International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 11 | Nov 2019 www.irjet.net p-ISSN: 2395-0072

Pattern Recognition Process, Methods and Applications in Artificial Intelligence

Susmita Karyakarte[1], Prof. Ila Savant[2]

[1]Student, *Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India*
[2]*Professor, Department of Computer Engineering, Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India*

---***---

Abstract - *Pattern recognition can be seen as a classification process.Pattern recognition has become more and more popular and important to us attracting attention as it finds it's applications into widespread areas of research. Pattern recognition plays an important role: reading texts, identifying people, retrieving objects, or finding the way in a city. Once patterns are established, however learned, we are able to classify new objects or phenomena into a class of known patterns. In Artificial Intelligence using Pattern Recognition we can explicitly program machines to learn these these patterns by themselves and use this information to perform specific tasks. It's the very first step involved in development of most AI systems that makes it an inseparable part of the process thus finding its way into various subdomains of AI. The objective of this paper is to discuss various steps involved in pattern recognition and how they are common to almost every system application of artificial intelligence and which algorithms are popularly used to implement pattern recognition in various subdomains.*

Key Words: **Pattern Recognition, Artificial Intelligence, Preprocessing, Feature Extraction, Classification, Neural Networks, Support Vector Machine, Image Processing, Data Analytics.**

1. INTRODUCTION

The field of pattern recognition is concerned with the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories. Its ultimate goal is to optimally extract patterns based on certain conditions and to separate one class from the others. The application of Pattern Recognition can be found everywhere. Examples include disease categorization, prediction of survival rates for patients of specific disease, fingerprint verification, face recognition, iris discrimination, chromosome shape discrimination, optical character recognition, texture discrimination, speech recognition,etc. The design of a pattern recognition system should consider the application domain. It is striking and interesting to observe that artificial recognition devices, especially the ones that learn from examples, are almost not, or just superficially based on a modeling of the human perception and learning abilities. One of the reasons is that the artificial systems may serve different purposes and they need to be more stable and should sometimes be faster and larger, at the cost of a reduced flexibility.

Human beings are pattern recognizers, not just because of this recognition ability, but especially because we are aware of it. We can handle it and also teach the patterns to others and discuss with them our observations. The ability to judge the similarity between objects or events is called generalization. The question of how our mind travels from observations to memory and to generalization is thereby the basic scientific question of pattern recognition and how this process can be integrated in and taught to a computer, a challenge for it's algorithms. Pattern Recognition can be implemented with the use of Machine Learning algorithms. These algorithms perform classification of data based on knowledge already gained or on statistical information extracted from patterns and/or their representation. Pattern recognition is the ability to detect arrangements of characteristics or data that yield information about a given system or data set. Predictive analytics in data science work can make use of pattern recognition algorithms to isolate statistically probable movements of time series data into the future. In a technological context, a pattern might be recurring sequences of data over time that can be used to predict trends, particular configurations of features in images that identify objects, frequent combinations of words and phrases for natural language processing (NLP), or particular clusters of behaviour on a network that could indicate an attack — among almost endless other possibilities. A basic PR system completely relies on data and derives any outcome or model from data all by itself recognizing familiar patterns quickly and accurately. The basic components of a pattern recognition process are preprocessing, feature extraction, and classification used in every PR system which are discussed further in detail.

2. Pattern Recognition Process

Pattern recognition has been under constant development for many years. It includes lots of methods impelling the development of numerous applications in different fields. The basic components in pattern recognition are preprocessing, feature extraction, and classification. Once the dataset is acquired, it is preprocessed, so that it becomes suitable for subsequent sub-processes. The next step is feature extraction, in which, the dataset is converted into a set of feature vectors which are supposed to be representative of the original data. These features are used in the classification step to separate the data points into different classes based on the problem.[1]

