relationship. This training can be performed with a gradient descent algorithm similar to the one used for the standard neural network.

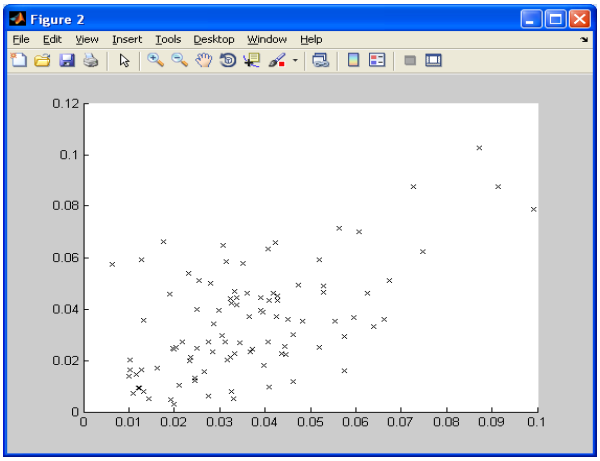

Figure 4.1: Lymphoma dataset

Subspace clustering is an expansion of characteristic selection that tries to find clusters in different subspaces of the same dataset. The Neuro-Fuzzy classifiers were tested on the well-known set of lymphoma data introduced by Fisher which consists of a three class problem based on four parameters of the genes, i.e. the petal length and width and the sepal length and width. One type of the genes can be separated linearly from the two other types whereas the other two types of the genes cannot be separated linearly from each other

5. Experimental Results

In general the Bayes classifier and the KNN[2] classifier could not handle the massive data as good as the Neuro-Fuzzy classification systems. This effect was not experimented with the lymphoma data which can be contributed to the different statistical properties of the two data sets. With the lymphoma data all system had problems to find the similar four outliers which limited the achievable recognition rate to 95.85 %. This well-maintained a smooth relationship between sensitivity, specificity and produced maximum recognition rates.

- Number of selected genes,
- Predictive accuracy on selected genes,
- Extracted knowledge from the trained models.

Table 5.1: Comparison of the classification performance of different Classifiers on lymphoma cancer data set

Methods	*Gene Selection*	*Lymphoma*
Neuro-fuzzy	Information gain	95.85
SVM	Information gain	N/A
	Information gain	72.64

6. Conclusion

Clustering large data sets is a ubiquitous task. Astronomers work to classify stars into similar sets based on their images. Search engines on the web seek to group together similar documents based on the words they contain or based on their citations. Marketers seek clusters of similar shoppers based on their purchase history and demographics.

The task of a Neuro-Fuzzy classification is to provide with a computational solution to the feature selection problem motivated by a certain definition of *relevance*. This algorithm should be reliable and efficient. The many Neuro-Fuzzy classification is proposed in the literature are based on quite different principles (as the evaluation measure used, the precise way to explore the search space, etc) and loosely follow different definitions of relevance.

The Neuro-Fuzzy classification system, which is based on a built clustering algorithm reached recognition rates above in comparison to the Bayes classifier) and the KNN classifier. Our experiment results recommend that Neuro-Fuzzy classification algorithms have the capability a lot to progress common classification systems that can be used in ultrasonic tissue characterization.

