

24/11/2019

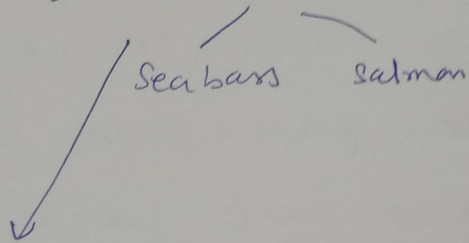
Problem: Sorting incoming fish on a conveyor belt.

Assumption: two kinds of fish

- i) sea bass
- ii) salmon.

Step to be follow:

- i) Preprocessing — Enhancement, boundary of each fish (edge detection)
- ii) Feature Extraction → 1. length as feature and draw histograms
- iii) Classification



1. draw a decision boundary that minimizes the error

2. two types

i) Generative

$X$  - input vector (pattern)

$w$  - class label (class)

- model the joint probability;  $P(X, w)$
- By using Bayes rule to calculate  $P(w/X)$  → observed
- Pick the most likely label  $w_i$

ii) Discriminative

- Estimate  $P(w|x)$  directly (learn a direct map from input  $x$  to class label  $w$ ).
- pick the most likely label ( $w$ )

iii) Generalization

iv) Ensembles of classifiers.



### Curse of dimensionality:

For quick grasp, consider this example. Say you dropped/lost a coin (gold) on a 100 meter line (road)

How do you find it? Simple, just walk on the line and search. But what if it's  $100 \times 100$  sq.m field?

It's already tough, trying to search a football ground for a single coin. But, what if it's  $100 \times 100 \times 100$  cu.m space? You know, football ground now has thirty-story building. Good luck finding a coin there! That, in essence is "Curse of dimensionality".

def:

In Applied maths, curse of dimensionality refers to the problem caused by the exponential increase in volume associated with adding extra dimensions to a mathematical space.