SMAI-S25-02: Representation, Transformations and Classification

C. V. Jawahar

IIIT Hyderabad

January 10, 2025

Agenda

1. Recap/Repeat of L01
   - https://www.dropbox.com/scl/fi/1d2ytsgk2oo930ltcrd7y/SMAI_S_2025-L1.pdf?rlkey=zec0jqkqk0lvcce58ktr62l1o&dl=0
2. This lecture:
   - Representation as a vector
   - Nearest Neighbour Algorithm
   - Linear Classification
3. Discussions and Extensions
4. Reviews and Next Step

Course Evaluation

- Exams/Quiz (40%)
- ClassWork (30%)
  - MCQ/Poll/Form
  - Problem Solving
- HomeWorks (30%)
  - Two Assignments (fixed questions)
  - One open Project (3 members in a team). **Form your team now**

Recap: Technical Summary

- We had seen how data can lead to learnable parameterized functions $f(\mathbf{W}, \mathbf{x})$
- The notion of
  - "Training" and
  - "Testing" (or inference/evaluate)

  in the context of "Supervised Learning"
- We started with an overview on how data can help in solving problems.

Representation

- Examples on how to represent in the form of $\mathbf{x} \in R^d$
  - A good day to go for outing/games/picnic
  - Representing an email for spam classification
  - Orange vs Apple on a conveyor belt
  - A "beautiful" photo and an ordinary photo (*)

Nearest Neighbour

- Nearest neighbor Algorithm for Classification
  - Classify as the label of the nearest neighbour.
- K-NN (K Nearest Neighbour) Algorithm
  - Classify based on the majority labels in the K neighbourhood.
- Questions:
  - What should be the value of K?
  - How do we know which sample is near?

Linear Classifier

- A simple Linear Binary Classifier
  - Classify as $sign(\mathbf{w}^T\mathbf{x})$
  - Either +ve or -ve.
- Watch (6 min)
  - `https://www.youtube.com/watch?v=P92mkhzt6Hg&feature=youtu.be` or
  - `https://tinyurl.com/4upj2fjs`

Feature Transformation

- In general, a feature transformation

$$\phi : \mathbf{x} \to \mathbf{x}'$$

is a useful trick.
  - **Goal:** Get better features starting from "raw" measurements/features.
- The linear feature transformation is like:

$$\mathbf{x}' = \mathbf{W}\mathbf{x}$$

can lead to dimensionality reduction, when the matrix $\mathbf{W}$ has more columns than rows.

Problem-1

Consider a two class classification problem with positive examples as

$$\{[1,1]^T,[2,2]^T,[3,2]^T\}$$

and negative examples as

$$\{[1,-1]^T,[2,-2]^T,[3,-2]^T\}$$

We are interested in classifying $[4,1]^T$. What is the prediction for $K=1,3$?
What about for $[4.0]^T$

Problem-2

We are given a set of 2D points from two classes, as shown in the Figure.
Q: To make the computations efficient, we want to do a dimensionality reduction from 2D to 1 D with the help of a 1×2 matrix $\mathbf{W}$

$$\mathbf{x}' = \mathbf{W}\mathbf{x}$$

What should be the $\mathbf{W}$ matrix be in this case?
(Indeed the goal is to get good classification performance in the new feature space $\mathbf{x}'$, while the computations could be efficient)

(a) $[1, 0]$ (b) $[-1, 0]$ (c) $[2, 0]$ (d) $[1, 1]$ (e) $[0, 1]$ (f) $[0, -1]$ (g) $[1, 0]^T$

Problem-3

We are given a set of 2D points from two classes, as shown in the Figure. We want to "rorate" the data so that points are spread across first (x) axis (i.e., something like rotate clockwise by 45°)

$$\mathbf{x}' = \mathbf{W}\mathbf{x}$$

What should be the 2×2 matrx $\mathbf{W}$ be in this case?

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
(f) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ (g) Any one of the above (h) None of the above

Problem-4

In the context of K-NN with Euclidean distance in 2 dimension,

- If vectors are in different magnitude, does it matter? Why does it happen?
- What about if we normalize such that $||x|| = 1$, does that help? How do we do that? (write an expression) Validate with an example.
- What about if each dimension (feature, attribute) is normalized to [-1,+1] range? Does it help? How do we do? (write an expression) Validate with an example.

Problem-5

As a follow up of the previous question:

- Can we achieve the same by modifying the distance function?
- Write the expression for a weighted Euclidean Distance and Show how it works.