1.2 Pattern Recognition

Pattern is everything around in this digital world. A pattern can either be seen physically or it can be observed mathematically by applying algorithms. **Example:** The colors on the clothes, speech pattern, etc. In computer science, a pattern is represented using vector feature values.

Pattern recognition is the process of recognizing patterns by using a machine learning algorithm. Pattern recognition can be defined as the classification of data based on knowledge already gained or on statistical information extracted from patterns and/or their representation. One of the important aspects of pattern recognition is its application potential.

Overall, there are two major parts of pattern recognition algorithms:

- explorative - used to recognize commonalities in the data;
- descriptive - used to categorize the commonalities in a certain manner;

The combination of these two elements is used to extract insights out of the data, including the use in big data analytics. The analysis of the common factors and their correlation uncovers details in the subject matter that may be critical in understanding it.

Examples: Speech recognition, speaker identification, multimedia document recognition (MDR), automatic medical diagnosis.

In a typical pattern recognition application, the raw data is processed and converted into a form that is amenable for a machine to use. Pattern recognition involves the classification and cluster of patterns.

1. In classification, an appropriate class label is assigned to a pattern based on an abstraction that is generated using a set of training patterns or domain knowledge. Classification is used in supervised learning.
2. Clustering génerated a partition of the data which helps decision making, the specific decision-making activity of interest to us. Clustering is used in unsupervised learning.

Features may be represented as continuous, discrete, or discrete binary variables. A feature is a function of one or more measurements computed so that it quantifies some significant characteristics of the object.

Example: consider our face then eyes, ears, nose, etc are features of the face. A set of features that are taken together, forms the **features vector**.

Example: In the above example of a face, if all the features (eyes, ears, nose, etc) are taken together then the sequence is a feature vector([eyes, ears, nose]). The feature vector is the sequence of a feature represented as a d-dimensional column vector. In the case of speech, MFCC (Mel-frequency Cepstral Coefficient) is the spectral feature of the speech. The sequence of the first 13 features forms a feature vector.

1.2.1 Why Is Pattern Recognition Important?

Nowadays, pattern recognition serves as a basis for a number of technologies used in everyday life. Face recognition can be one of the most common examples of implementing pattern

recognition on a complex level, as it involves processing a large set of visual elements that make a person's face unique.

Face recognition, as well as other biometrics technologies, have already tremendously influenced the process of identity verification and will continue to influence our society.

Besides, pattern recognition is an irreplaceable analytical tool as well. Complex big data analyses, like stock market prediction, business analytics, or medical diagnostics rely on pattern-recognizing algorithms. Without seamless pattern recognition, drawing meaningful conclusions from large sets of data would be impossible.

1.2.2 Training and Learning Models in Pattern Recognition

Training and Learning is the building block model of Pattern Recognition. Learning is a phenomena through which a system gets trained and becomes adaptable to give result in an accurate manner. Learning is the most important phase as how well the system performs on the data provided to the system depends on which algorithms used on the data.

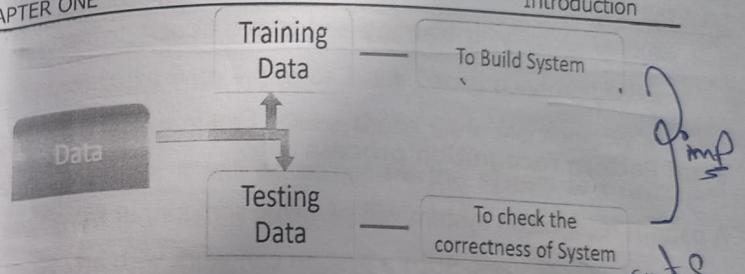
The model need to undergo from two phases and dataset is divided into two categories, one which is used in training the model and called as Training set and the other is used in testing the model after training called as Testing set.

- **Training set**

Training set is used to build a model. It consists of the set of images which are used to train the system. Training rules and algorithms used give relevant information on how to associate input data with output decision. The system is trained by applying these algorithms on the dataset, all the relevant information is extracted from the data and results are obtained. Generally, 80-85% of the data of the dataset is taken for training data.

