International Journal of Computer Applications (0975 – 8887)
Volume 19– No.3, April 2011

Classification of Trees by Pattern Recognition

Janmenjoy Nayak
Lecturer
Modern Engineering and Management Studies,Balasore, India

Ashanta Ranjan Routray
Lecturer
Fakir Mohan University
Balasore,India

Flt. Lt. Dr. Munesh Chandra Adhikary
Reader
Fakir Mohan University
Balasore,India

ABSTRACT
Pattern recognition has become more & more popular and it induces attractive attention coming from a wider areas. Pattern recognition is able to describe the actual problems via mathematical models. In this paper, we are classifying various south east trees by the help of pattern recognition. The training data are the cone and needle length for which we can easily classify the corresponding trees. This paper reviews the basic pattern recognition procedures for classifying various trees by taking their patterns.

General Terms
Pattern recognition, Dram Stack, Casuarina, Castor, Deodar

Keywords
Needle, Cone, Z-score Normalization, Feed forward Neural Network

1. INTRODUCTION
Pattern recognition is an interdisciplinary subject, covering developments in the areas of statistics, Engineering, Artificial Intelligence, Computer Science, Psychology and Physiology etc. Pattern recognition is a field concerned with machine recognition of meaning regularities in noisy of complex environments [1]. The word Pattern is derived from the same root as the word patron, which means something which is set up as a perfect example to be imitated [2]. It is a classification of input data via extraction important features from a lot of noisy data [3]. Pattern recognition includes a lot of methods, which impelling the development of numerous application in different field. The structure of a simple pattern classifier is given in figure 1.1.

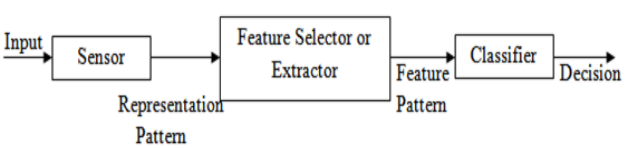

Figure 1.1 Pattern Classifier

Complete pattern recognition system consists of:

a) A sensor: It gathers the information to be classified.

b) A feature selector or Extractor: Feature selection is the process of selecting a subset of a given set of variables. The feature extractor mechanism takes a possible non linear combination of the original variables to form new variables.

c) A classifier: It classifies or describes the observations relying on the extracted features.

The aim of pattern classification is to utilize the information acquired from pattern analysis to discipline the computer in order to accomplish the classification. The composition of a pattern classification system [1] is shown in figure 1.2.

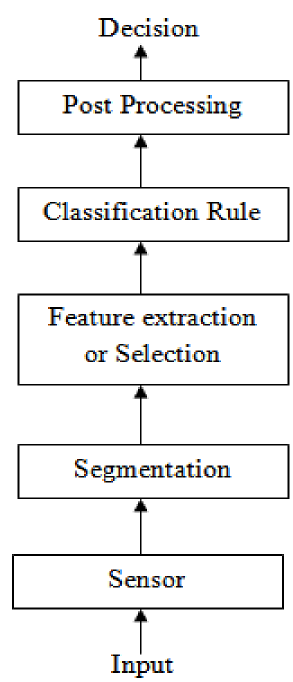

Figure 1.2 Components of a Pattern Recognition System

2. EXTRACTION OF FEATURES
Pattern recognition involves taking an input and mapping it to a desired recognition class[4]. We will first illustrate the pattern recognition process with an example that uses cones and leaves to identify four different species of evergreen trees in the south east (figure 1.3).

This figure shows the features to be used to identify the species are needle length & cone length. The desired outputs are the four species: Dram Stack, Casuarina tree, Castor Plant and Deodar tree. We begin by collecting samples of needles and cones for each species and measuring the lengths, resulting in the data shown in table 1.

