COMPUTER ENGINEERING DEPARTMENT

SUBJECT: MACHINE LEARNING

COURSE: T.E. **YEAR: 2020-2021** **SEMESTER: VI**
DEPT: COMPUTER ENGINEERING
SUBJECT CODE: CSDLO6021 **EXAMINATION DATE: 11/06/2021**

==========

MACHINE LEARNING
ANSWER SHEET

NAME : AMEY MAHENDRA THAKUR

SEAT NO. : 61021145

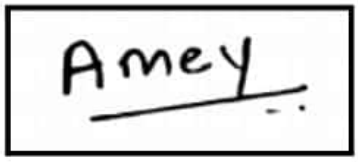
EXAM : SEMESTER VI

SUBJECT : MACHINE LEARNING

DATE : 11-06-2021

DAY : FRIDAY

STUDENT SIGNATURE:



NAME: AMEY THAKUR | BRANCH: COMPUTER | SEAT NO.: 61021145

SUBJECT: ML | EXAM: SEMESTER VI | PAGE NO.: 1 /10

Q 3 A]

i]

Logit Function

- A logit function, also known as the log-odds function, is a function that represents probability values from 0 to 1. and negative infinity to infinity.
- The function is an inverse to sigmoid function.

Importance of logit function:

- Logit function is used similar to the sigmoid in neural network
- The sigmoid is an activation function whereas logit function takes a probability and produces real number between negative and positive infinity.
- Like sigmoid function, logistic functions are often placed at the last layer.

Example:

- A logit function is used in final layer in classifier task.
- As the network determine probabilities for classification, the logit function can transform those probabilities to real number.

11-06-2021

STUDENT SIGNATURE: Amey

Logistic regression example: Spam Detection

- Spam detection is a binary classification problem where we are given an email and we need to classify whether or not it is spam.
- If the email is spam, we label it 1; if it is not spam, we label it 0.
- In order to apply logistic regression to the spam detection problem, the following features are extracted from email:

1. Sender of the email
2. Number of types in email
3. Occurance of words / phrases like "offer", "pizza", "free", etc.
4. etc.

- The resulting feature vector is then used to train a logistic classifier which emits the score in the range of 0 to 1.

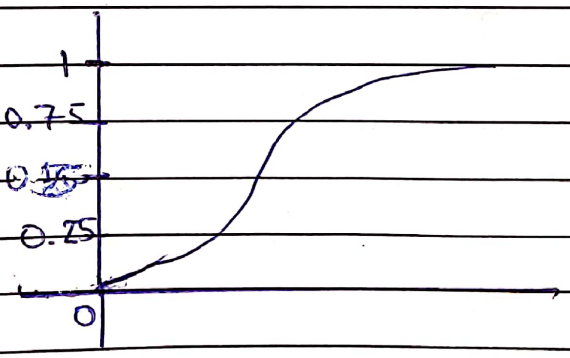
- If the score is more than 0.5 we label the email as spam. Otherwise we don't label it as spam.

11-06-2021 STUDENT SIGNATURE: Amey

Importance of logit function in logistic regression

Logistic Regression

- Logistic Regression is one of the basic and popular algorithm to solve a classification problem.
- It is the go-to method for binary classification problems (Problems with two class values).
- Linear regressions algorithm are used to predict/forecast values but logistic regression is used for classification tasks.
- The term "logistic" is taken from the logistic function that is used in this method for classification.
- The logistic function, also called the sigmoid function describe properties of population growth in ecology, rising quickly and maxing out the carrying capacity of the environment.
- It's an s-shaped curve that can take any real-valued number and map it into a value between 0 and 1, but never exactly at those limits.
- Example of logistic function.



11-06-2021 STUDENT SIGNATURE: Amey

Sigmoid Function (Logistic Function)

- Logistic regression algorithm also uses a linear equation with independent predictors to predict a value.
- The predicted value can be anywhere between negative infinity to positive infinity
- We need the output of the algorithm to be class variable. i.e. 0 - no, 1 - yes.
- Therefore, we are squashing the output of the linear equation into a range of [0,1].
- To squash the predicted value between 0 and 1, we use the sigmoid function.

Linear Equation: $z = \theta_0 + \theta_1 \cdot x_1 + \theta \cdot x_2 + \ldots$

Sigmoid Function: $y(x) = \frac{1}{1 | e^{-x}}$

Squashed Output - h: $h - y(z) = \frac{1}{1 + e^{-x}}$

11-06-2021

STUDENT SIGNATURE: Amey

Q3 A]

iii]

Covariance :

- Covariance is a measure of the relationship between two random variables. The metric evaluates how - much - to what extent - the variables change together. In other words, it is essentially a measure of the variance between two variables.