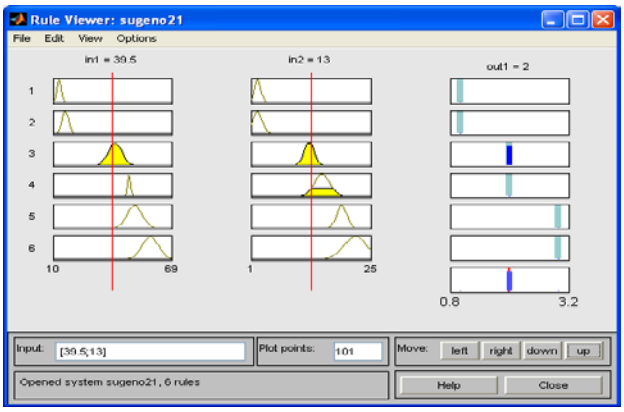

Figure 6.1(a): Fuzzy Neuro classification

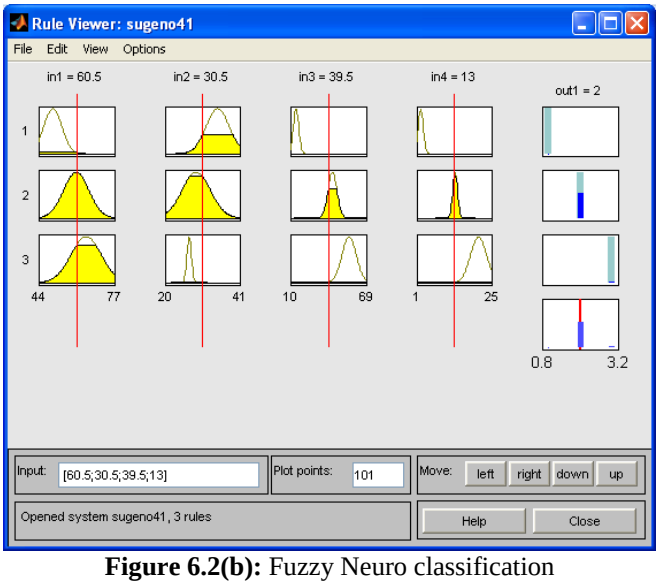

Figure 6.2(b): Fuzzy Neuro classification

Licensed Under Creative Commons Attribution CC BY

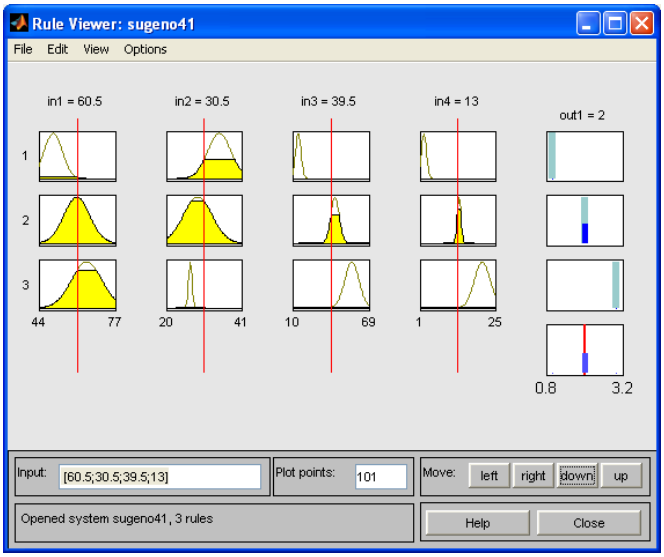

Figure 6.3(c): Fuzzy Neuro classification

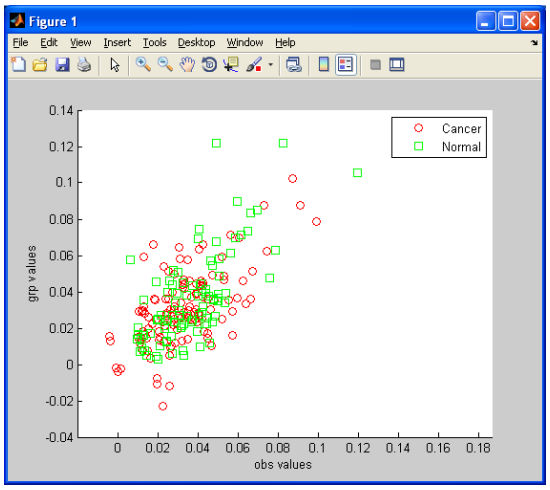

Figure 6.2: Overall data clustering

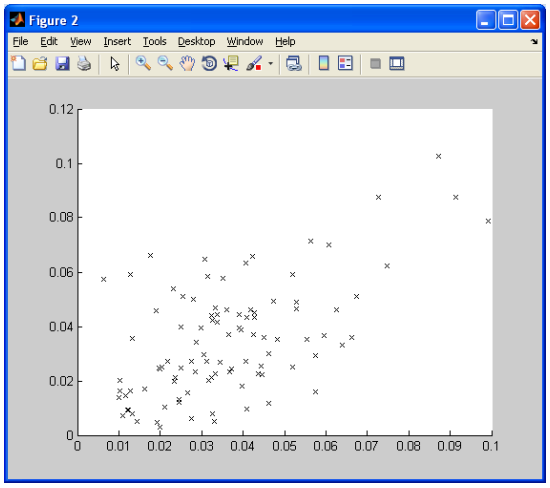

Figure 6.3: Cancer data classification

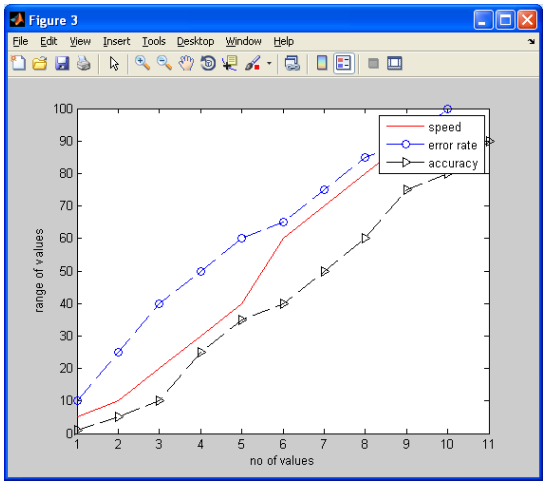

Figure 6.4: Metric range of the graph

7. Scope for Future Work

This proposed Neuro-Fuzzy classification algorithms based on various methods to arrange and classify biological data sets by the development of a interference system. These results prove that Neuro-Fuzzy algorithms have the capability to improve classification methods for the use in ultrasonic tissue characterization In future this study can be stretched in many methods in order to provide better evaluations such as continuous data, missing values, and the use of combined evaluation measures.

References

[1] F. N. Afrati, A. Gionis, and H. Mannila.Approximating a collection of frequent sets. In Proc. 2004 ACM SIGKDD Int. Conf. Knowledge Discovery in Databases (KDD'04), pages 12–19, Seattle,WA, Aug. 2004.