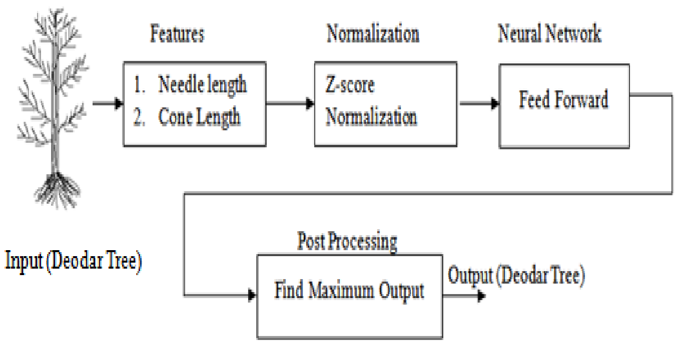

Figure 1.3 Pattern recognition for distinguishing four evergreen trees in south east

Tree	Raw Data		Normalized Data	
	Cone Length	Needle Length	Cone Length	Needle Length
Casuarina	35 mm	14 mm	-0.23	-0.28
Casuarina	36 mm	14 mm	-0.29	-0.37
Casuarina	34 mm	11 mm	-0.33	-0.42
Casuarina	37 mm	17 mm	-0.26	-0.36
Castor Plant	20 mm	19 mm	-0.43	-0.32
Castor Plant	22 mm	22 mm	-0.45	-0.30
Castor Plant	21 mm	14 mm	-0.46	-0.31
Castor Plant	23 mm	16 mm	-0.47	-0.05
Dram Stack	43 mm	23 mm	0.09	0.42
Dram Stack	44 mm	22 mm	0.41	0.43
Dram Stack	45 mm	26 mm	0.42	0.43
Dram Stack	42 mm	22 mm	-0.12	0.34
Deodar	53 mm	28 mm	0.47	-0.07
Deodar	50 mm	32 mm	0.46	-0.11
Deodar	51 mm	33 mm	0.44	-0.04
Deodar	55 mm	31 mm	0.43	-0.02

Table 1. Raw & normalized data for distinguishing four evergreen trees in south east

Here we have taken the features of the trees like needle length and cone length. We have taken 340 samples from several trees. For testing these samples split the collections of the data for the test set.

The use of good features cannot be emphasized enough. The vast majority of the work to be performed in pattern classification to obtain a good set of features[5].

Once that is accomplished, the classifier system used is not overly critical. Once the data are collected normalized and the associated class identified for each feature vector, the data are presented to a neural network are initially set to small random values[6].

3. NORMALIZATION

Here the raw data are the cone length and the needle length of four different trees. For training set we have taken 220 samples (68% of total), for validation 110 samples (30% of total). The raw data have been normalized to assist the neural network during training. The feature space for the data collected for this application example is depicted in figure 1.4.

This demonstrates that there is virtually no overlap in the feature space among the four species to be identified. This means that the classifier will not have to work hard to separate the various species from each other and then to use that information to identify a given feature vector with its associated tree species. Here the data are normalized 4000 iterations by using the z-score normalization.

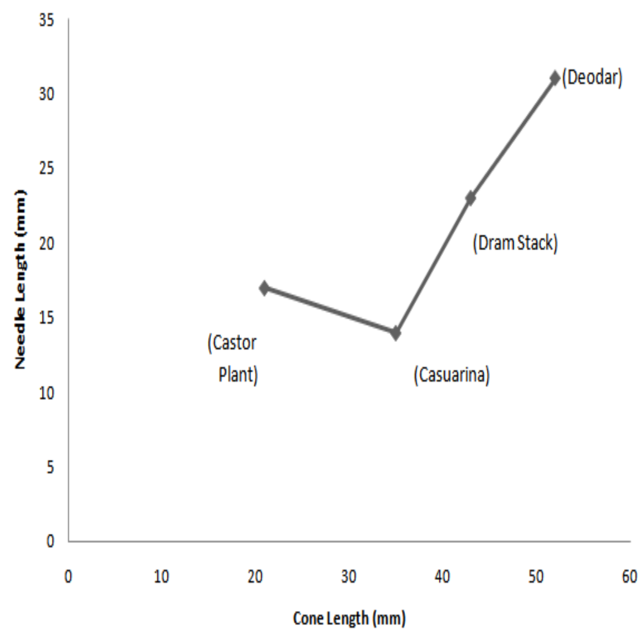

Figure 1.4 Feature space for distinguishing different evergreen trees in south east

