Building a Shallow Neural Network using Logistic Regression

in fulfilment of the requirements

for the completion of Neural Network Project

by

G.Naga Laxmi(N150390)

Department of Electronics and Communication Engineering

Rajiv Gandhi University of Knowledge Technologies

Nuzvid, India - 521202

June 2021

Certificate

It is certified that the work contained in this term project entitled 'Building a Shallow Neural Network using Logistic Regression' by 'G.Naga Laxmi' has been carried out under the supervision of Chakravarthi Jada and that it has not been submitted elsewhere for a degree.

Project Guide

June 2021

Chakravarthi Jada

Asst.Professor

RGUKT NUZVID

Abstract

Name of the student:

G.Naga Laxmi

Degree for which submitted: **B.Tech** Department: **ECE**

Term project title: **Building a Shallow Neural Network using Logistic Regression**

Project guide: **Chakravarthi Jada**

Month and year of thesis submission: **June 2021**

Present world will running on new technologies.With the technology human work will also be reduced rapidly.One of them is Neural Networks.This paper presents a object prediction using shallow neural network with one hidden layer with the help of logistic regression cost function.Shallow neural networks consist of only 1 or 2 hidden layers. Understanding a shallow neural network gives us an insight into what exactly is going on inside a deep neural networks.Here To detect the whether given image belongs to cat or not.steps involved in this are Initialization of parameters,finding the cost function,forward the backward optimization and prediction.

Key Words:: Neural Network ·Shallow Neural Network ·Logistic Regression ·Gradient Descent · Optimization ·updation ·prediction ·classification

Acknowledgements

I would extend my sincerest gratitude... to my project guide who guides and give best instructions to do this project and make it successful.