[2] D.L. Boley.Principal direction divisive partitioning.*Data Mining and KnowledgeDiscovery*, 2(4):325–344, 1998.

[3] I.S. Dhillon, S. Malella, and R. Kumar.Enhanced word clustering forhierarchical text classification.In*KDD-2002*, 2002.

[4] S. Cho and H. Won, "Machine learning in dna microarray analysis forcancer classification," in *APBC*, vol. 34, 2003, pp. 189–198

[5] U. Fayyad and K. Irani, "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning," Proc. 13thInt'l Joint Conf. Artificial Intelligence, pp. 1022-1027, 1993.

[6] D.H. Fisher, L. Xu, and N. Zard, "Ordering Effects in Clustering,"Proc. Ninth Int'l Workshop Machine Learning, pp. 162-168, 1992.

[7] Alizadeh, A.A. et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403, 503-511 (2000).

Author Profile

Mr. R. Mahalingam, M.Sc., M.Phil., Department of Computer Science, Kovai Kalaimagal College of Arts & Science, Coimbatore-109, India.

Licensed Under Creative Commons Attribution CC BY

Mr. S. Omprakash., M.Sc., M.Phil., Assistant Professor, Department of Computer Science, Kovai Kalaimagal College of Arts & Science, Coimbatore-109, India.

Licensed Under Creative Commons Attribution CC BY