PATTERN RECOGNITION

Lecture : 1

Introduction :

what is meant by pattern?

Recognition of a signal (1D or 2D) by a machine.

How do a picture is recognised? (by human)

" The concept of abstract ideas are known to us apriori, through a Mystic connection." (Plato)

— May ~~was~~ wrongly identify similar pictures

Identifying the underlined structure in the data

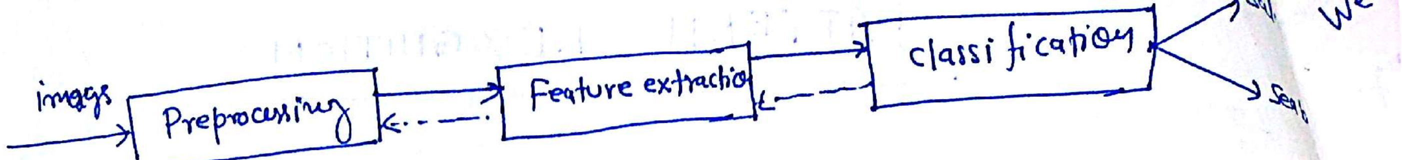
Examples: ECG pattern, speech recognition, and speaker recognition, finger print, face, gait recognition etc.

Pattern recognition - the act of taking in raw data and making an action based on the "category" of the pattern

Example : Suppose that a fish packing plant wants to automate the process of sorting incoming fish on a conveyor belt according to species

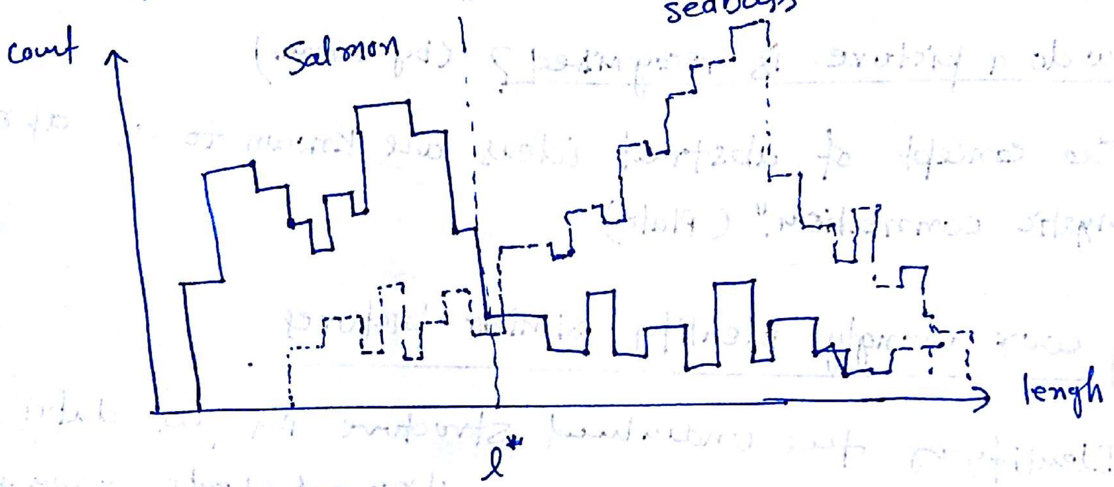
→ A camera system is set up to capture the image of two type fishes (salmon and seabass)

Many ~~Many~~ properties (features) can be used : length, lightness, width, number and shape of fins, position of the mouth.



Before feature extraction (e.g. length, lightness etc) preprocessing steps such as controlling variation in lighting (and segmentation (position of fish) is performed.

Suppose lengths of fishes is recorded as histogram (seabass is longer than salmon)



We can choose a threshold value l^* to achieve classification. Although it gives poor classification (miss classification). No matter how we choose l^*, we cannot reliably separate salmon and sea bass using length.

Cost: Till now we have assumed that all action (classification) have equal cost. Deciding the fish was a sea bass when in fact it was a salmon was just as undesirable as the converse.

In fact consumer won't mind having a "salmon" (better in taste) in packet labeled as sea bass, but they would react vigorously if piece of sea bass appear in the packet of salmon. Therefore we should lower the value of l^*.

Such considerations suggest that there is an overall single cost associated with our decision, and our true task is to make a decision rule (i.e. set a decision boundary) so as to minimize such a cost. This is the central task of decision theory

We may pick another feature, say lightness (salmon is darker)

It may be noted that lightness gives better classification, but still misclassification occurs. Another observation is that sea bass are typically wider than salmon.

We can have two features for classifying

$x_1 \rightarrow$ lightness

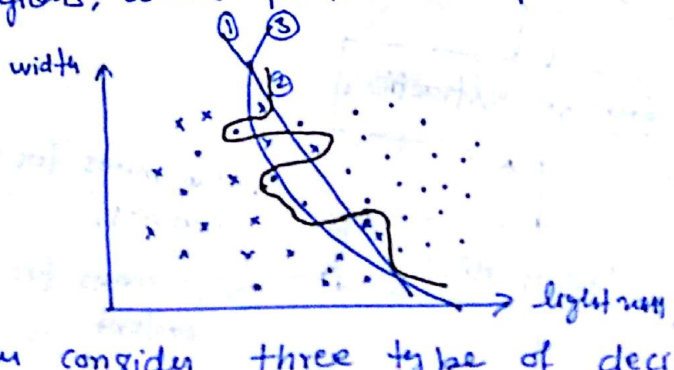
$x_2 \rightarrow$ width.