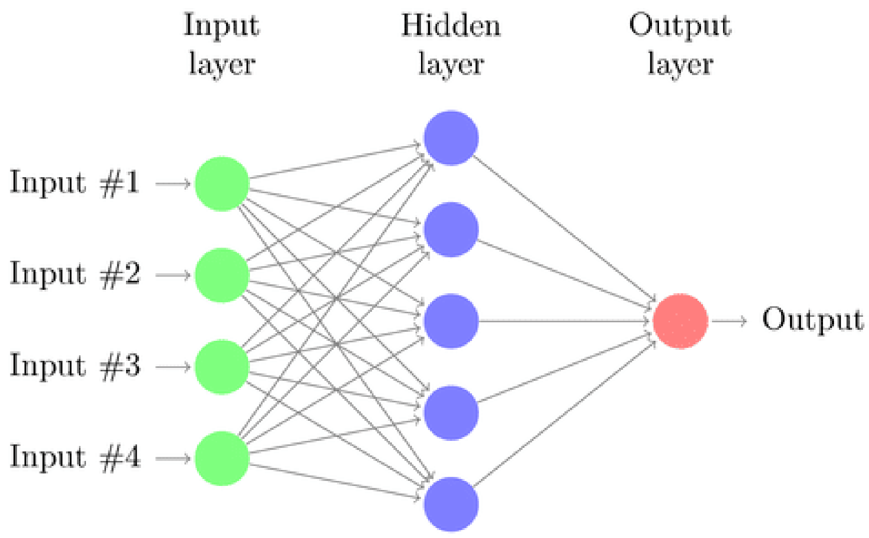
Contents

List of Figures

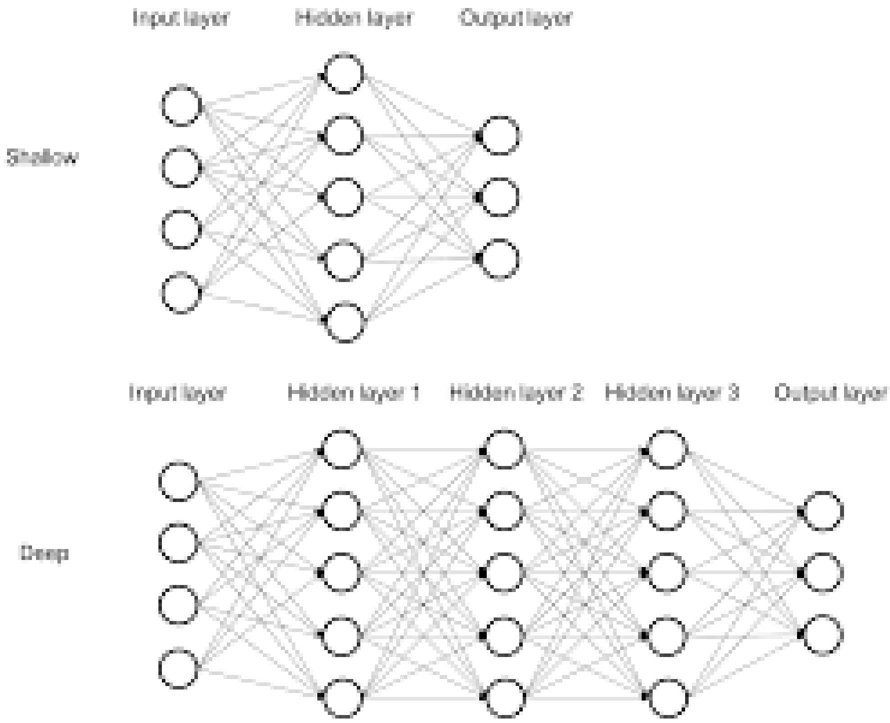
Chapter 1

Introduction

A Neural Network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading systems. "Shallow" neural networks is a term used to de-

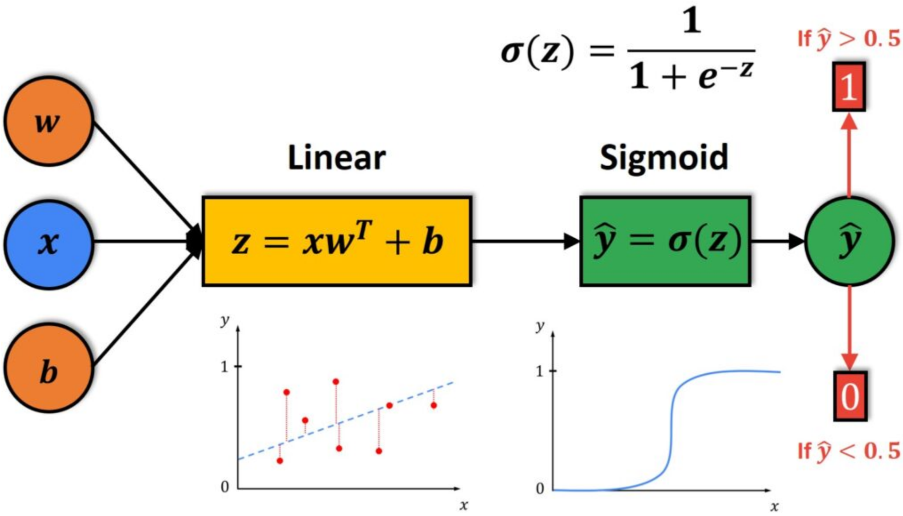


scribe NN that usually have only one hidden layer as opposed to deep NN which have several hidden layers, often of various types. ... In lay man terms : Shallow means : NOT DEEP that is no of hidden layer = 1. A shallow network has less number of hidden layers.Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.



1.1 Logistic Regression

Logistic regression is a classification algorithm used to assign observations to a discrete set of classes. Some of the examples of classification problems are Email spam or not spam, Online transactions Fraud or not Fraud, Tumor Malignant or Benign. Logistic regression transforms its output using the logistic sigmoid function to return a probability value.We can call a Logistic Regression a Linear Regression model but the Logistic Regression uses a more complex cost function, this cost function can be defined as the 'Sigmoid function' or also known as the 'logistic function' instead of a linear function. The hypothesis of logistic regression tends it to limit the cost function between 0 and 1. Therefore linear functions fail to represent it as it can have a value greater than 1 or less than 0 which is not possible as per the hypothesis of logistic regression.



$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} Cost(h_\theta(x^{(i)}), y^{(i)})$$

$$J(\theta) = \frac{1}{m}[\sum_{i=1}^{m} -y^{(i)} log(h_\theta(x^{(i)})) + (1 - y^{(i)}) log(1 - h_\theta(x^{(i)}))]$$