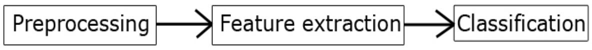

Figure 1: Basic pattern recognition process common to most models

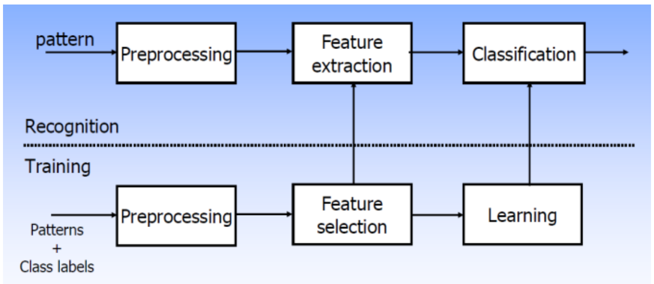

Figure 2: Pattern recognition process in a Statistical PR Model

2.1 Preprocessing

The role of preprocessing is to segment the interesting pattern from the background. It is used to reduce variations and produce a more consistent set of data. Preprocessing should include some noise filtering, smoothing and normalization to correct the image from different errors, such as strong variations in lighting direction and intensity. In some applications, segmentation of the interesting pattern of a given image from the background is very important, for example, dealing with diseases detection in agriculture applications needs segmentation of the infected region of the diseased plant images.

2.2 Feature Extraction

Feature extraction is used to overcome the problem of high dimensionality of the input set in pattern recognition. Therefore, the input data will be transformed into a reduced representation set of features, also named feature vector. Only the relevant information from the input data should be extracted in order to perform the desired task using this reduced representation instead of the full size input. Features extracted should be easily computed, robust, rotationally invariant, and insensitive to various distortions and variations in the images. Then optimal features subset that can achieve the highest accuracy results should be selected from the input space.

Two kinds of features are used in pattern recognition problems. One kind of features has clear physical meaning, such as geometric or structural and statistical features. Another kind of features has no physical meaning. We call these features mapping features. The advantage of physical features is that they need not deal with irrelevant features. The advantage of the mapping features is that they make classification easier because clear boundaries will be obtained between classes but increasing the computational complexity. Most of feature selection algorithms involve a combinatorial search through the whole space. Usually, heuristic methods, such as hill climbing, have to be adopted, because the size of input space is exponential in the number of features. Other methods divide the feature space into several subspaces which can be searched easily. There are basically two types of feature selection methods filter and wrapper. Filters methods select the best features according to some prior knowledge without thinking about the bias of further induction algorithm. So these methods performed independently of the classification algorithm or its error criteria.

In feature extraction, most methods are supervised. These approaches need some prior knowledge and labeled training samples. There are two kinds of supervised methods used: Linear feature extraction and nonlinear feature extraction. Linear feature extraction techniques include Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), projection pursuit, and Independent Component Analysis (ICA). Nonlinear feature extraction methods include kernel PCA, PCA network, nonlinear PCA, nonlinear auto-associative network, Multi-Dimensional Scaling (MDS) and Self-Organizing Map (SOM), and so forth.

2.3 Classification

During the classification task, the system uses the features extracted in the previous stage from each of the patterns to recognize them and to associate each one to its appropriate class. Two types of learning procedure are found in the literature. The classifiers that contain the knowledge of each pattern category and also the criterion or metric to discriminate among patterns classes, which belong to the supervised learning. The unsupervised learning in which the system parameters are adapted using only the information of the input, and constrained by prespecified internal rules, it attempts to find inherent patterns in the data that can then be used to determine the correct output value for new data instances.

For example, when determining whether a given image contains a face or not, the problem will be a face/non-face classification problem. Classes, or categories, are groups of patterns having feature values similar according to a given metric. Pattern recognition is generally categorized according to the type of learning used to generate the output value in this step. This step enables us to recognize an object or a pattern by using some characteristics (features) derived from the previous steps. It is the step which attempts to assign each input value of the feature vector to one of a given set of classes. There are different classification methods used in the pattern recognition process that have it's own abilities and characteristics. They are listed and described in short as follows-

