VOL. 9, NO. 8, AUGUST 2014 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

PATTERN RECOGNITION TECHNIQUES: STUDIES ON APPROPRIATE CLASSIFICATIONS

Sasan Karamizadeh, Shahidan M. Abdullah, Mazdak Zamani and Atabak Kherikhah
Advanced Informatics School (AIS), Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
E-Mail: ksasan2@live.utm.my

ABSTRACT

Pattern recognition techniques are divided into categories of supervised, unsupervised and semi supervised. Supervised pattern recognition methods are utilized in the examination of various sources' chemical data such as sensor measurements, spectroscopy, and chromatography. The unsupervised classification techniques use algorithms to classify and analyze huge amounts of raster cells. Semi-Supervised Learning is an approach that is in the middle ground between supervised and unsupervised learning and guarantees to be better at classification by involving data that is unlabeled. In this paper, we tried to categories pattern recognition methods and explain about each of them and we compared supervised method with unsupervised method in terms of types and location of features.

Keywords: pattern recognition, semi-supervised learning.

INTRODUCTION

Pattern recognition techniques are divided into categories of supervised, unsupervised and semi supervised. This is dependent on the analyst's intention of the information that needs to be utilized or that is available regarding the samples comprising of the data matrix. In the supervised methods, or the classification method, prior description is made on the classes as the concept or the attribute used to classify the samples into subsets are already known [1]. In the unsupervised method, the classification is removed by considering only the variations and resemblances among the samples, without utilizing any of their details. The semi-supervised method is in the middle ground between the supervised and unsupervised analysis and assures to be a better classification using the non-labeled details [2].

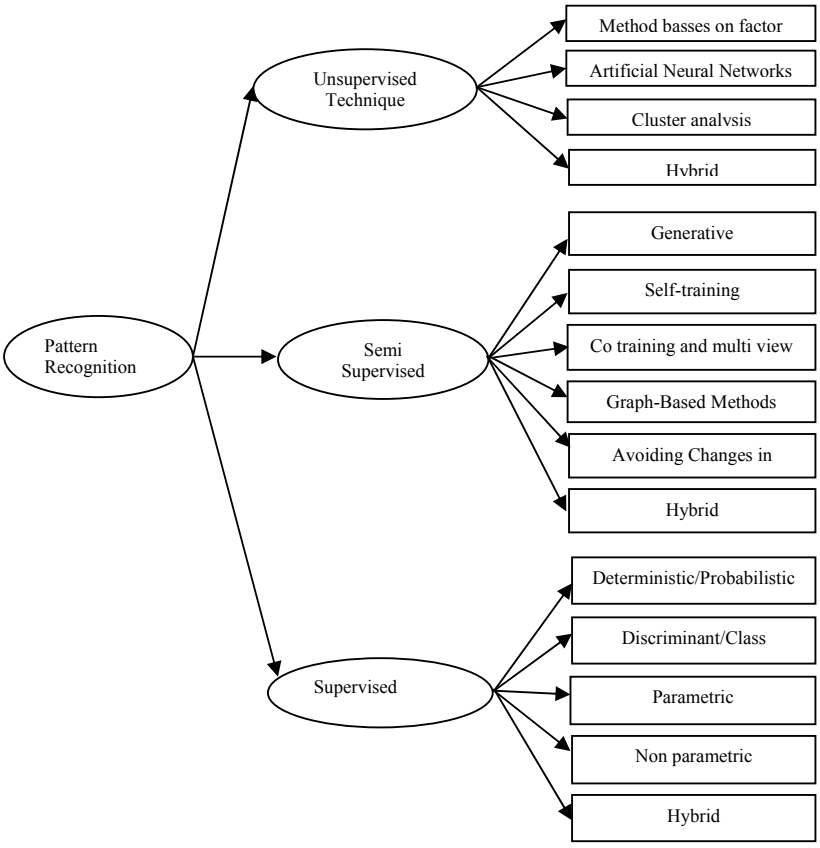

Figure-1. Classification of Pattern-Recognition Techniques.

SUPERVISED METHODS

Supervised pattern recognition methods are utilized in the examination of various sources' chemical data such as sensor measurements, spectroscopy, and chromatography. Various supervised techniques exist which have been widely utilized in the analytical chemistry [2]. In all the cases, the most suitable method is reliant on the problem that needs to be addressed since the criteria and bases of the techniques differ significantly according to the problems faced. As revealed in Figure-1, different criteria can be utilized to apply the supervised methods. Several of the common methods are elaborated below:

Parametric and non-parametric methods

Metric methods utilize the mathematical models that have parameters that can be adjusted to perform classification of samples. These methods involve SIMCA (Soft Independent Modeling of Class Analogy), LDA (Linear Discriminate Analysis), DA (Discriminate Analysis), and SVM (Support Victor Machine) [3].