4. NEURAL NETWORK

The artificial neural network paradigm is constantly motivated and constrained by neuronal analogies, rather than by the attempt to model real neurons. Artificial neural networks are fine-grained parallel implementations of non linear static or dynamic systems. They are generally adaptive in nature, means "learning by example" replaces "programming" in solving problems[9].

A neural network is first and foremost a graph, with patterns represented in terms of numerical values attached to the nodes of the graph and transformations between patterns achieved via simple message-passing algorithms[10]. Certain of the nodes in the graph are generally distinguished as being input nodes or output nodes, and the graph as a whole can be viewed as a representation of a multivariate function linking inputs to outputs. Numerical values (weights) are attached to the links of the graph, parameterizing the input/output function and allowing it to be adjusted via a learning algorithm.

The statistical approach of neural network helps in understanding the capabilities and limitations of network models and in extending their range. Neural networks can be viewed as members of the class of statistical models known as "nonparametric," and the general theory of nonparametric statistics is available to analyze network behavior[11]. It is also of interest to note that many neural network architectures have close cousins in the nonparametric statistics literature; for example, the popular multilayer perceptron network is closely related to a statistical model known as "projection pursuit," and the equally popular radial basis function network has close ties to kernel regression and kernel density estimation

A more thoroughgoing statistical approach, with close ties to "semiparametric" statistical modeling, is also available in which not only the input and output nodes of a network but also the intermediate ("hidden") nodes are given probabilistic interpretations. The general notion of a mixture model, or more generally a latent variable model, has proved useful in this regard; the hidden units of a network are viewed as unobserved variables that have a parameterized probabilistic relationship with the observed variables (i.e., the inputs, the outputs, or both). This perspective has clarified the links between neural networksand a variety of graphical probabilistic approaches in other fields; in particular, close links have been forged with hidden Markov models, decision trees, factor analysis models, Markov random fields, and Bayesian belief networks.[7]

These links have helped to provide new algorithms for updating the values of nodes and the values of the weights in a network; in particular, EM algorithms and stochastic sampling methods such as Gibbs sampling have been used with success[12].

Neural networks have found a wide range of applications, the majority of which are associated with problems in pattern recognition and control theory[13]. Here we give a small selection of examples, focusing on applications in routine use.

The problem of recognizing species of trees is a challenging one that has been widely studied as a prototypical example of pattern recognition. Some of the most successful approaches to this problem are based on neural network techniques and have resulted in several commercial applications[8]. Mass screening of forest images is another area in which neural networks have been widely explored, where they form the basis for one of the leading systems for semi-automatic interpretations. As an example of pattern recognition we mention the problem of verifying the patterns of trees, based on the dynamics of the species captured by sensor, where the leading approach to this problem is again based on neural networks.

Practical Considerations for Back-Propagation Learning

Most of the practical considerations necessary for general Back-Propagation learning of single layer Perceptions are:

1. Do we need to pre-process the training data? If so, how?
2. How do we choose the initial weights from which we start the training?
3. How do we choose an appropriate learning rate h?
4. Should we change the weights after each training pattern, or after the whole set?
5. Are some activation/transfer functions better than others?
6. How can we avoid flat spots in the error function?
7. How can we avoid local minima in the error function?
8. How do we know when we should stop the training?

However, there are also two important issues must be considered

9. How many hidden units do we need?
10. Should we have different learning rates for the different layers?

How Many Hidden Units?