- **Testing set**

Testing data is used to test the system. It is the set of data which is used to verify whether the system is producing the correct output after being trained or not. Generally, 20% of the data of the dataset is used for testing. Testing data is used to measure the accuracy of the system.



1.2.3 Pattern Recognition techniques

There are three basic approaches that pattern recognition algorithms utilize:

- **Statistical.** This approach is based on statistical decision theory. Pattern recognizer extracts quantitative features from the data along with the multiple samples and compares those features. However, it does not touch upon how those features are related to each other.
- **Structural (a.k.a. syntactic).** This approach is closer to how human perception works. It extracts morphological features from one data sample and checks how those are connected and related.
- **Neural.** In this approach, artificial neural networks are utilized. Compared to the ones mentioned above, it allows

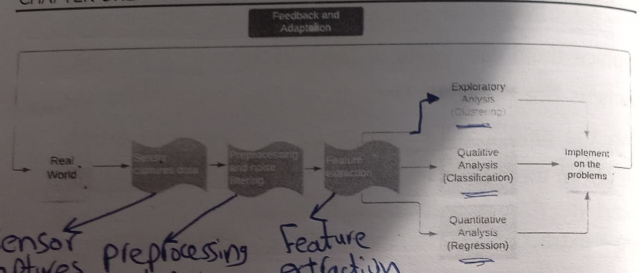


Fig (1): Step by step process

The following steps summarize the pattern recognition process

1. Collection of digital data
2. Cleaning the data from noise
3. Examining information for important features or familiar elements
4. Grouping of the elements into segments
5. Analysis of data sets for insights
6. Implementation of on the problems

The data itself can be anything:

- Text
- Images

- Sounds
- Sentiments, and others.

Any information on the sequential nature can be processed by pattern recognition algorithms, making the sequences comprehensible and enabling their practical use.

1.2.5 Features of Pattern recognition

Pattern recognition system should recognize familiar patterns quickly and accurate

1. Pattern recognition system should recognize familiar patterns quickly and accurate
2. Recognize and classify unfamiliar objects
3. Accurately recognize shapes and objects from different angles
4. Identify patterns and objects even when partly hidden
5. Recognize patterns quickly with ease, and with automaticity.

1.2.6 Pattern Recognition and Artificial Intelligence (AI)

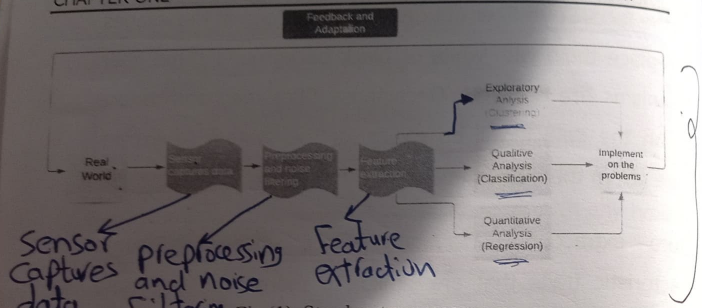


Fig (1): Step by step process

The following steps summarize the pattern recognition process

1. Collection of digital data
2. Cleaning the data from noise
3. Examining information for important features or familiar elements
4. Grouping of the elements into segments
5. Analysis of data sets for insights
6. Implementation of on the problems

The data itself can be anything:

- Text
- Images

- Sounds
- Sentiments, and others.

Any information on the sequential nature can be processed by pattern recognition algorithms, making the sequences comprehensible and enabling their practical use.

1.2.5 Features of Pattern recognition

Pattern recognition system should recognize familiar patterns quickly and accurate

1. Pattern recognition system should recognize familiar patterns quickly and accurate
2. Recognize and classify unfamiliar objects
3. Accurately recognize shapes and objects from different angles
4. Identify patterns and objects even when partly hidden
5. Recognize patterns quickly with ease, and with automaticity.

1.2.6 Pattern Recognition and Artificial Intelligence (AI)

Artificial Intelligence (AI) refers to the simulation of human intelligence, where machines are programmed to think like humans and mimic their actions. Most prominently, fields of artificial intelligence aim to enable machines to solve complex human recognition tasks, such as recognizing faces or objects. Accordingly, pattern recognition is a branch of Artificial Intelligence.