Discriminate and class modeling assessment

Supervised pattern recognition methods differentiate the variables' hyperspaces that distinguish the samples into various classifications. Utilizing the discriminant methods, when a new sample is put into the hyperspace classifications, it is identified with that classification, however, when it is put outside, this does not happen. There is a lack of an in-between or middle ground [4]. The techniques used here are KNN, LDA, ANN, and DA. The analysis on class modeling considers the samples that fit the model as part of the class, whereas rejected non-members are the objects that do not fit. In the event of modeling more than one class, three various circumstances can be identified; for example every sample can be designated into a single classification, or more than one classification or none of the classifications [1].

Deterministic/ probabilistic methods

No statement is made regarding the reliability of the decision when a deterministic system is utilized to designate each sample's class. Probabilistic methods, however, do measure the classification's reliability. Deterministic methods are namely KNN, and the probabilistic methods are namely DA, LDA and ANN [1].

UNSUPERVISED METHODS

The unsupervised classification techniques use algorithms to classify and analyze huge amounts of raster cells. These procedures need set values for several of the operating parameters; however, the classifying method goes on without any intervention from the users. The efficacy of the unsupervised techniques is dependent on the basis that the input raster dataset has natural statistical groups of spectral patterns that consist of specific forms of physical characteristics [1]. The entire unsupervised classification techniques, aside from the Simple One-Pass Clustering, utilize the interactive procedure to examine a set of sample input cells and decide on the set of class centers and other related statistical features [5]. All the input raster set is then analyzed, and a classification rule is utilized to designate each raster cell to a defined class. The techniques that are normally utilized can be classified into four significant groups as revealed in Figure-1.

Cluster Analysis (CA)

Cluster Analysis or CA has been the most commonly utilized used method of pattern recognition until several years ago. This particular method designates samples to similar clusters based on the level of similarity in the variables (characteristics) which have been utilized to identify the objects, and, simultaneously designate samples that are dissimilar to various other clusters. It is commonly utilized to design a new category of samples being studied study; however, it can also be used to confirm an existing group. Researchers have introduced a complete monograph of cluster analysis based on analytical chemical data [6].

Artificial Neural Networks (ANN)

ANNs are mathematical methods that follow the workings of the nervous system in humans, by making up pattern recognition models. They are normally very successful in addressing the challenges often faced during the process of classification. ANNs begin from a data training set, that contains characteristics such as spectra or concentration levels that measure samples that are indifferent, to measure probabilities of samples that are a class member (output variables)[3]. ANNs are utilized in both supervised and unsupervised pattern identification, however, since their usage not as simple as CA, their usages are limited somewhat [4].

Techniques according to factor models

The aim of these techniques is to constrain the n-dimensional information about objects to a limited and more inclusive aspect. In this way, all the samples can be depicted graphically in a two or three-dimensional (2D or 3D) arena, simplifying identification of the major characteristics. PCA (Principal Component Analysis) are some of the commonly used methods that operate the 2D data tables and multi-set methods [5].

SEMI SUPERVISED

Semi-Supervised Learning (SSL) is an approach that is in the middle ground between supervised and unsupervised learning and guarantees to be better at classification by involving data that is unlabeled. Since getting labeled data is costly and complicated, by causing unlabeled data to be less expensive to get in many applications [6], SSL tries to gain better classification function by utilizing both unlabeled and labeled data. The self-training technique is one of the first algorithms suggested for utilizing the unlabeled data. Another two significant methods include the transductive S3VM and co-training.

Generative

One of the oldest semi-supervised learning techniques is the generative models. The model's assumption is that p(x, y) = p(y) p (x|y) where p (x|y) is a distribution of identifiable mixture; for instance the Gaussian mixture models. Having a huge sum of unlabeled data, the mixture elements can be recognized; and typically, only one labeled example per element is needed to completely decide on the distribution of the mixture [7].

Self-training

Self-training or decision-directed or self-labeling learning is the easiest and frequently used SSL approach. This wrapper algorithm utilizes the forecasting of a supervised learning technique to label the unlabeled data. The classifier utilizes its own forecasts to teach itself. Initially, it begins by training a separated hyper plane with only the labeled data. At every stage, the algorithm chooses a portion of the unlabeled samples for labeling, based on the target or a decision task. After that, the technique adds on these objects to the set of training. Lastly, the classifier retrains itself and the procedure is repeats once more [8].

The self-learning algorithm is easy and can be utilized as an algorithm for meta-learning. However, it depends on the goodness- of-fit of the obtained classifier, taking into consideration that errors tend to strengthen themselves. One other drawback of self-learning is the complexity of examining it generally, but there have been several researches on the convergence of particular base learners [9]. Self-training will be used as one of the semi-supervised tactics to develop the models on credit scoring.

