Pattern Classification Methods: A Survey

Bhavana H T
Assistant Professor, Dept. of ECE
BMS College of Engineering
Bangalore

Srikanth H T
Application system Engineer
Wells Forgo India Pvt. Ltd.
Bangalore

DOI: 10.29322/IJSRP.9.08.2019.p9299
http://dx.doi.org/10.29322/IJSRP.9.08.2019.p9299

ABSTRACT

In this paper overview of few pattern classification methods and comparison of some aspects of pattern recognition system are presented. The paper also contains the research topics and applications of pattern recognition system. Pattern recognition is the research area that studies the operation and design of systems that recognize patterns in data. New and emerging applications, such as data mining, web searching, retrieval of multimedia data, face recognition, and cursive handwriting recognition, require robust and efficient pattern recognition techniques.

Keywords: *Pattern Recognition, classification, probabilistic*

1. INTRODUCTION

Pattern recognition is finding and labeling smaller desired patterns in a big pattern. Pattern recognition is emerging as exciting and challenging field; in this paper discussion and comparison of some features of pattern recognition are presented. The basic types of pattern recognition are supervised and unsupervised classification [1]. The neural network techniques extracted from statistical learning theory are getting more attention in recent pattern recognition methodologies. The following issues need to be considered while designing a pattern recognition system: definition of pattern classes, sensing environment, pattern representation, feature extraction and selection, classifier design and learning, selection of training and test samples and performance evaluation [2]. The most common problems in the pattern recognition field are identifying complex patterns with arbitrary orientation, location, and scale remains unsolved [3].

The organization of this paper is as follows: In section 2, information on several classifiers is given focusing on Statistical Pattern recognition system. The classifiers are categorized according to the design methodology; as using similarity maximization, probability, and geometric information on deciding. Later in section 3 the different pattern classification techniques are discussed and in section 4 comparison of pattern classification methods are discussed along with conclusion [4].

2. STATISTICAL PATTERN RECOGNITION

A number of commercial recognition systems have been designed based on statistical pattern recognition [5]. In this method a set of d features (attributes) are used to represent a pattern, the pattern can be viewed as d-dimensional feature vector. Statistical theory concepts are used to create decision boundaries between pattern classes. Training and Classification are two modes in statistical recognition system. The preprocessing module segments the pattern of interest from the background, removes noise and normalizes the pattern. Preprocessing unit will do all operations to define clear representation of the pattern. The feature extraction/selection module will search for the appropriate features to represent the input patterns and to train the classifier. The designer can optimize the preprocessing and feature extraction/selection strategies using the feedback path [6]. Finally in classification mode the input patterns

are categorized as one of the pattern classes based on the measured features. Fig. 1 gives the components of a statistical pattern recognition system.

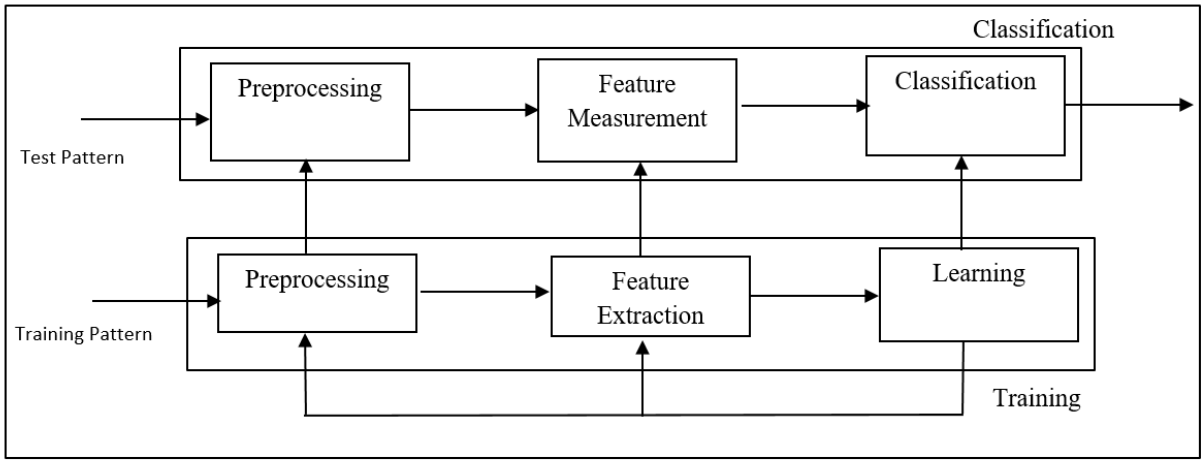

Figure 1: The components of a statistical pattern recognition system.