1.2.7 Pattern Recognition and Machine Learning

Today, in the era of Artificial Intelligence, pattern recognition and machine learning are commonly used to create ML models that can quickly and accurately recognize and find unique patterns in data. Pattern recognition is useful for a multitude of applications, specifically in statistical data analysis and image analysis.

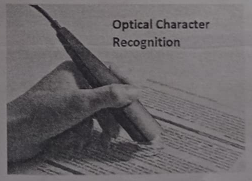
Most modern use cases of pattern recognition are based on artificial intelligence technology. Popular applications include speech recognition, text pattern recognition, facial recognition, movement recognition, recognition for video deep learning analysis, and medical image recognition in healthcare.

1.3 Pattern Recognition examples and projects

1- Stock market prediction

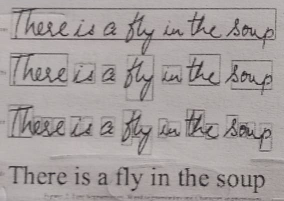
Using pattern recognition for stock market prediction applications is a classical yet challenging task with the purpose of estimating the future value of a company stock or other traded assets. Both linear and machine learning methods have been studied for decades. Only lately, deep learning models have been introduced and are rapidly gaining in popularity.

2- Optical character recognition (OCR)



Optical character recognition (OCR) is the process of classification of optical patterns contained in a digital image. The character recognition is achieved through image segmentation, feature extraction, and classification. Read our article about OCR to learn more about the recognition of texts.

3- Handwriting recognition



Handwriting recognition is used to compare patterns across handwritten text or signatures to identify patterns. Various applications are involved in the computer recognition of pen-input handwritten words. However, handwritten word recognition and spotting is a challenging field because handwritten text involves irregular and complex shapes.

4- Face recognition and visual search

Image recognition algorithms aim to detect patterns in visual imagery to recognize specific objects (Object Detection). A typical image recognition task is image classification, which uses neural networks to label an image or image segment based on what is depicted. This is the basis of visual search, where users can easily search and compare labeled images.

Advantages:

1. Pattern recognition solves classification problems
2. Pattern recognition solves the problem of fake biometric detection.
   - It is useful for cloth pattern recognition for visually impaired blind people.
3. It helps in speaker diarization.
4. We can recognize particular objects from different angles.

Disadvantages:

1. The syntactic pattern recognition approach is complex to implement and it is a very slow process.
2. Sometimes to get better accuracy, a larger dataset is required.

3. It cannot explain why a particular object is recognized.
   Example: my face vs my friend's face.

1.4 Types of Patterns

Logic, number, sound, image, and word patterns are all around us. Logic patterns help us classify similar objects, while number patterns help us predict a sequence. Word patterns help us make sense of language. Here is an example of a logic pattern.

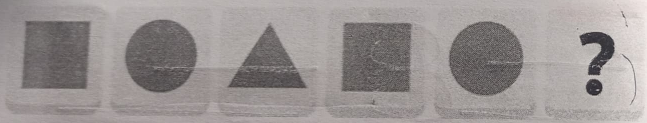


Image patterns help in classifying information in images. Melodious music patterns can be identified with music sequences.

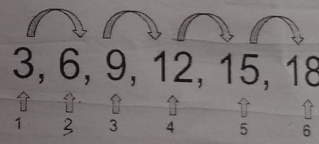
Design Patterns

Architect Christopher Alexander described first **pattern language** and **design patterns**.

A Design Pattern is a blueprint that provides a general solution to the similar type of issues you will encounter over and over in any field of activity. Here is an example.

Mathematics and Patterns

Mathematics is sometimes called the science of patterns. The most important concept in mathematics is a function. A function is an abstract representation of a pattern. Similarly, every field of activity has patterns.



Here, the function $y = f(x) = 3x$ is shown as number pattern above. The pattern shows values of the sequence, but a function can output any value of the pattern sequence directly.