It may be realize that the features have actually reduces the image of fish into a point in feature space.

$$\vec{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

That is, each fish is represented as a vector $\vec{X}$ in a two-dimensional feature space.

Now our problem reduces to partitioning of this feature space into two regions, where points in one partition are sea bass and points in other part are salmon



We can consider three type of decision boundaries, 1 a straight line, simplest decision boundary (poor separation) 2 zigzac curve to best separate two type of object 3 smooth curve, an optimal partition.

The complex decision boundary like 2 gives poor generalization i.e. they may produce good results for particular training sample (as they are tuned to these), but ~~sin~~ unlikely to produce proper classification for new samples.

Simpler curve like 3 gives better generalization.

Lecture : 2

Dimensionality

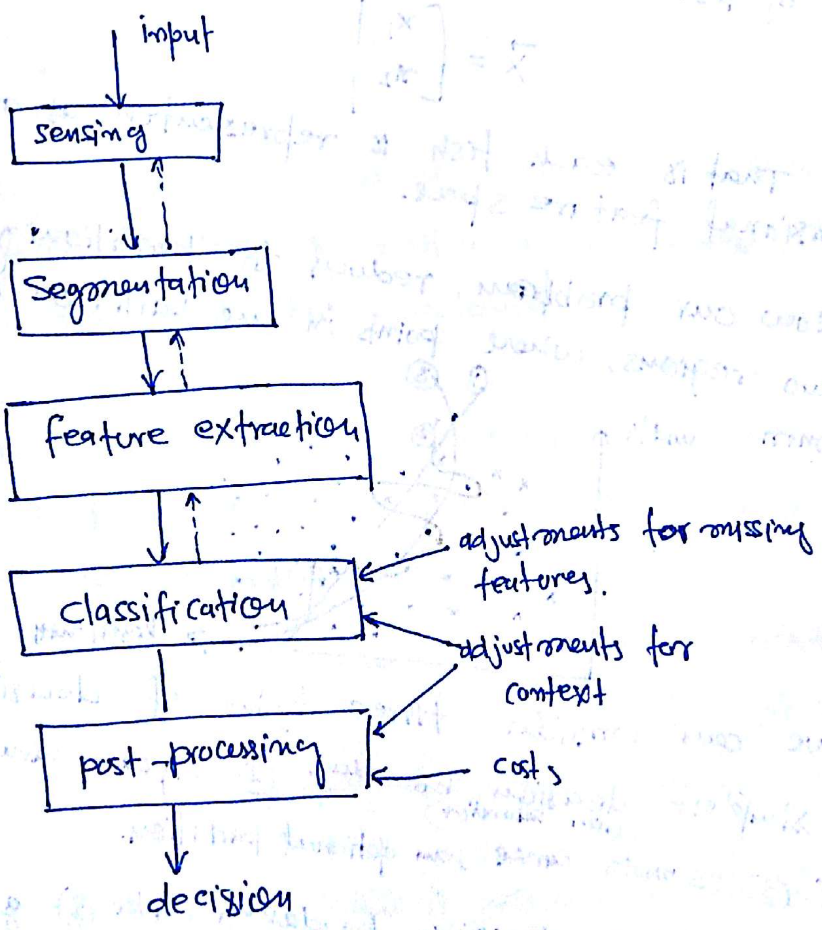
More features i.e. feature vector with higher dimension may be used to achieve better classification.

Although more features, not always produces better results. Sometime even degrades the performance of the classifier, if feature are not selected proper. As discussed earlier if we include length also, classification would be poor.

Redundancy :- Some feature may provides related information (i.e. by increasing no. of features, ~~is~~ same information or (no addition) information is achieved.). This results in redundant features.

PATTERN RECOGNITION System

Design -



Example of ~~pri~~ finger print recognition system

Feature Extraction

An ideal feature extractor would yeild a representation that makes the job of the classifier trivial;

conversely, an omnipotent classifier would not need the help of a sophisticated feature extractor.

The traditional goal of the feature extractor is to characterize an object to be recognized by measurements whose value are very similar for objects in the same category, and very different for different category objects.

It is desirable to have distinguishing features that are invariant to irrelevant transformations of the inputs

classification

The task of the classifier component is to use the feature vector provided by feature extractor to assign the object to a category

The degree of difficulty of the classification problem depends on the variability in the feature values for objects in the same category relative to the difference between feature values for the objects in different category

Noise & missing features

Noise: Any property of the sensed pattern which is not due to the true underlying model but instead to randomness in the world or the sensors.

Missing features: It may not always be possible to determine the values of all of the features for a particular input. for example in fish classification, width of fish in some cases may not be determined because of occlusion by another fish.

misclassification error & risk → minimized total cost

Design cycle