$$m = number\ of\ samples$$

Chapter 2

Building a shallow Neural Network

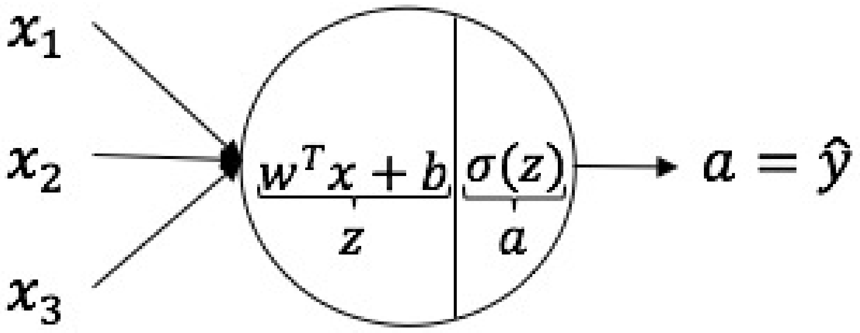
2.1 Sigmoid Function

Special cases of the sigmoid function include the Gompertz curve (used in modeling systems that saturate at large values of x) and the ogee curve (used in the spillway of some dams). Sigmoid functions have domain of all real numbers, with return (response) value commonly monotonically increasing but could be decreasing. Sigmoid functions most often show a return value (y axis) in the range 0 to 1. Another commonly used range is from 1 to 1.

A wide variety of sigmoid functions including the logistic and hyperbolic tangent functions have been used as the activation function of artificial neurons. Sigmoid curves are also common in statistics as cumulative distribution functions (which go from 0 to 1), such as the integrals of the logistic density, the normal density, and Student's t probability density functions. The logistic sigmoid function is invertible, and its inverse is the logistic function.

2.2 Optimization

Initializing parameters:W,b Forward and Backward propagation: To learning the parameters Farward Propagation:X,A,J(cost Function) Backword propagation:dw,db You have initialized your parameters. Update the parameters using gradient descent. Gradient Descent:The Gradient descent algorithm multiplies the gradient by a number (Learning rate or Step size) to determine the next point. W=W-(dW) b=b-(db)



$$z = w^T x + b$$

$$a = \sigma(z)$$

2.3 Gradient Descent Algorithm

Gradient descent is an optimization algorithm that's used when training a machine learning model. It's based on a convex function and tweaks its parameters iteratively to minimize a given function to its local minimum. W=W-(dW) b=b-(db)

2.4 Algorithm

Initialize (w,b) Optimize the loss iteratively to learn parameters (w,b): computing the cost and its gradient updating the parameters using gradient descent Use the learned (w,b) to predict the labels for a given set of examples

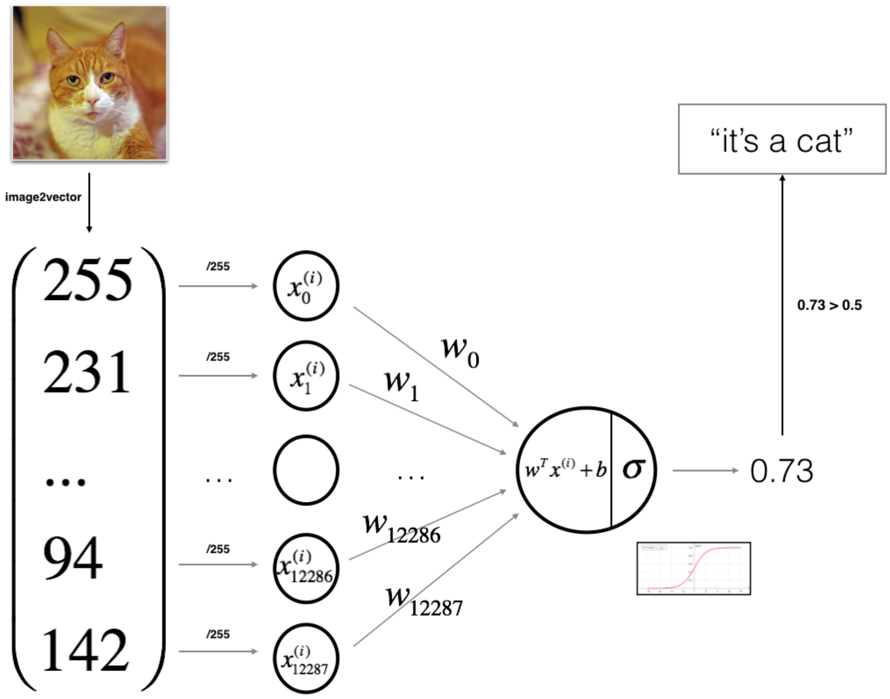


image2vector
255
231
...
94
142
/255
/255
/255
/255
$x_0^{(i)}$
$x_1^{(i)}$
$x_{12286}^{(i)}$
$x_{12287}^{(i)}$
w_0
w_1
w_{12286}
w_{12287}
$w^T x^{(i)} + b$
σ
0.73
0.73 > 0.5
"it's a cat"

