Pattern Recognition – Introduction

Last Updated : 05 May, 2025

In this digital world, patterns can be found all around us. They can be seen physically in the colors of the clothing or the rhythm of the speech, or mathematically through the algorithms. In computer science, patterns are represented using vector feature values. And these patterns play an important role in understanding this world. Thus, the ability to identify these patterns becomes essential. This is where **pattern recognition** comes into play.

In this article, we will be familiarizing ourselves with the concept of pattern recognition. We will look for ways we can apply pattern recognition in our lives to solve our problems.

What is Pattern Recognition?

Pattern Recognition is the process of using **machine learning algorithms** to recognize patterns. It means sorting data into categories by analyzing the patterns present in the data. One of the main benefits of pattern recognition is that it can be used in many different areas.

In a typical pattern recognition application, the raw data is processed and converted into a form that a machine can use. Pattern recognition involves classifying and clustering patterns.

- **Classification:** Classification is when we teach a system to put things into categories. We do this by showing the system examples with known labels (like “apple” or “orange”) so it can learn and label new things. This is part of supervised learning, where we give the system the answers to learn from.
- **Clustering:** Clustering is when the system groups similar things together without any labels. It looks at the data and tries to find

natural groups. This is part of unsupervised learning, where the system learns by itself without knowing the answers beforehand.

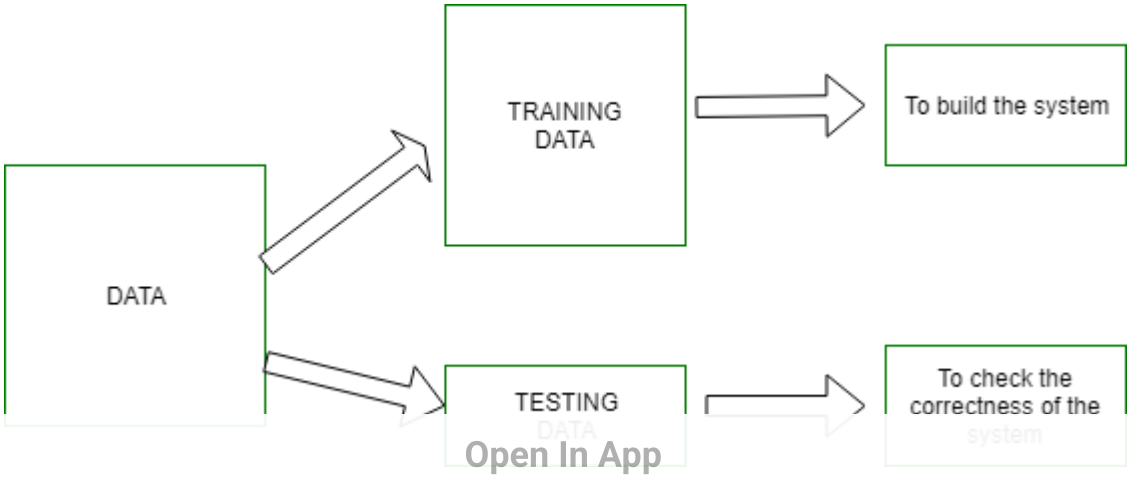
Pattern recognition possesses the following features:

- A pattern recognition system should recognize familiar patterns quickly and accurately.
- Recognize and classify unfamiliar objects.
- Accurately recognize shapes and objects from different angles.
- Identify patterns and objects even when partly hidden.

Training and Learning in Pattern Recognition

Learning is a phenomenon through which a system gets trained and becomes adaptable to give accurate results. The entire dataset is divided into two categories, one of which is used in training the model, i.e., Training set, and the other that is used in testing the model after training, i.e., Testing set.

- **Training set:** The training set is used to build a model. It consists of a set of images that are used to train the system. Training rules and algorithms are used to give relevant information on how to associate input data with output decisions. Generally, 80% of the data in the dataset is taken for training data.
- **Testing set:** Testing data is used to test the system. It is the set of data that is used to verify whether the system is producing the correct output after being trained or not. Generally, 20% of the data in the dataset is used for testing.



Example of Pattern Recognition

While talking about various types of balls, a description of a ball is a pattern. In the case balls are considered as patterns, the classes could be football, cricket ball, table tennis ball, etc. Given a new pattern, the class of the pattern would be determined. The choice of features and representation of patterns is a very important step in pattern classification.

An obvious representation of a pattern will be a **vector**. Each element of the vector can represent one feature of the pattern. The first element of the vector will contain the value of the first feature for the pattern being considered.

While representing spherical objects, (25, 1) may be represented as a spherical object with 25 units of weight and 1 unit of diameter. The class label can form a part of the vector. If spherical objects belong to class 1, the vector would be (25, 1, 1), where the first element represents the weight of the object, the second element, the diameter of the object, and the third element represents the class of the object.

