Sigmoid Neuron - a solution to non linear data

With reference from session by Prof. Mitesh Khapra and Pratyush Kumar

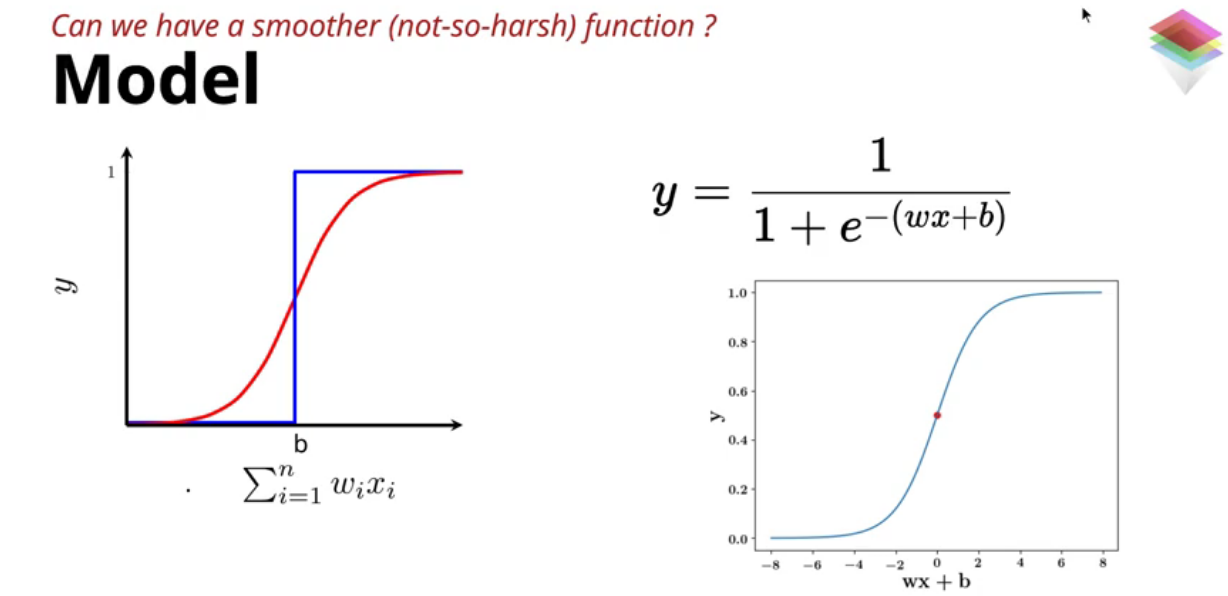
The Sigmoid Neuron model is another member of the Neuron family models in deep learning. Its is often used to work on data that is not linearly seperable. The Sigmoid Neuron model is implemented according to the [Six jars of Machine Learning](https://medium.com/@nitinnaidu/six-jars-of-machine-learning-2dd5a72ca1b" \t "/home/nitin/Documents\\x/_blank) in the earlier publish.

1. Data and task-

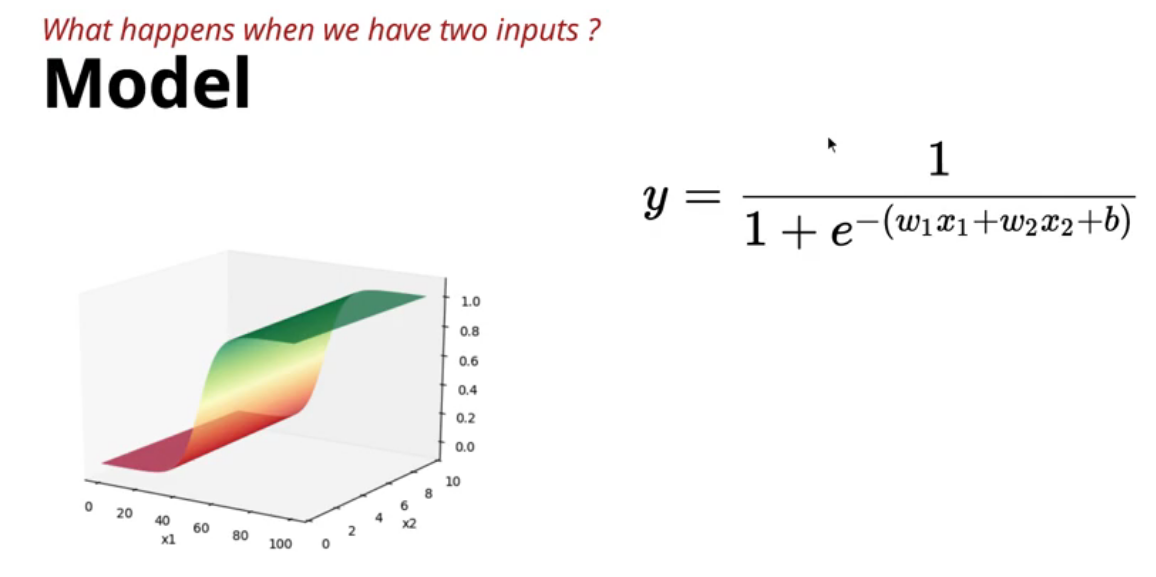
The data can be binary or real values. The input data may further be standardized or binarised as per need. The output data is continuous which may be a value between 0 and 1. The task is normally regression i.e regressing a probability for data. Suppose you are given x which contains various attributes about a location in the ocean (such as, density, pressure, salinity, temperature, depth, etc). Using this x, you can predict y which is the probability of finding y at the given location.