Co-training and multi view learning

Co-training techniques depend on three assumptions. Firstly, it is stated that must be a natural variables split in two of the subsets. Secondly, every subset must be large enough to train a good classifier. Lastly, the technique presumes that both of the subsets are conditionally independent considering the class. This technique trains two various classifiers; one for each subset and uses just the labeled data. After that, each of the classification tasks categorizes a portion of the unlabeled data and trains the other classifier. Both classifiers will be retrained using this new labeled data handed out by the other classifier (cross information) in an iterative method [10].

Graph-based methods

Semi-supervised techniques that are graph-based refer to a graph where the nodes consist of both the labeled and unlabeled samples in the dataset, and edges (may be weighted) show the samples' similarity. These techniques normally assume that there is smoothness of label across the graph. Graph techniques are non-parametric, transductive and discriminative [11, 12].

Avoiding changes in dense areas

Discriminative techniques function directly on p (y|x). This causes the danger of leaving p(x) outside the parameter's estimation loop, if p(x) and p (y|x) do not have similar parameters. Normally, p(x) is all that can be retrieved from the unlabeled data. It is suggested that if p(x) and p (y|x) do not have similar parameters, semi-supervised learning is unable to assist emphasizes this fact [13].

A method of binary classification that locates the optimal linear decision surface between two classifications is known as the Support Vector Machine. The decision surface is a weighted mixture of the supported vectors. The SVM in these utilizations need to be imputed with an individual's images, which will contain one class and the other class will contain images of other individuals besides the first individual. The SVM will then create a linear decision surface [14, 15, 16].

HYBRID

The hybrid models are defined as the models for credit scoring that have been developed by integrating two or more existing models. The benefit of these models is that the creditor can gain from having two or more models aside from reducing the weakness of the model by combining them with other models. However, these techniques are difficult to plan and execute in comparison to other methods that are easier [14] claim that the hybrid method faces faster compared to the traditional concept of neural networks. Several successful credit scoring prototypes of hybrid techniques have also been developed in current years [13]. Examine the hybrid model empirically by implementing two real groups of domain information.

COMPARISON TECHNIQUE BETWEEN SUPERVISED, UNSUPERVISED AND SEMI SUPERVISED

Unsupervised classification techniques use algorithms to classify and analyze huge amounts of raster cells. These procedures need set values for several of the operating parameters; however, the classifying method goes on without any intervention from the users. The efficacy of the unsupervised techniques is dependent on the basis that the input raster dataset has natural statistical groups of spectral patterns that consist of specific forms of physical characteristics. The entire unsupervised classification techniques, aside from the Simple One-Pass Clustering, utilize the interactive procedure to examine a set of sample input cells and decide on the set of class centers and other related statistical features. All the input raster set is then analyzed, and a classification rule is utilized to designate each raster cell to a defined class [14]. The supervised methods of classification are carried out according to the user-defined classes and subsequent representative sample sets. The training raster data sets specify the sample sets, which must be developed before imputing the Automatic Classification procedure. The activation of the Training Data button is carried out when

a supervised classification technique is selected, which shows that selection of training is needed and to set the raster. The Feature Mapping procedure offers the tools required to develop a training raster, as shown in the segment known as designing the Training Set Raster. The training sectors are first examined to decide on the statistical characteristics of each classification. In the last stage of classification, every cell in the input raster set is designated to each of the training classes by utilizing a suitable decision rule [17].

Supervised techniques result in superior outcomes when the classification idea is translated into specific groupings that are represented well by training sections and suitable for input raster's. In the Table-1 shows advantage and disadvantage of techniques in pattern recognition.

Table-1. Advantage and disadvantage of Pattern Recognition Techniques.

	Advantage	**Disadvantage**	**Reference**
PCA	▪ It is used to reduce the dimension of the data. ▪ It gives high accuracy and low computational cost. ▪ PCA gave better results for varying poses.	▪ It is very time consuming. ▪ High order dependencies still exist in PCA analysis.	[18, 19]
BTC	▪ The algorithm is independent of the size of a face image. ▪ Simple image coding technique.	▪ Larger size of the feature vector at BTC level 4 compare with other levels.	[18]
DCT	▪ DCT is used to reduce image information redundancy. ▪ DCT has been implemented in a single integrated circuit because of input independency. ▪ DCT packing the most information into the fewest coefficients for most natural images, and Minimizing block like appearance.	▪ DCT based features are sensitive to changes in the illumination direction.	[18]
LDA	▪ More efficient if model correct, borrows strength from p(x).	▪ Bias if model is incorrect.	[20]
SVM	▪ Produce very accurate classifiers. ▪ Less over fitting, robust to noise. ▪ SVM is defined by a convex optimization problems (no local minima) for which there are efficient methods.	▪ SVM is a binary classifier. To do a multi-class classification, pair-wise classifications can be used (one class against all others, for all classes). ▪ Computationally expensive, thus runs slow.	[21]

CONCLUSIONS

In this paper, we have divided pattern recognition techniques in three categories in order to supervised, unsupervised and semi supervised. We have elaborated each category and finally we compare supervised and unsupervised and comparing of methods show that unsupervised method is better than supervised when we do not have good knowledge of the surface, set of training classes are not involved all significantly distinctive types of surface materials, and each training area is not representative of its intended class.