The best number of hidden units depends in a complex way on many factors, including:

1. The number of training patterns
2. The numbers of input and output units
3. The amount of noise in the training data
4. The complexity of the function or classification to be learned
5. The type of hidden unit activation function
6. The training algorithm

Too few hidden units will generally leave high training and generalisation errors due to under-fitting. Too many hidden units will result in low training errors, but will make the training unnecessarily slow, and will result in poor generalisation unless some other technique (such as regularisation) is used to prevent over-fitting. Virtually all species of trees you recognize about are actually nonsense. A sensible strategy is to try a range of numbers of hidden units and see which works best.

The neural network used for this classifier is feed forward network in which two inputs are given for generating four outputs. The learning method is supervised. The patterns are trained by using the back propagation technique[14]. The two inputs to the neural network are cone length and needle length and the corresponding outputs at the outer layer of the neural network are Casuarina, Castor plant, Dram Stack and Deodar. The weights in the neural network are initially set to small random values. We have placed the initial hyperplanes used by the feed forward neural network on the feature space map as shown in figure 1.5.

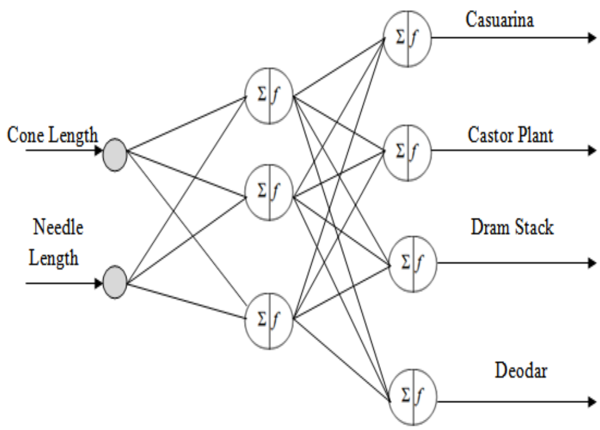

Figure 1.5 Neural Network used to classify four evergreen trees

As the neural network is trained, the hyperplanes shift and begin to carve out decision regions. Feed forward neural networks, with sigmoidal transfer functions, use the hyperplanes to form what mathematicians term half-spaces i.e. one side of the hyperplane contains one decision region and the other side contains the not of that decision region[17]. The hyperplanes that define the various tree species do not have limits in the feature space. This means that any new tree species introduced will be recognized as one or more of the learned classes. While these patterns were solved using a feed forward neural network employing hyperplanes, [18] it could just easily been solved with a clustering algorithm such as Adaptive Resonance Theory (ART). The distance metrics for determining the range of support for each cluster center we have used the Euclidean distance. (Figure 1.6)

5. POST PROCESSING

In post processing the threshold is applied to each output followed by a maximum picker to find the most likely classes of the patterns[15]. After testing all the patterns the output layer gives the clear distinguishable feature spaces of four different trees for which the result can be hundred percent. Beside taking the Euclidean distance, if the result sets will be tested by using some different distance metrices like (Mahalanobis) for separating the clusters then the result may be slightly different[16].

6. CONCLUSION

From the above discussion, it may be concluded that pattern recognition is the heart of all scientific inquiry, including understanding ourselves and the real-world around us. And the developing of pattern recognition is increasing very fast, the related fields and the application of pattern recognition became wider and wider. In this paper we expatiate pattern recognition in the round, include the definition of Pattern Recognition, the applications of Pattern Recognition, and finally how the Pattern Recognition can be frequently used for identifying different trees from their collected species. Also Pattern Recognition can be used to recognize the handwritten numbers, the electronic nose, the airport scanner texture recognition etc. In addition, it is an important trend to use pattern recognition on engineering; we should make efforts on this. And pattern recognition scientists should pay attention to new technique of Pattern Recognition, and enlarge the application areas of Pattern Recognition.