Chapter 3

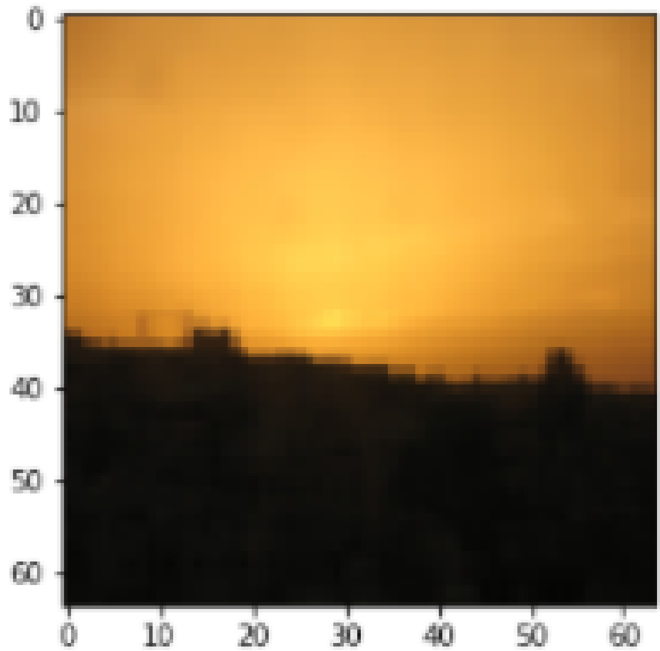
Results and Simulations

3.1 Software

Using the Jupyter NoteBook we have simulated the following results

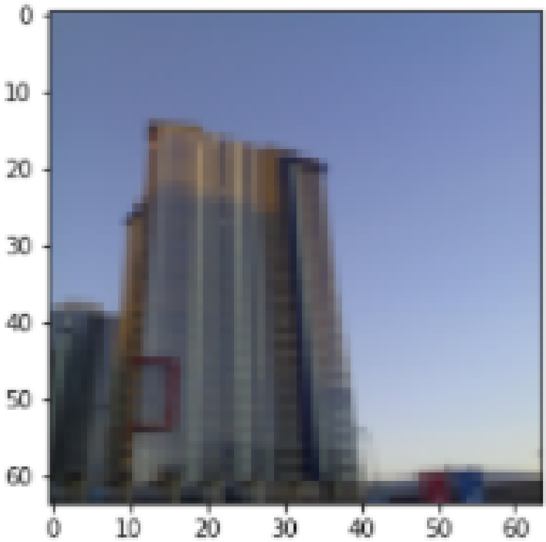
```
In [8]: # Example of a picture
        index = 80
        plt.imshow(train_set_x_orig[index])
        print ("y = " + str(train_set_y[:, i
```

```
y = [0], it's a 'non-cat' picture.
```



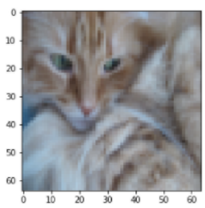
```
In [9]: # Example of a picture
        index = 90
        plt.imshow(train_set_x_orig[index])
        print ("y = " + str(train_set_y[:, i
```

```
y = [0], it's a 'non-cat' picture.
```




```
In [16]: # Example of a picture
         index = 2
         plt.imshow(train_set_x_orig[index])
         print ("y = " + str(train_set_y[:, index]) + ", it's a '" + classes[np.squeeze(train_set_y[:, index])].decode("utf-8") +  "' pic
```

```
y = [1], it's a 'cat' picture.
```



Chapter 4

Conclusions and Future works

4.1 Conclusions and Future works

The experimental results indicating that accuracy is 75 percent using Artificial Neural networks with the logistic regression.Here clearly know the implement of the 1 hidden layer shallow neural network.In future It can be developed to have Deep neural network which has multiple hidden layer which increases the accuracy of the result.By combing ANN with Convolution which gives the CNN results much easier way of extraction of the feature.It can be used to develop detection of live video image.

4.2 REFERENCES

4.2.1 www.google.com

4.2.2 www.wikipedia.com

4.2.3 www.researchgate.com

4.2.4 https://www.geeksforgeeks.org

Appendix A

Appendix A

Write your Appendix content here.

Appendix B

Appendix B

Write your Appendix content here.