ACKNOWLEDGEMENT

The work we presented in this paper has been supported by the Universiti Teknologi Malaysia.

REFERENCES

[1] Karamizadeh Sasan, Abdullah Shahidan M, Manaf Azizah A, Zamani Mazdak, and Hooman Alireza, "An Overview of Principal Component Analysis," Journal of Signal and Information Processing, vol. 4, p. 173, 2013.

[2] Sasan Karamizadeh, Shahidan M Abdullah, and Mazdak Zamani, "An Overview of Holistic Face Recognition," IJRCCT, vol. 2, pp. 738-741, 2013.

[3] Jigar M Pandya, Devang Rathod, and Jigna J Jadav, "A Survey of Face Recognition approach," International Journal of Engineering Research and Applications (IJERA), vol. 3, pp. 632-635, 2013.

[4] Kresimir Delac and Mislav Grgic, "A survey of biometric recognition methods," In: Electronics in Marine, 2004. Proceedings Elmar 2004. 46th International Symposium, 2004, pp. 184-193.

[5] N Jeffery R Price and Timothy F Gee, "Face recognition using direct, weighted linear discriminant analysis and modular subspaces," Pattern Recognition, vol. 38, pp. 209-219, 2005.

[6] Zhu, X. Semi-supervised learning. In: Encyclopedia of Machine Learning .pp. 892- 897, Springer.2010.

[7] Vincent Vandewalle, Christophe Biernacki, Gilles Celeux, and G©rard Govaert, "A predictive deviance criterion for selecting a generative model in semi-supervised classification," Computational Statistics and Data Analysis, 2013.

[8] Andrew Carlson, Justin Betteridge, Richard C Wang, Estevam R Hruschka Jr, and Tom M Mitchell, "Coupled semi-supervised learning for information extraction," In: Proceedings of the third ACM international conference on Web search and data mining, 2010, pp. 101-110.

[9] Saharon Rosset, Ji Zhu, Hui Zou, and Trevor J Hastie, "A method for inferring label sampling mechanisms in semi-supervised learning," In: Advances in neural information processing systems, 2004, pp. 1161-1168.

[10] Zhi-Hua Zhou and Ming Li, "Semi-Supervised Regression with Co-Training," In: IJCAI, 2005, pp. 908-916.

[11] Hariharan Narayanan, Mikhail Belkin, and Partha Niyogi, "On the relation between low density separation, spectral clustering and graph cuts," In: Advances in Neural Information Processing Systems, 2006, pp. 1025-1032.

[12] Hetal Bhavsar and Amit Ganatra, "Variations of Support Vector Machine classification Technique: A survey," International Journal, vol. 2, 2012.

[13] Matthias Seeger, "Learning with labeled and unlabeled data," technical report, University of Edinburgh 2001.

[14] Ashis Pradhan, "Support Vector Machine-A Survey," 2012.

[15] Bhumika G Bhatt and Zankhana H Shah, "Face feature extraction techniques: a survey," In: National conference onrecent trends in engineering and technology, 2011, pp. 13-14.

[16] Shahidan M. Abdullah; Azizah A. Manaf; Mazdak Zamani; "Capacity and quality improvement in reversible image watermarking approach", IEEE Networked Computing and Advanced Information Management (NCM), pp. 81-85, 2010.

[17] Feifei Long and Haifeng Xu, "Data mining technique of Acoustic Emission signals under supervised and unsupervised mode," in Natural Computation (ICNC), 2011 Seventh International Conference on, 2011, pp. 752-755.

[18] M M Hayat, "III-V Compound Avalanche Photodiodes", Comprehensive Semiconductor Science and Technology.

[19] Gore Sonal Sali Tanuja S, "A Review of Feature Extraction Techniques BTC, DCT, Walsh and PCA with FDM and BDM for Face Recognition," 2013.

[20] Shahidan M. Abdullah; Azizah A. Manaf; "Multiple Layer Reversible Images Watermarking Using Enhancement of Difference Expansion Techniques", Networked Digital Technologies, Springer Berlin Heidelberg, pg 333-342, 2010.

[21] Trevor Hastie, Robert Tibshirani, and Jerome Friedman, "Linear methods for classification," In: The Elements of Statistical Learning, ed: Springer, 2009, pp. 101-137.