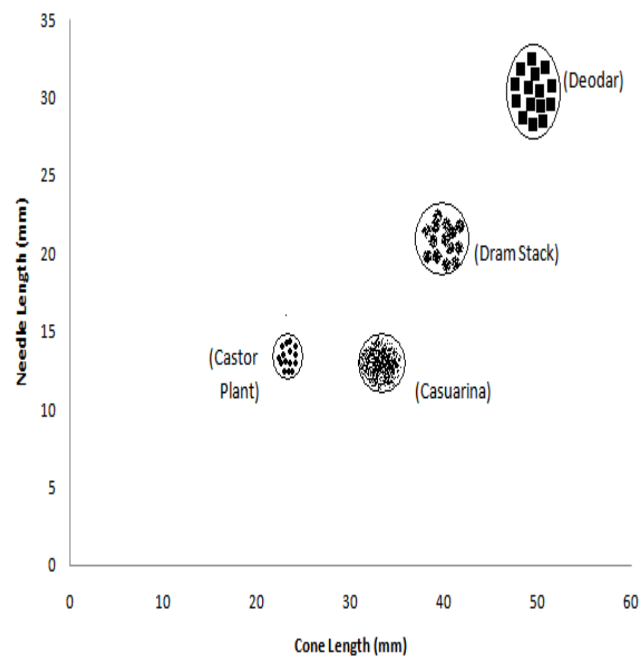

Figure 1.6 Clustering Neural Network Solutions with Euclidean distance metrics

7. REFERENCES

[1] Editorial, Advances in Pattern Recognition, Pattern Recognition Letters 26,395-398, 2005

[2] T. Pavlidis. Structural Pattern Recognition. Springer Verlag, Berlin Heidelberg New York, 1977.

[3] Gonzalez, R.C.Thomas, M.G. Syntatic Pattern Recognition: An Introduction, Addison Wesley, Reading, MA, 1978

[4] Watanabe, Pattern Recognition: Human and Mechanical.Wiley, New York, 1985

[5] Srihari, S.N., Covindaraju, Pattern recognition, Chapman & Hall, London, 1034-1041, 1993

[6] B. Ripley, Pattern Recognition and Neural Networks, Cambridge University Press, Cambridge, 1996

[7] Towell G., Shawlik J.W. (1994): Knowledge-based artificial neural networks. Artificial Intelligence, 70 (1-2): 119-165 (October).

[8] M.petrou, Learning in pattern recognition: some thoughts, Pattern Recognition Letters 22,3-13, 2001

[9] Frank T.Allen*,Jason M.kinser, H. John Gaulfield, A neural bridge from syntactic pattern recognition to statistical pattern recognition ,Neural Networks 12, 519-526, 1999

[10] B. Verma (1997), "Fast Training of Multilayer Perceptrons." IEEE Transactions on Neural Networks 8 (6): 1314-1320.

[11] D. W. Ruck, S. K. Rogers and M. Kabrisky (1990a). "Feature selection using a multilayer perceptron." Journal of Neural Network Computing 2 (2): 40-48.

[12] Robert, L.G., 1965, Machine perception of three-dimensional solids. Tippett,J.T. et al.(Eds), Optical and Electro Optical Information Processing. MIT Press,Cambrige,MA, pp.159-197,1965

[13] Guzman, A. Decomposition of a visual scene into three dimensional bodies , AFIPS Conference Proceeding s. Thomplson,Washington,DC, Vol 33,pp.291-304,1968

[14] Alan Wee_Chung Liew, Hong Yan,Mengsu Yang, Pattern Recognition techniques for the emerging field of bioinformatics:A review, Pattern Recognition, 2005

[15] Pattern recognition applied to mineral characterization of Brazilian coffees and sugar-cane spirits, Mirian C. santos,Sherlan G. Lemos, Ana Rita A. Nogueira , SPECTROCHIMICA ACTA PART B ,2005

[16] Bezdek, J.C. and S.K. Pal, Eds , Fuzzy Models for pattern recognition. IEEE Press New York, 1992

[17] Keller,J.M., Fuzzy sets in computer vision. Proc. VI IFSA World Congress, Sap Paulo, Brazil, Vol.I, 7-10, 1995

[18] Marr, D.Vision. W.H. Freeman, San Francisco, CA, 1982 Witold pedrycz, Special issue on fuzzy set technology in pattern recognition, Pattern Recognition Letters, 1996.