Several iterations are given for training to optimize the selected features. The ability to classify the pattern correctly depends upon the output of the feature extraction algorithm. The input data also contributes to the quality of extracted features. The main aim of the classifier is to group objects in a category based on variance in their measured features. Objects in the same category possess similar feature values and in different categories the similarity reduces. The classifiers can handle several inconsistencies in feature sets; still each classifier has its pros and cons when dealing with precise feature sets [7].

Another contrast learning method in statistical pattern recognition is that of supervised learning (labeled training samples) versus unsupervised learning (unlabeled training samples).The category of a pattern is represented by its label. In unsupervised learning the number of classes must be learnt along with the structure of each class. The several categories that belong to statistical pattern recognition are shown in fig. 2.

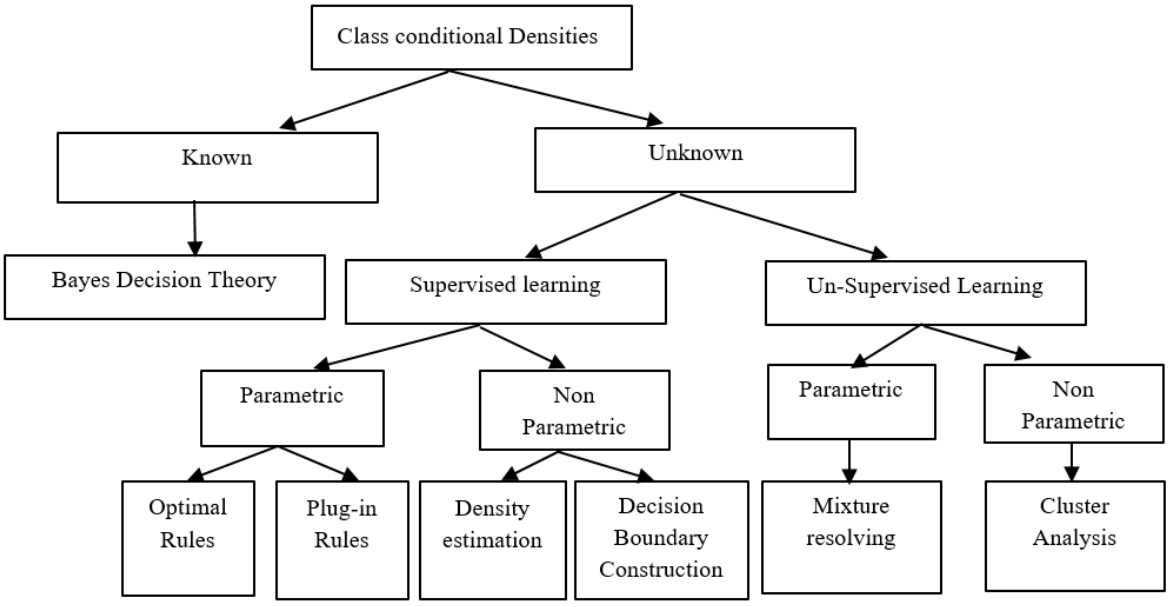

Fig. 2. Various approaches in statistical pattern recognition.

The information available to the system designer reduces as we move from top to bottom and left to right in fig. 2. This results in increased difficulty in classification problems. Several categories in statistical pattern recognition try to implement Bayes decision

rule. In the nonparametric and unsupervised learning mode cluster analysis handles decision making problems, where the number of categories or clusters is not specified [8].

The statistical pattern recognition can be further classified as geometric approach and probabilistic density-based approach based on whether the decision boundaries are obtained directly or indirectly as shown in Fig. 2. In probabilistic approach the density functions have to be estimated first, and then discriminant functions which specify the decision boundaries have to be constructed. The geometric approach constructs decision boundaries directly from optimizing certain cost functions. All classification or decision rules require precise training using available training samples. Hence the competence of classifier depends on both the number of available training samples and specific values of the samples. Finally the pattern recognition system should classify future test samples which are probable to be different from the training samples. Therefore optimizing classifier on the training set may not always result in the desired performance on a test set [9].

The efficiency of pattern recognition system in classifying test patterns which were not used during the training stage is known as generalization ability of classifier. The following are some of the reasons for poor generalization ability of a classifier.

1) The number of features is too large relative to the number of training samples [10],
2) The number of unknown parameters associated with the classifier is large,
3) A classifier is too intensively optimized on the training set (overtrained);

The novel work by Cover [5] on capacity of classifier and complexity provides a good understanding of mechanisms by overtraining.

3. PATTERN CLASSIFICATION

Basically there are three types in classifiers [11]. In all types the training set and the feature selection algorithm decides outcomes. Each classifier has its own advantage. These classifiers depend upon similarity maximization methods, probabilistic methods, and geometric methods, respectively.