There are three types of patterns that are commonly used in mathematics:

3. It cannot explain why a particular object is recognized.
   Example: my face vs my friend's face.

1.4 Types of Patterns

Logic, number, sound, image, and word patterns are all around us. Logic patterns help us classify similar objects, while number patterns help us predict a sequence. Word patterns help us make sense of language. Here is an example of a logic pattern.

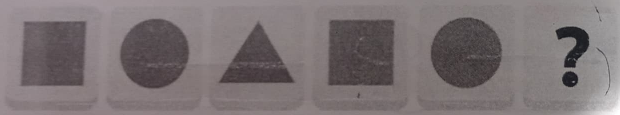


Image patterns help in classifying information in images. Melodious music patterns can be identified with music sequences.

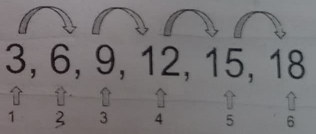
Design Patterns

Architect Christopher Alexander described first **pattern language** and **design patterns**.

A Design Pattern is a blueprint that provides a general solution to the similar type of issues you will encounter over and over in any field of activity. Here is an example.

Mathematics and Patterns

Mathematics is sometimes called the science of patterns. The most important concept in mathematics is a function. A function is an abstract representation of a pattern. Similarly, every field of activity has patterns.



Here, the function $y = f(x) = 3x$ is shown as number pattern above. The pattern shows values of the sequence, but a function can output any value of the pattern sequence directly.

There are three types of patterns that are commonly used in mathematics:

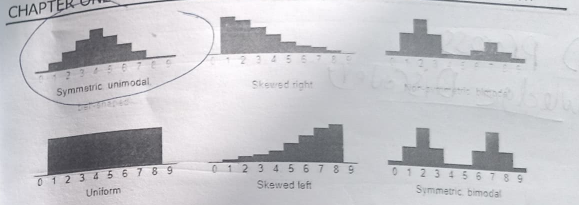
- Repeating Pattern - A pattern that keeps repeating over and over again in the sequence of numbers is called the repeating pattern.
- Growing Pattern - If the numbers or objects are arranged in an increasing order in a sequence, that pattern is called a growing pattern.
- Shirking Pattern - A shirking pattern is a pattern where numbers or objects are arranged in a decreasing order.

Example: **Find the pattern rule for the series: 81, 27, 9.**

1.4.1 Data Patterns in Statistics

Graphic displays like histograms in statistics are useful for seeing patterns in data. Patterns in data are commonly described in terms of center, spread, shape, and unusual features.

Some common distributions have special descriptive labels, such as symmetric, bell-shaped, skewed, etc. This is useful in exploratory data analysis. Probability is used to anticipate the patterns in data.



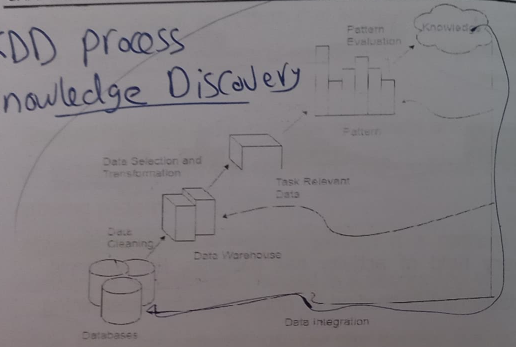
Similarly, every field of activity has patterns. For example, Cancer cell patterns are studied by oncologists for determining prophylaxis.

So, identifying patterns is an easy way to understand, organize, and classify information.

1.4.2 Patterns in Data Mining

Data nowadays is in both structured (database, spreadsheet, etc) and unstructured form (images, documents, etc). Finding relevant data is a big challenge for stakeholders.

Data Mining tools perform data analysis to uncover important data patterns contributing greatly to business strategies and scientific research.



1.4.3 Patterns in Machine Learning

Machine learning uses mathematics, statistics, and domain-specific knowledge and data to solve complex problems.

What is machine learning? Here is a very simple definition.

