Course Name: Machine Learning and Deep learning - Fundamentals and Applications

Professor Name: Prof. M. K. Bhuyan

Department Name: Electronics and Electrical Engineering

Institute Name: Indian Institute of Technology, Guwahati

Week-12

Lecture-47

Hello everyone. Welcome to today's lecture of Machine Learning and Deep Learning Fundamentals and Application. In this lecture, this would be a continuation of the problem solving session as we did in the previous class. So here also we will be looking into few numericals related to different concepts of machine learning. So without delay, let us start the class. So the first question is there are 18 points in an axial plane such that this set of points belongs to class 1 and this set of points belongs to class 2 and also there is another set of points given by this which belongs to class 3.

Now there is a new point P which is equal to 4.2 and 1.8 and it is introduced into this plane. Now we have to find out to which class does this point P belongs to.

Now we are going to use KNN technique with K equal to 5. So how to solve this problem? So first we have to find Euclidean distance between the point P and the other points. So that is this Euclidean distance between point P and the other points. So the Euclidean distance is given by d (X, Y). So x and y are the points and it is equal to $\sqrt{(x1-y1)^2+(x2-y2)^2}$

So this is the formula. So now let us consider this one to be the X1 point, X2, X3, X4, X5, X6, X7, X8 X9, X10, X11, X12, X13, X14, X15, X16, X17 and finally X18. So now we have the points now. So now let us find the Euclidean distance between the point P and X1 first. So it is given by d(X1, P) which is equal to $\sqrt{(0.8-4.2)^2+(0.8-1.8)^2}$. So if we calculate this we will get the value 3.54. So similarly let us find the distance between the other points and that is X2 to X18 and P. So let us move on to the next page. So now let us write the distance first here X1 and P. So it is 3.54. So this we have already calculated. So similarly we can write for other points as well. So d (X2, P) it would be 3.32. So this I have already calculated.

If you use the same formula you will get the same result. d(X3, P) = 3.16, d (X4, P) = 3.45. And d (X18, P) = 1.12. So we have calculated the euclidean distance of each point from X1 to X18 and P. Now since k = 5 which is already given in the equation so we find the nearest neighbors to P. So nearest neighbors to P are so we will find those points which has the least distance. So least 5 distance would be considered. So if we look at this so the first or the lowest or the least distance is going to be this one 0.82. If you look at this distances this is the least value and second we have this point alright and third is this. Similarly the fourth point is going to be this one and finally the fifth point is going to be this one. So we can write the nearest neighbors are X17 in ascending order X8, X11, X16, and X7 and if we see the classes so the classes of these points are 3, 2, 2, 3, and 2 alright. So now when we do the voting so if we do the voting here so we get 3 points belong to class 2 and 2 points belongs to class 3. So according to the majority voting we can conclude that the point P belongs to class 2.

So this is the solution alright. So let us move on to the next problem. So this problem is consider the data points given by this set and find out the principal component for this set of points and plot it on a graph. So this is going to be our second problem. So let us write down the given data points.

Data points are given as (2, 1) so we are just writing in the vector form (3, 5), (4, 3), (5, 6), and (6, 7) right. Now we find the mean of these points so the mean is given by mean mu is equal to (4, 4.4). Now let us calculate (xi – μ) for i = 1 to 5. 5 because there are 5 points alright. So if we do this we will get this result - 2 point sorry (-2, -3.4) when we do the subtraction of mean from (2, 1) we get this result and similarly for other points we get (-1, 0.6), (0, -1.4), (1, 1.6), (2, 2.6). Now since we are done calculating (xi – μ) let us find out calculate (xi – μ) (xi – μ)^T alright. So for the first point we can write (-2, -3.4), (-2, -3.4) so this is equal to $\begin{bmatrix} 4 & 6.8 \\ 6.8 & 11.56 \end{bmatrix}$. So this part is simply (x1 – μ) (x1 – μ)^T alright. So similarly if we do this multiplication we get the following results. So for the first point we have I am just writing it again $\begin{bmatrix} 4 & 6.8 \\ 6.8 & 11.56 \end{bmatrix}$ right for the second point we have $\begin{bmatrix} 1 & -0.6 \\ -0.6 & 0.36 \end{bmatrix}$, for the third point we have $\begin{bmatrix} 0 & 0 \\ 0 & 1.96 \end{bmatrix}$, $\begin{bmatrix} 1 & 1.6 \\ 1.6 & 2.56 \end{bmatrix}$, and finally, we have $\begin{bmatrix} 4 & 5.2 \\ 5.2 & 6.76 \end{bmatrix}$ alright. Now from this we can calculate the covariance matrix.