Advantages and Disadvantages of Pattern Recognition

Advantages	Disadvantages
Solves classification problems	Syntactic pattern recognition is complex and slow
Detects fake biometrics	Requires a larger dataset for better accuracy
Helps visually	Cannot explain why a particular

Advantages	Disadvantages
impaired recognize cloth patterns	object is recognized
Assists in speaker diarization	May struggle with noisy or incomplete data
Recognizes objects from different angles	Can be computationally expensive
Useful in image and speech recognition	Performance depends on the quality of the training data
Enhances decision-making through clustering	Requires careful feature selection and preprocessing
Enables automation in various fields (e.g. medicine)	Difficult to adapt to completely new or unseen patterns

Applications of Pattern Recognition

Application	Description
Image Processing	Pattern recognition is used to give machines the ability to recognize human features in image processing.
Computer Vision	Used to extract meaningful features from image/video samples for applications like biological and biomedical imaging.
Seismic Analysis	Helps in discovering, imaging, and interpreting temporal patterns in seismic recordings.
Radar Signal Classification	Used in radar signal classification for applications like AP mine detection and identification.
Speech Recognition	Pattern recognition is used in speech recognition algorithms to handle larger units instead of phonemes.
Fingerprint Identification	Widely used in biometric systems for fingerprint matching and identification.
Medical Diagnosis	Pattern recognition is used to analyze medical images and assist in diagnosing diseases like cancer.
Autonomous Vehicles	Applied in self-driving cars to recognize obstacles, road signs, and pedestrians for navigation.

Python Implementation for Pattern Recognition

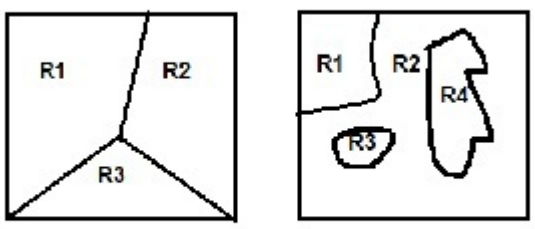
Imagine we have a dataset containing information about apples and oranges. The features of each fruit are its color (red or yellow) and its shape (round or oval). We can represent each fruit using a list of strings, e.g,. [‘red’, ’round’] for a red, round fruit.

Pattern Recognition | Basics and Design Principles

Last Updated : 10 Jan, 2023

Prerequisite – Pattern Recognition | Introduction **Pattern Recognition System** Pattern is everything around in this digital world. A pattern can either be seen physically or it can be observed mathematically by applying algorithms. In **Pattern Recognition**, pattern is comprises of the following two fundamental things:

- Collection of observations
- The concept behind the observation
    - Differentiate between good and bad features.
    - Feature properties.
        1. In a statistical-classification problem, a **decision boundary** is a hypersurface that partitions the underlying vector space into two sets. A decision boundary is the region of a problem space in which the output label of a classifier is ambiguous.**Classifier** is a hypothesis or discrete-valued function that is used to assign (categorical) class labels to particular data points.
        2. **Classifier** is used to partition the feature space into class-labeled decision regions. While **Decision Boundaries** are the borders between decision regions.



Classifier and decision boundaries

- **A Sensor :** A sensor is a device used to measure a property, such as pressure, position, temperature, or acceleration, and respond with feedback.
- **A Preprocessing Mechanism :** Segmentation is used and it is the process of partitioning a data into multiple segments. It can also be defined as the technique of dividing or partitioning an data into parts called segments.
- **A Feature Extraction Mechanism :** feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. It can be manual or automated.
- **A Description Algorithm :** Pattern recognition algorithms generally aim to provide a reasonable answer for all possible inputs and to perform “most likely” matching of the inputs, taking into account their statistical variation
- **A Training Set :** Training data is a certain percentage of an overall dataset along with testing set. As a rule, the better the training data, the better the algorithm or classifier performs.

3. Statistical Approach and
4. Structural Approach
5. **Descriptive Statistics:** It summarizes data from a sample using indexes such as the mean or standard deviation.
6. **Inferential Statistics:** It draw conclusions from data that are subject to random variation.
7. Sentence Pattern

8. Phrase Patterns
9. Formulas
10. Idioms

Pattern recognition is a subfield of machine learning that focuses on the automatic discovery of patterns and regularities in data. It involves developing algorithms and models that can identify patterns in data and make predictions or decisions based on those patterns.

There are several basic principles and design considerations that are important in pattern recognition:

1. Feature representation: The way in which the data is represented or encoded is critical for the success of a pattern recognition system. It is important to choose features that are relevant to the problem at hand and that capture the underlying structure of the data.
2. Similarity measure: A similarity measure is used to compare the similarity between two data points. Different similarity measures may be appropriate for different types of data and for different problems.
3. Model selection: There are many different types of models that can be used for pattern recognition, including linear models, nonlinear models, and probabilistic models. It is important to choose a model that is appropriate for the data and the problem at hand.
4. Evaluation: It is important to evaluate the performance of a pattern recognition system using appropriate metrics and datasets. This allows us to compare the performance of different algorithms and models and to choose the best one for the problem at hand.
5. Preprocessing: Preprocessing is the process of preparing the data for analysis. This may involve cleaning the data, scaling the data, or transforming the data in some way to make it more suitable for analysis.
6. Feature selection: Feature selection is the process of selecting a subset of the most relevant features from the data. This can help to