Performance of Artificial Neural Network on Woman's Fertility Dataset

Shrikant patro,

School of Computer Science and Engineering, VIT Chennai, Tamil Nadu, India 600127

Email:shrikantjagannath.2018@vitstudent.ac.in

Abstract- deep learning is getting a lot of press and is without any doubt the hottest topic in the machine learning field. Deep learning can be understood as a set of algorithms that were developed to train artificial neural networks with many layers most efficiently. This Paper will include a conceptual understanding of multi-layer neural networks, Training Neural Network for Image Classification, implementation of powerful backpropagation. It discuss the Artificial Neural Network that work better than the single layer perceptron neural network.

Introduction

When Multiple Single layer neuron are connected together to form multi-layer feed forward neural network. This is also called as multi-layer perceptron (MLP). In Single Layer Perceptron, it is very difficult to fit in case of non-linearly separable pattern. It is achieved by mapping non-linearly separable x-space into high dimensional z-space. It is painful task. So, deal with such problem efficiently we need more than one single perceptron connected together it is called MLP. The figure1 explains the concept of an MLP consisting of three layers: one input layer, one hidden layer, and one output layer. The units in the hidden layer are fully connected to the input layer, and the output layer is fully connected to the hidden layer, respectively. If such a network has more than one hidden layer, we also call it a *deep* artificial neural network.

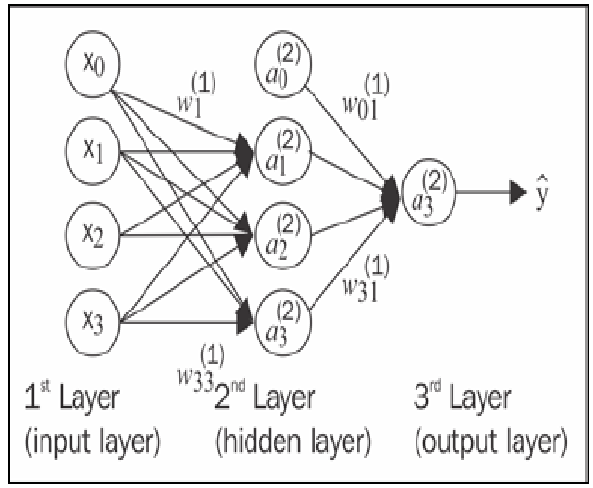

Fig.1 Multi-Layer perceptron

The *i*th activation unit in the *l*th layer as a_1^l, and the activation units a_0^1 and a_0^2 are the **bias units**, respectively, which we set equal to 1. The activation of the units in the input layer is just its input plus the bias units:

$$a^{(1)} = \begin{bmatrix} a_0^{(1)} \\ a_1^{(1)} \\ \vdots \\ a_m^{(1)} \end{bmatrix} = \begin{bmatrix} 1 \\ x_1^{(i)} \\ \vdots \\ x_m^{(i)} \end{bmatrix}$$

Each unit in layer 'L' is connected to all units in layer 'L+1' through weight coeffcient. The connection between the kth unit in layer 'L' to jth Unit to Layer 'L+1' would be denoted as $w_{j,k}^{(l)}$. While one unit in the output layer would suffice for a binary classification task, we saw a more general form of a neural network in the preceding figure, which allows us to perform multi-class classification via a generalization of the One-vs-All (OvA) technique.

Database- Woman's Fertility dataset

There are different CSV file for different phenomenon associated with woman fertility dataset. This dataset is collected by WHO with the help of 100 volunteers a semen sample analyzed according to WHO 2010 criteria. Sperm concentration are related to socio-demographic data, environmental factors, health status, and life habits. Dataset characteristics: Multivariate, Attribute characteristics: Real, Associated task: Classification, Regression, Number of instances: 100, Number of attribute: 10, Missing values: NA, Area: Life, Date Donated: 2013-01-17 and number of web hits: 142847. The further description is specifically regarding the "Blood Transfusion Service Center". This include field like Recency (Months), Frequency (Times), Monetary (c.c. blood), Time (months) and donated. Donated is binary attributed stating in terms of 1 and 0 to represent whether donated or not.

Methodology

The process of activation of the neural network through forward propagation. This section will include the process of forward propagation to calculate the output of MLP Model. The learning process of output is summarized into following steps:

1. Starting at the input layer, we forward propagate the patterns of the training data through the network to generate an output.
2. Based on the network's output, we calculate the error that we want to minimize using a cost function that we will describe later.
3. We backpropagate the error, find its derivative with respect to each weight in the network, and update the model.

Finally, after repeating the steps for multiple epochs and learning the weights of the MLP, we use forward propagation to calculate the network output and apply a threshold function to obtain the predicted class labels in the one-hot representation, which was discussed earlier. Each unit in the hidden unit is connected to all units in input layer the calculation of a_1^2 as follows:

$$z_1^{(2)} = a_0^{(1)} w_{1,0}^{(1)} + a_1^{(1)} w_{1,1}^{(1)} + \cdots + a_m^{(1)} w_{1,m}^{(1)}$$

$$a_1^{(2)} = \phi\left(z_1^{(2)}\right)$$

Here, Z is the net input and a_1^2 is the activation function, which has to be differentiable to learn the weights that connect the neurons using a gradient-based approach. To be able to solve complex problems such as image classification, we need nonlinear activation functions in our MLP model, for example, the sigmoid (logistic) activation function that we used in logistic regression. As we can remember, the sigmoid function is an S-shaped curve that maps the net input Z onto a logistic distribution in the range 0 to 1, which passes the origin at Z=0.5 is shown in Fig 2.

$$\phi(z) = \frac{1}{1+e^{-z}}$$

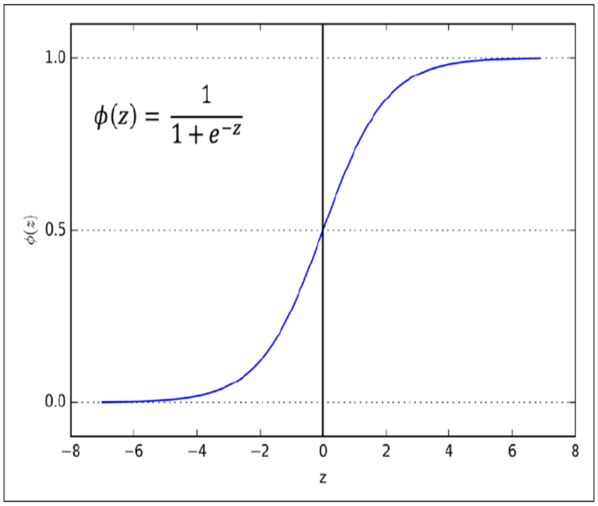

Fig 2. Sigmoid Function

The MLP is a typical example of a *feedforward* artificial neural network. The term *feedforward* refers to the fact that each layer serves as the input to the next layer without loops, in contrast to *recurrent neural networks*. The name multi-layer perceptron is confusing but the individual unit in MLP is Logistic regression unit. The output of each neuron is continuous real valued number between 0 and 1.

Conclusion

The performance of the multi-layer Perceptron is significantly better than Single Layer Perceptron for the classification of the result into more than one classes and it is basically collection of multiple neurons such that each individual is the Logistic Regression units.

References:

[1] Sebastian Raschka, “Python Machine Learning “