
PATTERN
RECOGNITION

CONTENTS

INTRODUCTION

- Pattern Recognition is a branch of artificial intelligence.
- Humans can recognize the faces without worrying about the various illuminations.
- When implementing such recognition artificially, it becomes a very complex task.
- The field of artificial intelligence has made this complex task possible.

PATTERN

- A pattern is a set of objects or phenomena or concepts where the elements of set are similar to one another in certain ways or aspects.
- A pattern is an entity, that could be a given name.

Examples: Fingerprint image, handwritten word, human face, speech signal, DNA sequence, etc.

EXAMPLES

TASK	INPUT	OUTPUT
Read Handwritten Characters	Characters	Read
Face Recognition	Face	Name of that person
Speaker Recognition	Voice	Name of the speaker
Medical Diagnosis	Symptoms	Disease

PATTERN RECOGNITION

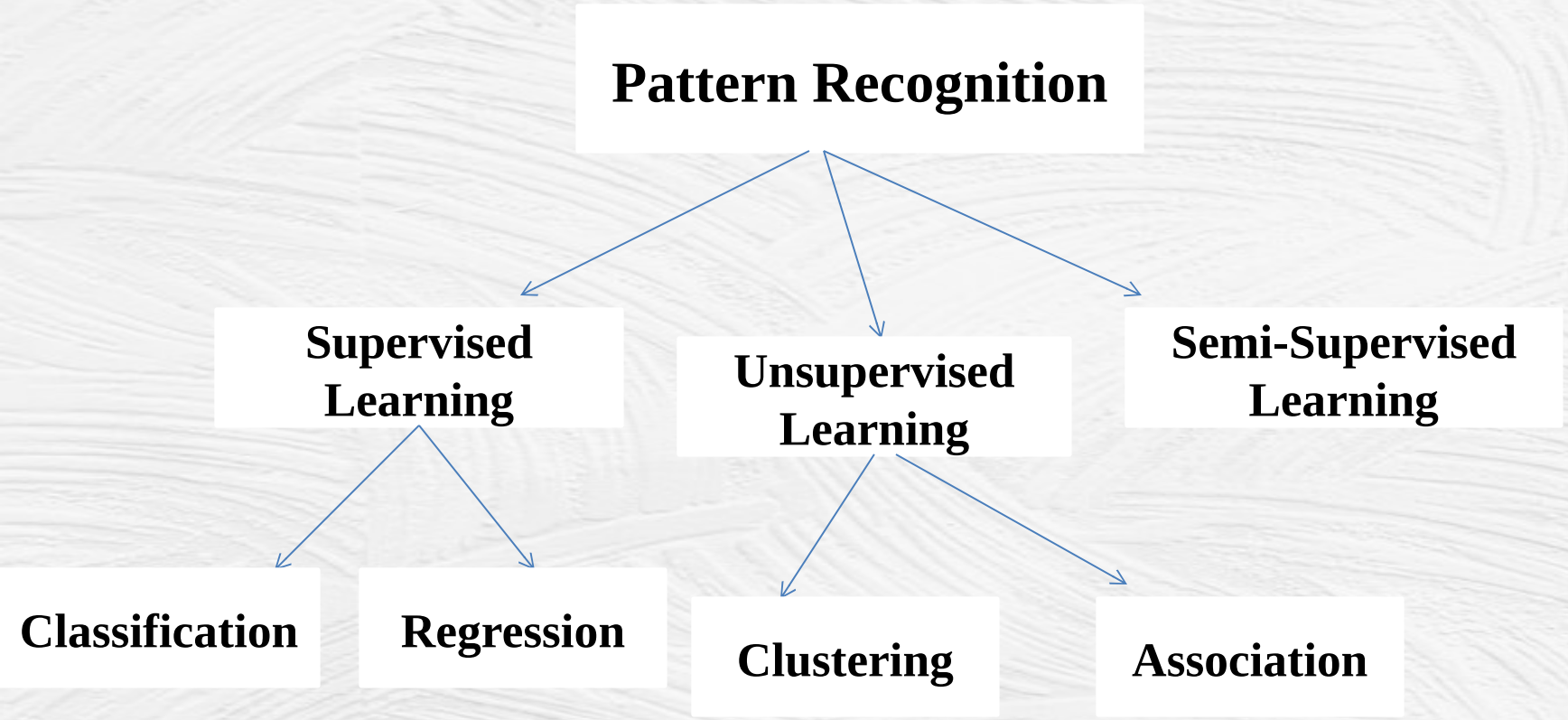
- Process of establishing a close match between some new stimulus and previously stored stimulus pattern.
- **Perceive + Process + Prediction:** It is the study of-

✓ Perceive: Observe the environment(i.e. interact with the real world)

✓ Process: Learn to distinguish patterns of interest from their background

✓ Prediction: make sound and reasonable decisions about the categories of the patterns.

Parts of Pattern Recognition



Supervised Learning

- When we have an input data(x) and output data(y) also known as labeled training data.

Types:

1. Classification
2. Regression

Supervised Learning

1. **Classification:** When the probability of output is two values only.