- Fuzzy ART
- Neural Networks

- Markov Random Fields
- Support Vector Machine

2.3.1 Fuzzy ART

Fuzzy ART neural networks can be used as an unsupervised vector classifier. Adaptive Resonance Theory (ART) is compatible with the human brain in processing information, it has the ability to learn and memorize a large number of new concepts in a manner that does not necessarily cause existing ones to be forgotten. ART is able to classify input vectors which resemble each other according to the stored patterns. Also, it can adaptively create a new corresponding to an input pattern, if it is not similar to any existing category. ART1 was the first model of ART, it can stably learn how to categorize binary input patterns presented in an arbitrary order. Plus, Fuzzy sets theory can imitate thinking process of human being widely and deeply. [6]So the Fuzzy ART model, which incorporates computations from fuzzy set theory into the ART1 neural network, is capable of rapid stable learning of recognition categories in response to arbitrary sequences of analog or binary input patterns.

2.3.2 Neural Networks

The neural approach applies biological concepts to machines to recognize patterns. It is a promising and powerful tool for achieving high performance in pattern recognition. The outcome of this effort is the invention of artificial neural networks which is set up by the elicitation of the physiology knowledge of human brain. Neural networks are composed of a series of different, associate unit. It is about mapping device between an input set and an output set. Since the classification problem is a mapping from the feature space to some set of output classes, we can formalize the neural network, especially two-layer Neural Network as a classifier. A basic Artificial Neural Network(ANN) can consist of the following layers-the input layer, one or more hidden layers and the output layer.[7]While the usual scheme chooses one best network from amongst the set of candidate networks, better approach can be done by keeping multiple networks and running them all with an appropriate collective decision strategy. Multiple neural networks can be combined for higher recognition rate.

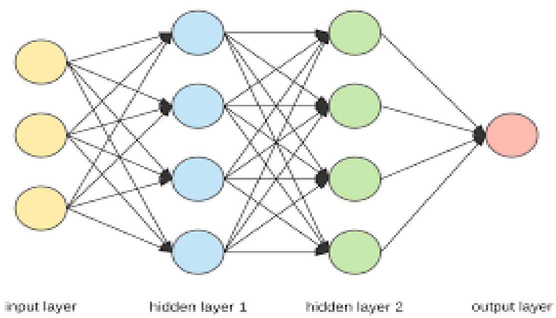

Figure 3: A simple Artificial Neural Network(ANN)

2.3.3 Markov Random Fields

Markov random fields (MRFs) are a kind of probabilistic model which encodes the model structure as an undirected graph. Two variables are connected by an edge if they directly influence each other. Markov random field (MRF) models are multi-dimensional in nature, for pattern recognition, they combine statistical and structural information. States are used to model the statistical information, and the relationships between states are used to represent the structural information. Only the best set of states should be considered. The global likelihood energy function can be rewritten with two parts, one used to model structural information that is described by the relationships among states, and the second models the statistical information because it is an output probability for the given observation and state.[9] The recognition process is to minimize the likelihood energy function that is the summation of the clique functions. It consists of an undirected graph G = (N, E) in which the nodes N represent random variables and the edges E encode conditional independence relationships via some defined rules.

2.3.4 Support Vector Machine (SVM)

The Support Vector Machine (SVM) classifier has been proved to be very successful in many applications.

The strength of the SVM is its capacity to handle not only linearly separable data, but also non-linearly separable data using kernel functions. The kernel function can map the training examples in input space into a feature space such that the mapped training examples are linearly separable. The frequently used SVM kernels are: polynomial, Gaussian radial basis function, exponential radial basis function, spline, wavelet and autocorrelation wavelet kernel. Theoretically, features with any dimension can be fed into SVM for training, but practically, features with large dimension have computation and memory that cost to the SVM training and classification process, therefore, feature extraction and selection is a crucial step before the SVM classification.[4][5]

3. Pattern Recognition Models and their Applications in Artificial Intelligence

General purpose pattern recognition is a very difficult problem hence use of object models, constraints and context is necessary for identifying complex patterns and varies depending upon the application. No single recognition approach has been found to be optimal for all pattern recognition problems. Following are the three main models of Pattern Recognition –

- Statistical Model: It identifies a specific piece of data by studying and learning from the already classified training data using statistics. This model uses supervised machine learning methods.