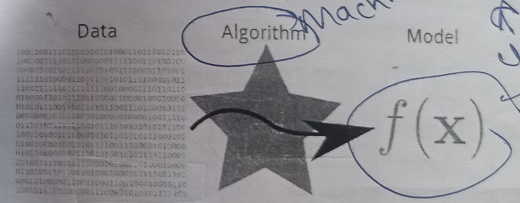
Machine learning is turning things (data) into numbers and **finding patterns** in those numbers.

For finding patterns, algorithms are used. An algorithm is a specific set of steps to perform a task.

An "algorithm" in machine learning is a procedure that is run on data to create a machine learning "model." A machine learning algorithm is written to derive the model. The model identifies the patterns in data that **fit** the dataset. Fit is a synonym to "**find patterns in data**".

A "model" in machine learning is the output of a machine learning algorithm run on data.

A model represents what was **learned** by a machine learning algorithm. It is basically a mathematical function that can adapt to new data by tweaking its parameters.

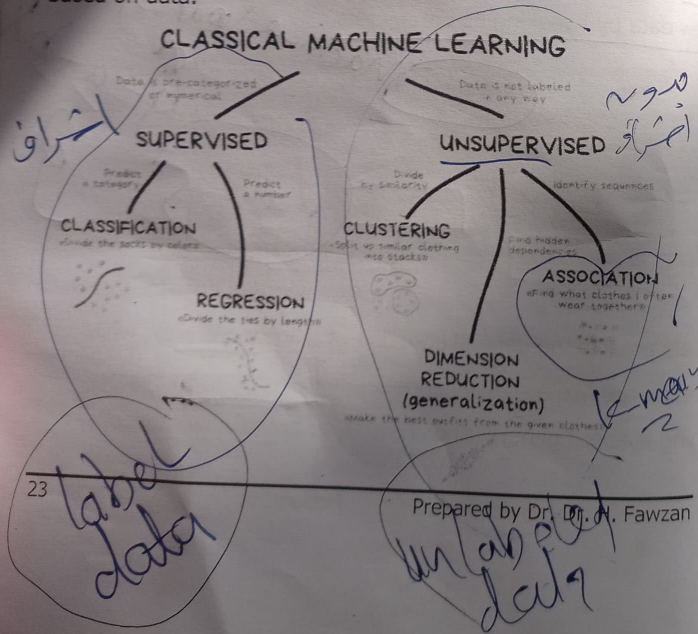


Models are like the general equation **of** a line y = a + bx, while **patterns** are like a specific equation, e.g. y = 5 + 2x. Machine learning is about generalizing correctly to brand-new situations.

The basic task of machine learning is to create a model that can predict or classify different patterns from data. One of the applications of this is the classification of spam or non-spam data.

1.5 Machine learning

Machine learning investigates how computers can learn (or improve their performance) based on data. A main research area is for computer programs to automatically learn to recognize complex patterns and make intelligent decisions based on data.



Machine learning is a form of pattern recognition which is basically the idea of training machines to recognize patterns and apply them to practical problems. Machine learning is a feature that can learn from data and iteratively keep updating itself to perform better but, Pattern recognition does not learn problems but, it can be coded to learn patterns.

For example, a typical machine learning problem is to program a computer so that it can automatically recognize handwritten postal codes on mail after learning from a set of examples.

You can see there are many similarities between data mining and machine learning. For classification and clustering tasks, machine learning research often focuses on the accuracy of the model. In addition to accuracy, data mining research places strong emphasis on the efficiency and scalability of mining methods on large data sets, as well as on ways to handle complex types of data and explore new, alternative methods.

1.5.1 Algorithms Grouped by Learning Style

There are different ways an algorithm can model a problem based on its interaction with the experience or environment or whatever we want to call the input data.

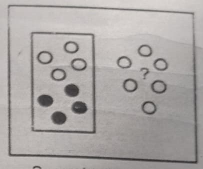
There are only a few main learning styles or learning models that an algorithm can have and we'll go through them here with a few examples of algorithms and problem types that they suit. This taxonomy or way of organizing machine learning algorithms is useful because it forces you to think about the roles of the input data and the model preparation process and select one that is the most appropriate for your problem in order to get the best result.