Compute covariance matrix

	X	Y
	2.5	2.4
	0.5	0.7
	2.2	2.9
	1.9	2.2
	3.1	3.0
	2.3	2.7
	2.0	1.6
	1.0	1.1
	1.5	1.6
	1.2	0.9
Mean	1.81	1.91

11-06-2021

STUDENT SIGNATURE: Amey

First find mean values.

$$X_m = \frac{\Sigma x}{N} \qquad N = \text{no. of data points} = 10$$

$$X_m = 1.81$$

$$Y_m = \frac{\Sigma y}{N}$$

$$Y_m = 1.91$$

Now find covariance matrix.

$$C = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix}$$

$$C_{xx} = \frac{\Sigma (x - x_m)^2}{N-1} = 0.6165$$

$$C_{xy} = C_{yx} = \frac{\Sigma (x - y_m)\cdot(y - y_m)}{N-1} = 0.61544$$

$$= 0.615449$$

$$C_{yy} = \frac{\Sigma (y - y_m)^2}{N-1} = 0.7165$$

$$\therefore \quad C = \begin{bmatrix} 0.6165 & 0.61544 \\ 0.61544 & 0.7165 \end{bmatrix}$$

11-06-2021 STUDENT SIGNATURE: Amey

Q 3 B]

ii]

Radial Basis Function (RBF)

- A Radial Basis Function (RBF) is a real valued function ϕ whose value depends only on the distance from the origin, so that $\phi(x) = \phi(\|x\|)$
- Alternatively on the distance from other point C, called a centre.
- So any function that satisfies the property is a radial function.
- The norm is usually Euclidean distance, although other distance functions are also possible.
- Sums of radia basis function are typically used to approximate given function.
- This approximation process can also be interpreted as a simple kind of neural network
- RBFs are also used as a kernel in support vector classification.
- Commonly used types of radial basis functions.

a. Gaussian: $\phi(r) = e^{-(\varepsilon r)^2}$

b. Multiquadratic: $\phi(r) = \sqrt{1 + (\varepsilon r)^2}$

c. Inverse Quadratic: $\phi(r) = \frac{1}{1 + (\varepsilon r)^2}$

d. Inverse Multiquadratic: $\phi(r) = \frac{1}{\sqrt{1 + (\varepsilon r)^2}}$

11 - 06 - 2021

STUDENT SIGNATURE: Amey

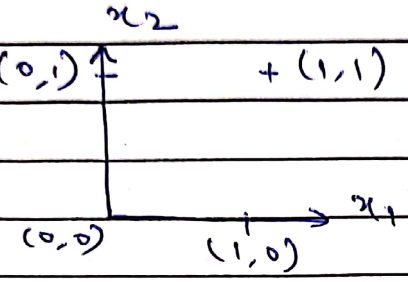
- Radial basis function network is an artificial neural network that uses radial basis functions as activation functions.
- The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters.
- Radial basis function networkes have many uses including function approximation, time series prediction, classification and system control.
- Radial basis function networks is composed of input, hidden and output layer. RBFN is strictly limited to have exactly one hidden layer. This hidden layer is called as feature vector.
- Radial Basis Function Networks increases dimension of feature vector.

11 - 06 - 2021

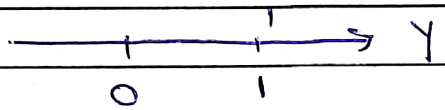
STUDENT SIGNATURE: Amey

Example: the XOR Problem

- Input space:



- Output Space :

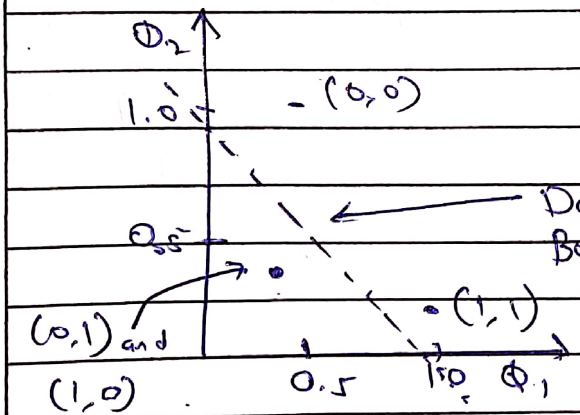


- In the feature (hidden layer) space :

$$\phi_1(\|x - t_1\|) = e^{-\|x-t_1\|^2}$$

$$\phi_2(\|x - t_2\|) = e^{-\|x-t_2\|^2}$$

with $t_1 = (1,1)$
$t_2 = (0,0)$



11 - 06 - 2021

STUDENT SIGNATURE: Amey

- When mapped into the feature space $<\Phi_1, \Phi_2>$ (hidden layer), C_1 and C_2 becomes linearly separable. So a linear classifier with $\Phi_1(x)$ and $\Phi_2(x)$ as input can be used to solve the XOR problem.

11 - 06 - 2021

STUDENT SIGNATURE: Amey