So, the covariance matrix is given by $\frac{1}{5}\sum_{i=1}^{5}$(xi – μ) (xi – μ)T because there are 5 points. So, we have already calculated this part now we just have to add all those matrices and divide it by 5. So, the result comes out to be $\begin{bmatrix} 5 & 2.6 \\ 2.6 & 4.64 \end{bmatrix}$ alright. So now this is going to be transformation matrix for us. Now we calculate the eigenvalues. So eigenvalues are calculated using this equation that is A X = λX alright. So, the A value we have already calculated which is the transformation matrix.

So, we can write $\begin{bmatrix} 5 & 2.6 \\ 2.6 & 4.64 \end{bmatrix}$ X = λX. Now we already know determinant of (A – λI) is equal to 0 alright. So, this gives $\begin{bmatrix} 5-\lambda & 2.6 \\ 2.6 & 4.64-\lambda \end{bmatrix} = 0$ which implies $\lambda^2 - 9.64\lambda + 16.44 = 0$. So, this is the quadratic equation that we get. Now solving this equation solving say 1 solving equation 1 we get λ_1 = 7.42 and λ_2 = 2.21 alright. Now since $\lambda_1 > \lambda_2$, we will calculate the eigenvector for λ_1 alright since we are more interested in the principal component alright. Thus, we can write $\begin{bmatrix} 5-7.42 & 2.6 \\ 2.6 & 4.64-7.42 \end{bmatrix}\begin{bmatrix} x1 \\ x2 \end{bmatrix} = 0$.

So, this is for λ_1 alright. So, this gives the following set of equation. So, - 2.42 x1 + 2.6 x2 = 0 alright and 2.6 x1 - 2.78 x2 = 0 alright we have this set of linear equations. Now its augmented matrix is given by augmented matrix is given by So, this augmented matrix is given by $\begin{bmatrix} -2.42 & 2.6 & | & 0 \\ 2.6 & -2.78 & | & 0 \end{bmatrix}$ alright. So, we are using the Gaussian elimination method here to obtain the eigenvector. So, using Gaussian elimination method alright. So, we also know this by a row reduction method ok. So, now doing this operation R2 + 1.07 × R1 and replace the R2 row. So, this gives us $\begin{bmatrix} -2.42 & 2.6 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$. So, we arrive at this.

Now in linear system it is represented as -2.42 x1 + 2.6 x2 = 0 alright. Now from this we will get x1 = $\frac{2.6}{2.42}$ x2 = 1.07 x2 alright. So, therefore, the eigenvector or say the principal component is (1.07, 1) alright. So, this is the answer.

Now let us see how it look in a graph. So, we are also asked here in the question to plot a graph right. So, this can be done this way. So, the principal component can be drawn as. So, let me draw a rough graph here. So, 1, 2, 3, 4, 5, 6, 7, 8, 9, and let me draw here as well alright.

So, and the points are (2, 1) right. So, (2, 1), (3, 5) somewhere around here (4, 3) let us approximately consider this is the point (5, 6), (6, 7) alright. So, we have the points here. So, let us consider the point (1.07, 1) which is the principal vector.

So, for that we plot the point 1.07 and 1. So, there will be somewhere around this point alright. So, now, we just draw a straight line joining the point (0, 0) and (1.07, 1). So, this would be approximately this one alright. So, this is going to be the principal component for our problem alright.

So, we have obtained the result here. So, let us move on to the next problem. So, the next problem is use LDA for 2 classes C1 and C2 to cluster into 2 groups. So, C1 and C2 are given by these matrices and the new transformation point will be. So, we have to find out the new transformation point.

