★ Perceptron Algorithm:- (slide - 3-72)

- The percaptron algorithm is a course correction type of algorithm i.e. it updates it's classifier every time the predicted & the actual label for a data point don't match.

- For a dataset of D_n with n data points, we first define how many iterations of this algorithm we want to perform. It's a matter of choice & this number is 'T'
- We first initialize $\theta = [0\ 0\ 0 \dots 0]^T$ & $\theta_0 = 0$
- Then, we run a loop 'T' times. This is the outer loop & 'T' is arbitary.
- Inside the loop, we have maintain a boolean variable which is set to False by default. It checks for updation of classifier.
- Then inside the inner loop (which runs over n data points) we check whether the signs of predicted & actual labels are same.
- If they mismatch, their product is negative (i.e. ≤ 0). Now, we make an update to θ & θ_0 such that this product may come out positive. And we update the bool variable to True.
- In the if statement, if the bool value was False (i.e. the loop ran over all data points without updation) that means we have found the perfect classifier & no need to search further. Hence, we break the loop & return θ & θ_0.

✱ Classifier Quality :- (slide - 4 -)

- By definition a good classifier classifies every single point in the dataset D_n successfully.

★ Maths facts - Margin of point (slide - 5 -)

- The margin is basically the signed distance of a point from the classifier line / plane defined by θ, θ_0.
- It also tells whether the point is a misclassification or not.
- It sort of tells us, how far away ~~on average~~ is the data from the classifier plane / line.

★ Classifier quality 2 & Margin of dataset (Slide - 5 -)

- Margin of the dataset is the minimum of all margins of individual points in the dataset. i.e. even if there's 1 misclassification, margin of dataset will be negative.
- If all margins of points are positive (perfect classifier) then the margin of dataset will tell the minimum distance between classifier plane & data points.

★ Theorem: Perceptron performance:- (Slide - 6 -)

- Assumption B basically states that the margin of the dataset must be greater than γ which is an arbitrary value. i.e. There must be sufficient distance between data points & classifier line.
- Assumption C describes distance R such that no point in the dataset has magnitude greater than R. All datapoints line IN the circle defined by radius R at origin.