Let's take a look at three different learning styles in machine learning algorithms:

1. Supervised Learning

- **Supervised learning** is basically a synonym for classification. The supervision in the learning comes from the labeled examples in the training data set. For example, in the postal code recognition problem, a set of handwritten postal code images and their corresponding machine-readable translations are used as the training

examples, which supervise the learning of the classification model.



Supervised Learning Algorithms

Input data is called training data and has a known label or result such as spam/not-spam or a stock price at a time.

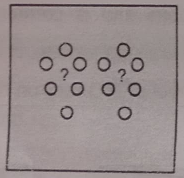
A model is prepared through a training process in which it is required to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data.

Example problems are classification and regression.

Example algorithms include: Logistic Regression and the Back Propagation Neural Network.

2. Unsupervised Learning

- **Unsupervised learning** is essentially a synonym for clustering. The learning process is unsupervised since the input examples are not class labeled. Typically, we may use clustering to discover classes within the data. For example, an unsupervised learning method can take, as input, a set of images of handwritten digits. Suppose that it finds 10 clusters of data. These clusters may correspond to the 10 distinct digits of 0 to 9, respectively. However, since the training data are not labeled, the learned model cannot tell us the semantic meaning of the clusters found.



Unsupervised Learning Algorithms

. A model is prepared by deducing structures present in the input data. Input data is not labeled and does not have a known result This may be to extract general rules. It may be through

a mathematical process to systematically reduce redundancy, or it may be to organize data by similarity.

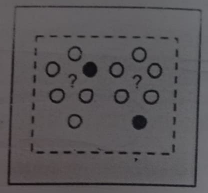
Example problems are clustering, dimensionality reduction and association rule learning.

Example algorithms include: the Priory algorithm and K-Means.

3. Semi-Supervised Learning

Semi-supervised learning is a class of machine learning techniques that make use of both labeled and unlabeled examples when learning a model. In one approach, labeled examples are used to learn class models and unlabeled examples are used to refine the boundaries between classes. For a two-class problem, we can think of the set of examples belonging to one class as the positive examples and those belonging to the other class as the negative examples. In Figure 1.12, if we do not consider the unlabeled examples, the dashed line is the decision boundary that best partitions the positive examples from the negative examples. Using the unlabeled examples, we can refine the decision boundary to the solid line.

Moreover, we can detect that the two positive examples at the top right corner, though labeled, are likely noise or outliers.



Semi-supervised Learning Algorithms

Input data is a mixture of labeled and unlabelled examples. There is a desired prediction problem but the model must learn the structures to organize the data as well as make predictions.

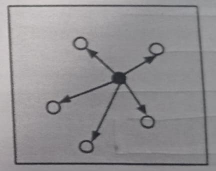
1.5.2 Algorithms Grouped By Similarity

Algorithms are often grouped by similarity in terms of their function (how they work). For example, tree-based methods, and neural network inspired methods.

This is a useful grouping method, but it is not perfect. There are still algorithms that could just as easily fit into multiple categories like Learning Vector Quantization that is both a

neural network inspired method and an instance-based method. There are also categories that have the same name that describe the problem and the class of algorithm such as Regression and Clustering.

1) Instance-based Algorithms



Instance-based Algorithms

Instance-based learning model is a decision problem with instances or examples of training data that are deemed important or required to the model.

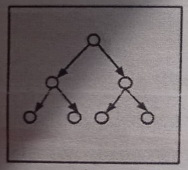
Such methods typically build up a database of example data and compare new data to the database using a similarity measure n order to find the best match and make a prediction. For this eason, instance-based methods are also called winner-take-all methods and memory-based learning. Focus is put on

the representation of the stored instances and similarity measures used between instances.

The most popular instance-based algorithms are:

- k-Nearest Neighbor (kNN)
- Learning Vector Quantization (LVQ)
- Self-Organizing Map (SOM)
- Support Vector Machines (SVM)

2) Decision Tree Algorithms



Decision Tree Algorithms

Decision tree methods construct a model of decisions made based on actual values of attributes in the data.

Decisions fork in tree structures until a prediction decision is made for a given record. Decision trees are trained on data for

classification and regression problems. Decision trees are often fast and accurate and a big favorite in machine learning.