For this case we find the mean of this one and this one. So, it will be given by $\mu1$ = (3, 3.6) and similarly we will obtain the mean for the second class as well. So, it is given by (4.67, 2) alright. Now, we need to find the scatter matrices. So, scatter matrices given by Si are $Si = \sum_{x \in \{C1, C2\}} (x - \mu i)(x - \mu i)^T$. So, there are 2 classes that is why x belongs to C1 and C2 and (x − μi) (x − μi)T. So, now if I include something like 1/n here. So, it would become covariance matrix, but that is not required for LDA.

So, we stick to scatter matrices alright. So, let us calculate S1 here. So, S1 is calculated like this (1, 2) - (3, 3.6). So, here we are just simply doing this (x − μi) (x − μi)T. So, this is similar to this operation and we would be adding that value for other points as well.

So, that would be (2, 3) - (3, 3.6) alright. So, let us do for the other cases as well. And finally, we have (5, 5) - (3, 3.6) whole square alright. So, like this we calculate the S1.

So, when we add it up we get the result $\begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix}$ alright. And similarly for S2 we get $\begin{bmatrix} 5.33 & 1 \\ 1 & 6 \end{bmatrix}$ alright. Now we know within class scatter matrix Sw is given by S1 plus S2 alright. So, what we will do? We simply add this two this one and this one alright.

So, we get Sw = $\begin{bmatrix} 15.33 & 9 \\ 9 & 13.20 \end{bmatrix}$ alright. Now we calculate the between class scatter matrix. Between class scatter matrix given by Sb = $(\mu1 - \mu2)(\mu1 - \mu2)^T$ alright.

So, we get Sb = $[(3, 3.6) - (4.67, 2)][(3, 3.6) - (4.67, 2)]^T$ alright. So, if we do this calculation we will get this between class scatter matrix as $\begin{bmatrix} 2.79 & -2.67 \\ -2.67 & 2.56 \end{bmatrix}$ alright. So, we have these values. After this we know w which would be the transformation point is given by $Sw^{-1}(\mu1 - \mu2)$ alright. So, we need to find the inverse of this matrix inverse we have to find out inverse. So, how we do that? So, this is a 2 × 2 matrix right.

So, we will just find the determinant of this matrix like this 15.33 × 13.20 - 81 alright. So, 81 is 9 × 9 alright and we will just interchange this part alright. So, interchange this we get 3.20 15.33 and simply include a negative here and here. So, this would result in the inverse of S w alright.

So, and we have 1.67 and 1.6. So, that is $\mu1 - \mu2$ alright. So, now, after we solve this we get Sw^{-1} as $\begin{bmatrix} 0.109 & -0.074 \\ -0.074 & 0.126 \end{bmatrix}$ alright and after that I have [-1.67, 1.6]. So, if we do the calculations we arrive at this - 0.3 and minus I am sorry 0.326. So, this is the transformation point that we need alright ok.

Now, coming to the fourth question. So, it is given as in the third iteration of Adaboost the weight assigned to a misclassified data point is 0.4. If the initial weight for all data points is 1 what is the misclassification rate of this data point at the end of the second iteration. So, we write what we already know. So, we know that $D_{t+1}(n) = D_t(n)\exp(\alpha_t)$ for misclassification and this is already taught to you right where $\alpha_t = \frac{1}{2} ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$ alright.

Let us consider this as equation 1 alright. According to the equation we have this expression as 0.4 = 1 $\exp(\alpha_t)$ and from this we can obtain the α_t = ln 0.4 alright. Now, substituting this α_t substituting α_t in equation 1 we get ln 0.4 = $\frac{1}{2} ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$ alright. So, this is equal to - 1.832 = $ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$and this is equal to 0.16 = $\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$ which gives ϵ_t = 0.862 equal to which gives 86.2 percent.

In this class we saw the numerical solutions to four different machine learning concepts. So, first we saw a problem related to KNN, then we had PCA, after that we saw LDA and then we finally did a numerical problem on Adaboost. So, I hope this helps you to understand or get an idea how to solve the numerical problems related to these four machine learning concepts and I hope you will explore other numerical solution as well so that you get a better understanding of the concepts.

So, with this note I would conclude today's lecture. Thank you and have a great day. Thank you.