- Syntactic / Structural Model: Used to define a more complex relationship between elements (for example, parts of speech). This model uses semi-supervised machine learning methods.

- Template Matching Model: Matches the object's features with the predefined template and identify the object by proxy. One of the uses of such model is plagiarism checking.

Artificial Intelligence

Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules for using the information), reasoning (using rules to reach approximate or definite conclusions) and self-correction. AI is ubiquitous today, used to recommend what you should buy next online, to understand what you say to virtual assistants such as Amazon's Alexa and Apple's Siri, to recognise who and what is in a photo, to spot spam, or detect credit card fraud.AI has a number of subdomains such as Computer Vision, Deep Learning, Natural Language Processing(NLP),Data Analytics, etc. In this paper we will be discussing applications of pattern recognition in the NLP, Image Processing and Data Analytics subdomains.

3.1 PR in Natural Language Processing(NLP)

Natural language processing techniques train computers to understand what a human speaks. Natural language processing gives machines the ability to read and understand the languages that humans speak. [8]NLP research aims to answer the question of how people are able to comprehend the meaning of a spoken/written sentence and how people understand what happened, when and where that happened or what is an assumption, belief or fact. The Syntactic / Structural Model of Pattern Recognition proves to be the most useful in NLP systems to classify words and comprehend the given textual data. The following diagram represents the flow of a Syntactical model –

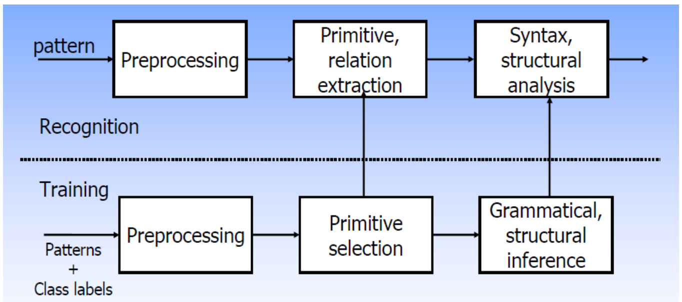

Figure 4: Syntactic / Structural Model in PR

Applications of the model in NLP are numerous and can be described as follows:-

- Text analysis - For content categorization, topic discovery and modeling (content marketing tools like Buzzsumo use this technique).

- Text summarization and contextual extraction – For finding the meaning of the text. There are many online tools for this task, for example, Text Summarizer.

- Text generation - For chatbots and AI Assistants or automated content generation (for example, auto-generated emails, Twitterbot updates, etc.);

- Text translation - In addition to text analysis and word substitution, the engine also uses a combination of context and sentiment analysis to make closer matching recreation of the message in the other language. The most prominent example is Google Translate.

- Text correction and adaptation - In addition to correcting grammar and formal mistakes, this technique can be used for the simplification of the text - from the structure to the choice of words. Grammarly, a startup that provides software for text correction is one of the most prominent examples of such NLP pattern recognition uses.

3.2 PR in Image Processing, Segmentation and Analysis

Pattern recognition is used to give human recognition intelligence to machine which is required in image processing. Template matching model is used in digital image processing for finding small parts of an image which match a template image.[2] It can be used in manufacturing control, a way to navigate a mobile robot or as a way to detect edges in images.[3] The following diagram helps to understand the concept of template matching :-