Model-

The perceptron model draws plots a line which is harsh and only works on linearly seperable data. Sigmoid neuron provides a solution to this by plotting a smoother curved line by working on non-linear data.



The model consists of logistic function with bias, input data and its variables. The function can take multiple input by the summation of the input in the logistic function.



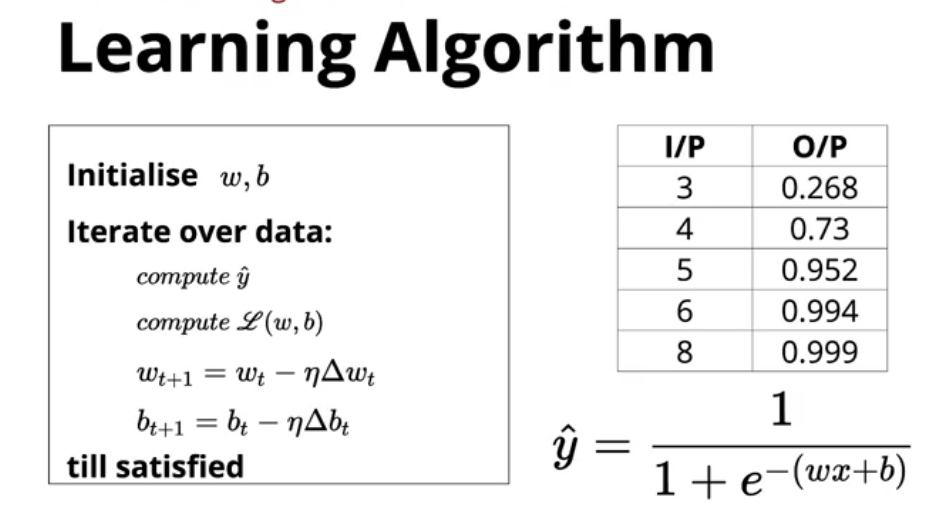
The model works well on data that is not linearly separable data. The best example is that of the probability of a person buying a car. In this case the predication as a percentage value would be more sensible rather than binary prediction. Sigmoid Neuron fits well for these sort of predictions.

Loss function-

The squared loss function is applied. It finds the difference between the predicted value and the actual prediction as a square. Error is more accurate by finding the points in a plot which contribute more to the overall error.

Learning Algorithm-

The learning algorithm consists of manipulating the parameters bias and the weight. Starting with random value for the parameters they are replaced until the error is minimized or the output is closer to the true output. A principled approach which relies on the loss function is used to find the ideal parameters. The loss steadily decreases when the approach is implemented. Tools like PyTorch, Tensorflow are used to find the ideal parameters with minimum loss.



Accuracy-

Accuracy for regression output is often taken by Root mean square error(RMSE). In case the output is binary it is given by the number of correct predictions divided by the total number of predictions.

Thus Sigmoid Neuron is a good model to work on data that is not linearly separable.

**References:**

1. Deep Learning by One Fourth Labs
2. Wikipedia