3.1 Similarity Maximization Methods

In this method classifiers will have some similarity metrics and assign class labels for maximizing the similarity. The similarity between patterns is used to decide on good classification. The nearest mean classifiers define the features of a class as a vector and represent the class with the mean of the elements of this vector. Thus, any unlabeled vector of features will be classified as the class with nearest mean value. This method finds problem in classifying face images as it needs to supply a template for each face label.

Nearest neighbor Algorithm [7, 8] is another important classifier where the data is represented as points in space, and classification is done based on the Euclidean distance of the data to the labeled classes. For the k-NN, the classifier checks the k nearest points and decides in favor of the majority

3.2 Probabilistic Methods

Bayes decision theory is used extensively in well-known probabilistic methods. The decision rule assigns class labels to samples which have maximum posterior probability. The posterior can be calculated by the well-known Bayes rule:

$$posterior = \frac{likelyhood\ X\ prior}{Evidence} \quad (1)$$

Bayesian Belief Nets [12, 13, and 14] represent the functional dependencies and independencies among model variables, i.e. features. Whenever some parameters take some values, the nodes of the network are affected and take a probability value, by the Bayes' rule [12].

3.3 Geometric Classifiers

Geometric classifiers build decision boundaries by directly minimizing the error criterion, since no related experiments are supplied. An example to these classifiers is Fisher's linear discriminant, which mainly aim to reduce the size of the feature space to lower dimensions in case of a huge number of features. It minimizes the mean squared error between the class labels and the tested instance. Additionally, neural networks are examples of geometric classifiers [13]

3.4 Decision Trees

The decision trees, take the instance described by its features as input, and outputs a decision, denoting the class information. Each node denotes a feature, and each iteration we go down to the lower depth, selecting a child node depending on the feature value

for the particular instance. There are several issues of decision trees, such as how to create a good one. There are several decision tree classifiers such as ID3 or C4.5 .

4. CONCLUSION

Pattern recognition is generally categorized according to the type of learning procedure used to generate the output value. In simple sense pattern recognition is the heart of all scientific inquiry, including understanding ourselves and the real-world around us. Now a day the development of pattern recognition is increasing very fast.

A survey on the pattern recognition has been presented. It has been shown that powerful methods exist, however, care has to be taken to build robust and consistent classifiers. The best approach for the inexperienced user seems to be the use of classical statistical tools, since plug and play works in this case. Pattern recognition can be done both in normal computers and neural networks. Computers use conventional arithmetic algorithms to detect whether the given pattern matches an existing one. It is a straightforward method. It will say either yes or no. It does not tolerate noisy patterns.

References

1. R. O. Duda, D. G. Stork, and P. E. Hart. Pattern Classification and Scene Analysis. John Wiley & Sons, Inc., 2nd edition, 2000.

2. S. Russell and P. Norvig. Artificial Intelligence - A Modern Approach. Pren-tice Hall, second edition, 2003

3. A. K. Jain, R. P. W. Duin, and J. Mao, Statistical Pattern Recognition: A Review, IEEE Trans. on Pattern Analysis and Machine Intelligence, 22(1):4- 37, January 2000.

4. S. R. Kulkarni, G. Lugosi, and S. S. Venkatesh, Learning Pattern Classifica-tionA Survey, IEEE Transactions Information Theory, vol. 44, no.6, O1998

5. T. M. Cover and P. E. Hart, Nearest neighbor pattern classification, IEEE Transactions Information Theory, vol. IT-13, pp. 2127, 1967.

6. Y. Chenyz Y. Hungyz C. Fuhz, Fast Algorithm for Nearest Neighbor Search Based on a Lower Bound Tree, to appear in Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada, July 2001.

7. W. L. Buntine, Operations for Learning with Graphical Models, Journal of
Artificial Intelligence Research, 2:159-225, 1994.

8. J. A. Bilmes, A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models, Technical Report TR-97-021, International Computer Science Institute, University of California, Berkeley, April 1998.

9. J. A. Anderson, Logistic Discrimination, Handbook of Statistics. P. R. Krish naiah and L. N. Kanal, eds., vol. 2, pp. 169-191, Amsterdam: North Holland, 1982.

10. E. Frank, M. Hall, and L. Trigg, Weka 3 - Data Mining with Open Source Machine Learning Software in Java.http://www.cs.waikato.ac.nz/ml/weka/index.html.

11. C. L. Blake and C. J. Merz, UCI Repository of machine learning databases, http://www.ics.uci.edu/~mlearn/MLRepository.html, University of Califor nia, Irvine, Dept. of Information and Computer Sciences, 1998.

12. http://www.diesel- ebooks.com/item/9780470845134/Webb-Andrew-R.-Statistical-Pattern-Recognition/1.html

13. http://media.wiley.com/product_data/excerpt/39/04708451/0470845139.pdf