The most popular decision tree algorithms are:

- Classification and Regression Tree (CART)
- Iterative Dichotomiser 3 (ID3)
- C4.5 and C5.0 (different versions of a powerful approach)

3) Bayesian Algorithms

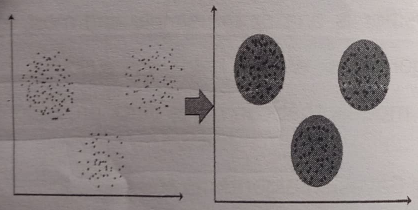
Bayesian methods are those that explicitly apply Bayes' Theorem for problems such as classification and regression. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.

The most popular Bayesian algorithms are:

- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Bayesian Network (BN)

4) Clustering Algorithms

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups. It is basically a collection of objects on the basis of similarity and dissimilarity between them. **For ex–** The data points in the graph below clustered together can be classified into one single group. We can distinguish the clusters, and we can identify that there are 3 clusters in the below picture.



The most popular clustering algorithms are:

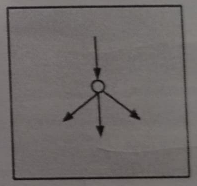
- k-Means
- k-Medians
- Expectation Maximisation (EM)
- Hierarchical Clustering

5) Association Rule Learning Algorithms

Association rule learning methods extract rules that best explain observed relationships between variables in data.

These rules can discover important and commercially useful associations in large multidimensional datasets that can be exploited by an organization.

6) Artificial Neural Network Algorithms



Artificial Neural Network Algorithms

Artificial Neural Networks are models that are inspired by the structure and/or function of biological neural networks.

They are a class of pattern matching that are commonly used for regression and classification problems but are really an

enormous subfield comprised of hundreds of algorithms and variations for all manner of problem types.

Note that I have separated out Deep Learning from neural networks because of the massive growth and popularity in the field. Here we are concerned with the more classical methods.

The most popular artificial neural network algorithms are:

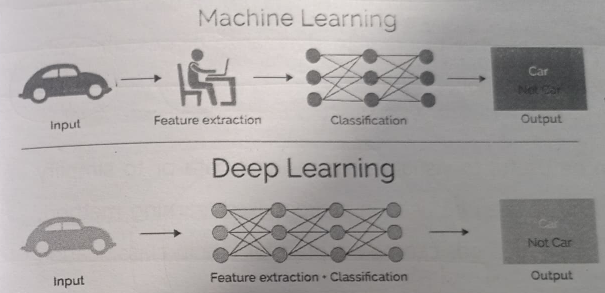
- Perceptron
- Multilayer Perceptrons (MLP)
- Back-Propagation
- Stochastic Gradient Descent
- Hopfield Network

7) Deep Learning Algorithms

Deep Learning methods are a modern update to Artificial Neural Networks that exploit abundant cheap computation.

Deep learning is a specialized form of machine learning that is inspired by the working of the brain. Deep learning uses artificial neural networks for more complex pattern tasks like audio and image processing, natural language processing, etc.

The below picture differentiates deep learning from other machine learning techniques.



They are concerned with building much larger and more complex neural networks and, as commented on above, many methods are concerned with very large datasets of labelled analog data, such as image, text. audio, and video.

The most popular deep learning algorithms are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNNs)

it is process of converting a data having vast dimensions into dataset with fewer dimensions

8) Dimensionality Reduction Algorithms

Like clustering methods, dimensionality reduction seek and exploit the inherent structure in the data, but in this case in an unsupervised manner or order to summarize or describe data using less information.

This can be useful to visualize dimensional data or to simplify data which can then be used in a supervised learning method. Many of these methods can be adapted for use in classification and regression.

- Principal Component Analysis (PCA)
- Principal Component Regression (PCR)
- Partial Least Squares Regression (PLSR)

Chapter 2

Preprocessing

2.1 Introduction

This chapter discusses several data-related issues that are important for successful pattern recognition:

The Type of Data: Data sets differ in a number of ways. For example, the attributes used to describe data objects can be of different types-quantitative or qualitative-and data