Example-“Yes” or “No”

“Disease” or “No Disease”

2. **Regression:** When the probability of output is more than two.

Example- “Name of the speaker”

Unsupervised Learning

- Where we have an input data(x) but no output data.

Example:- “Some ancient script which we can’t read”

➢ **Types:**

1. Clustering
2. Association

Unsupervised Learning

- **Clustering:** Here from a large data we group the data on a particular behavior.

Example- "Grouping customer on their purchase behavior"

- **Association:** is for discovering relationship between variables in large databases.

Example- "Amazon"-people who buy (x) tends to buy (y)

Semi-Supervised Learning

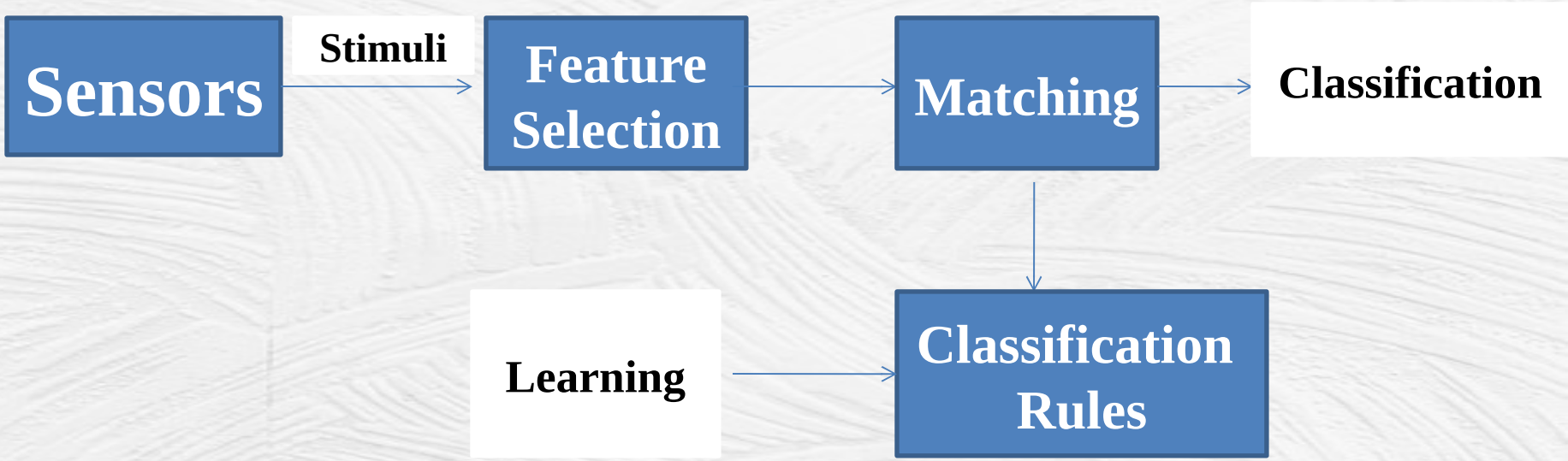
- Where we have large amount of input data but we can find output of only some of the data.
- Most of the real time applications are semi supervised

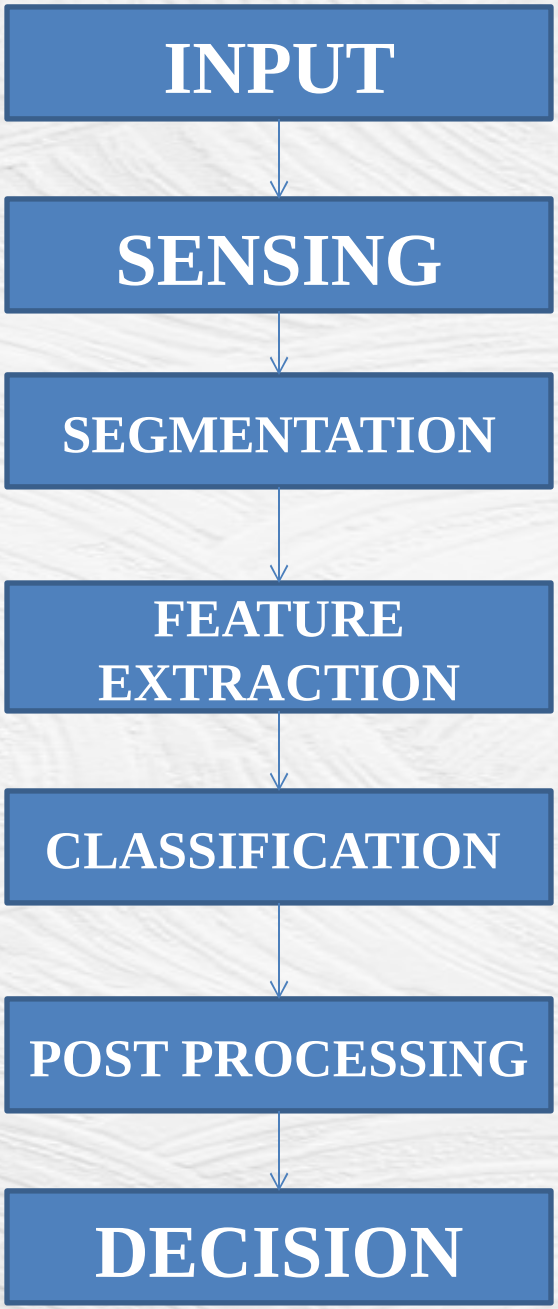
PATTERN RECOGNITION SYSTEM

- Design model of a pattern recognition system essentially involves the following 4 steps:

  - Data acquisition and preprocessing
  - Data representation
  - Feature extraction
  - Decision making

Pattern Recognition System





Pattern Recognition Process

DATA ACQUSITION BY SENSORS

- Measurement of physical variable need to be done.
- Sensor is a device which converts physical quantity to be measured into a signal which can be read, displayed, stored and used.
- Important things are bandwidth, resolution, latency, sensitivity, distortion etc.

Example: Classifying fruits in the supermarket we could use a color camera to capture the shape, color and texture features.

PREPROCESSING

- In this process noise in the data can be cleared.
- The patterns of interest can be isolated from the background.
- In text recognition, preprocessing adds many steps to make the document available for the next stage.
- The input is plain text document, the output is set of tokens for the vector modeling this stage.

Example: Recursive words are eliminated from the document, stop words are removed.

EXTRACTION OF FEATURES

- Measurements and their relations need to be extracted for the next stage of pattern recognition.
- Patterns are represented in basis of features. This process includes prior information, domain dependence, and discriminative features such as similar values for similar patterns.
- When the input is huge, it can be shorted as set of features. This process is called feature selection.

CLASSIFICATION

- Class is a group of objects having some same properties.
- It is denoted by class label.
- A classifier is a device or algorithm which takes inputs as object representation and outputs a class label.
- Classification is the process of assigning label to an object according to some representation of the objects properties

POST PROCESSING AND DECISION MAKING

- Post processing considers the cost of action. This process minimizes classification error rate and the risk.
- In this method any recursive data in the features are eliminated which in turn minimizes the risk in the process by reviewing the whole features or vectors.
- Features are evaluated based on the correctness.

Pattern Recognition Models

- There are four basic models followed in pattern recognition:

1. Statistical model
2. Syntactical or structural model
3. Template matching model
4. Neural network based model

Statistical Model

- Most intensively used model in pattern recognition systems because it is the simplest to handle.
- The statistical pattern recognition systems are based on statistics and probabilities.
- Here each pattern is described in terms of **feature sets**.
- Features are selected after analyzing the training patterns.

Syntactical Model

- Also named as structural models.
- Based on the relation between features.
- Here patterns are represented by the structures.
- Patterns used in this model forms a **hierarchical structure composed of sub-patterns.**

Template matching model

- This is a widely used model in image processing to determine the similarity between two samples, pixels or curves to localize and identify shapes in an image.
- In this model, a template or a prototype of the pattern to be recognized is available.
- Each pixel of the template is matched against the stored input image.

APPLICATIONS

- Bioinformatics
- Biometric Recognition
- Data Mining
- Document Classification
- Document Image Analysis
- Industrial Automation
- Multimedia Database Retrieval
- Remote Sensing
- Speech Recognition

Bioinformatics

- DNA sequence analysis
- DNA Mapping
- DNA micro data analysis research of heredity
- Gene Pattern Classification
- Voice Recognition
- Fingerprint Recognition
- Face Recognition
- Walking pattern analysis and classification

Medical Diagnosis

- Cancer Detection
- X-Ray mammography and image pattern recognition
- ECG signal analysis and classification

Natural Resource Study

- Geology: Earthquake analysis Rocks classification
- Forestry
- Environment
- Agriculture output analysis Soil evaluating

Biometrics

- Speech Recognition
- Face Recognition
- Character Recognition
- Natural Language Processing
- Image Processing
- Fingerprint Recognition

Problem Domain	Application	Input Pattern	Pattern Classes
Document image analysis	Optical character recognition	Document image	Characters, words
Document classification	Internet search	Text document	Semantic categories
Document classification	Junk mail filtering	Email	Junk/non-junk
Multimedia database retrieval	Internet search	Video clip	Video genres
Speech recognition	Telephone directory assistance	Speech waveform	Spoken words
Natural language processing	Information extraction	Sentences	Parts of speech
Biometric recognition	Personal identification	Face, iris, fingerprint	Authorized users for access control
Medical	Diagnosis	Microscopic image	Cancerous/healthy cell
Military	Automatic target recognition	Optical or infrared image	Target type
Industrial automation	Printed circuit board inspection	Intensity or range image	Defective/non-defective product
Industrial automation	Fruit sorting	Images taken on a conveyor belt	Grade of quality
Remote sensing	Forecasting crop yield	Multispectral image	Land use categories
Bioinformatics	Sequence analysis	DNA sequence	Known types of genes
Data mining	Searching for meaningful patterns	Points in multidimensional space	Compact and well-separated clusters