] Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE transactions on Neural Networks, 10(3), pp.626-634