Machine Learning Report

Page 1 : Logistic Regression (binary)

The logistic regression itself simply models probability of output in terms of input and doesn't perform statistical classification . But it can make a classifier by add a small concept "Threshold Value" . This Classifier is widely used in many fields like medicine ,finance ,marketing,image detection….etc .

For Example :

(medicine) informing the doctor if the sample of the patient has a certain disease or not .

(Marketing)If the customer will buy the product or not .

(Image Detection)if there is a certain object(cat,dog...etc)in the image or not .

How does the Logistic Regression Works ?

- **Logistic Function :**
  - Using Logistic Function(Sigmoid Function) to predict the probability of the output Y between [0,1] .
  - Y(X) = B0 + B1X1 + B2X2 + B3X3 + B4X4 + ……. + BnXn
  - P(Y = 1 | X) = 1/(1+e^(Y))
  - P(Y = 1 | X) = 1/(1+e^(B0 + B1X1 + B2X2 + B3X3 + B4X4 + ……. + BnXn))

    X1,X2,…,Xn : Independent Variable (Feature) Y : Dependent Value (Target)

    B0,B1,…,Bn : Coefficients of the model

- **Cost Function :**
  - Using Cost Function to find the best fit for the model by estimate the coefficients using an optimization techniques like Gradient Descent .
  - Cost function measures the average value of the summation of the difference between the predicted values and the actual values to inform the user how well the model match the actual outcomes .
  - It is good to make cost function small as possible.

    cost = 1/2m(∑(P(Y = 1 | X) - Y(i))^2) (1-->m)

- **Optimization Techniques :**
  - Using Gradient Descent(Optimization Techniques) to find the minimum value of cost function.
  - It uses an iterative optimization algorithm .
  - It start with random values for coefficients and update them continuously till Cost Function converges .
- After Training the model ,The outputs of the model will be classified to two categories according to the Threshold Value .
  - If the output is greater than the Threshold value then the whole output will be 1, and if the output is less than the Threshold value then the whole output will be 0

  for Example Threshold = 0.5 :

  each value is above 0.5 will replace by one and each value is below 0.5 will replace by 0

Activation Function

- **Sigmoid Function :**
  - A = 1/(1 + e^(x))
  - Usually used in output layer of a binary classification, where result is either 0 or 1, as value for sigmoid function lies between 0 and 1 only so, result can be predicted easily to be ***1*** if value is greater than 0.5 and ***0*** otherwise.
- **Tanh Function :**
  - A = (e^(x) - e^(-x))/(e^(x) + e^(-x))
  - Usually used in hidden layers of a neural network as it's values lies between **-1 to 1** hence the mean for the hidden layer comes out be 0 or very close to it, hence helps in *centering the data* by bringing mean close to 0. This makes learning for the next layer much easier.
- **RELU Function :**
  - A = max(0,x)
  - ReLu is less computationally expensive than tanh and sigmoid because it involves simpler mathematical operations. At a time only a few neurons are activated making the network sparse making it efficient and easy for computation.

What is backpropgation?

•In machine learning, backpropagation is an effective algorithm used to train artificial neural networks, especially in feed-forward neural networks.

•Backpropagation is an iterative algorithm, that helps to minimize the cost function by determining which weights and biases should be adjusted. During every epoch, the model learns by adapting the weights and biases to minimize the loss by moving down toward the gradient of the error. Thus, it involves the two most popular optimization algorithms, such as **gradient descent** or **stochastic gradient descent**.

•Computing the gradient in the backpropagation algorithm helps to minimize the **cost function** and it can be implemented by using the mathematical rule called chain rule from calculus to navigate through complex lay