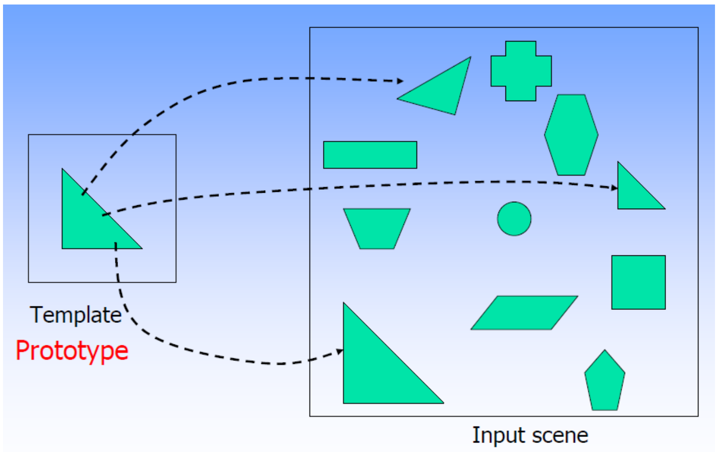

Figure 5: Template Matching

Applications are found for the detection of plant diseases to minimize the loss, and achieve intelligent farming. Plant diseases are one of the most important reasons that lead to the destruction of plants and crops. Detecting those diseases at early stages enables us to overcome and treat them appropriately. The development of an automation system using pattern recognition for classifying diseases of the infected plants is a growing research area in precision agriculture. For the identification of the visual symptoms of plant diseases, texture, color, shape and position of the infected leaf area might be used as discriminator features. Diseased regions such as spots, stains or strikes should be identified, segmented, preprocessed and a set of features should be extracted from each region.

As applications are found for identification of plant diseases, the same way pattern recognition can be used for early identification of diseases like cancer in patients with likely symptoms to help in medical diagnosis.

3.3 PR in Data Analytics

Pattern Recognition technology and Data Analytics are interconnected to the point of confusion between the two. An excellent example of this issue is stock market pattern recognition software, which is actually an analytics tool. Data analytic models are excellent examples of model following the statistical approach of pattern recognition. In the context of data analytics, pattern recognition is used to describe data, show its distinct features (i.e., the patterns itself) and put it into a broader context.

Let's look at two prominent use cases:

- Stock market forecasting - Pattern Recognition is used for comparative analysis of the stock exchanges and predictions of the possible outcomes. Yard Charts use this pattern recognition analysis.
- Audience research - Pattern Recognition refers to analyzing available user data and segmenting it by selected features. Google Analytics provides these features.

4. CONCLUSIONS

In this paper we elucidate pattern recognition in the round, including the definition, the methods of pattern recognition, the basic models of pattern recognition and it's applications in various subdomains of artificial intelligence. The applications of pattern recognition are increasing day by day thanks to the immense progress and development in machine learning algorithms used in all the PR models. It aims to classify a pattern into one of a number of classes. It appears in various fields like natural language processing (NLP), image processing, data analytics, etc. In this context, a challenge consists of finding some suitable description features since commonly, the pattern to be classified must be represented by a set of features characterizing it. These features must have discriminative properties: efficient features must be affined insensitive transformations. They must be robust against noise and against elastic deformations. The feature extracting method and the classifier should depend on the application itself. General purpose pattern recognition is very difficult and no single recognition approach has been found to be optimal for all pattern recognition problems. Hence future work will be done by searching for the right methods to develop a pattern recognition system that proves to be optimal for most pattern recognition problems.

REFERENCES

[1]Elie, Saliba. (2013). An overview of Pattern Recognition.

[2] Girolami, M.; Chao He; "Probability density estimation from optimally condensed data samples" Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 25, Issue: 10, pp: 1253 – 1264,Oct. 2003.

[3] Meijer, B.R.; "Rules and algorithms for the design of templates for template matching", Pattern Recognition, 1992. Vol.1. Conference A: Computer Vision and Applications, 11th IAPR International Conference on, pp: 760 – 763, Aug.1992.

[4] Vapnik, V., The Nature of Statistical Learning Theory, Springer, 1995.

[5] Julia Neumann, Christoph Schnorr, "SVM-based feature selection by direct objective minimization", 2004.

[6] Mun-Hwa Kim, Dong-Sik Jang, and Young-Kyu Yang. A robust-invariant pattern recognition model using fuzzy art. Pattern Recognition, 34(8):1685{1696, 2001.

[7] Vladimir N Vapnik. An overview of statistical learning theory. IEEE transactions on neural networks, 10(5):988{999, 1999.

[8] L. R. Ruiz, "Interactive Pattern Recognition applied to Natural Language Processing," Thesis, 2010.

[9] Jinhai Cai and Zhi-Qiang Liu. Pattern recognition using markov random field models. Pattern Recognition, 35(3):725{733, 2002.