MM-811

Assignment 3

Deep Learning

Submitted By:

Vaibhav Dixit

Submitted To:

Prof. Abram Hindle

Data Set Description:

The data set is taken from UCI machine learning repository. This data set includes description of hypothetical samples corresponding to 23 species of gilled mushroom in the Agaricus and Lepiota family. Here each species is identified as definitely edible, definitely poisonous or of unknown edibility and not recommended. There are 8124 number of instances and with 22 number of attributes in the data set.

The data set is saved with .csv format, change all the letters to numeric value.

The data set is available here: http://archive.ics.uci.edu/ml/datasets/Mushroom.

Problem Description:

For the mushroom data set, the task is to predict the edibility of the mushrooms.

Conditioned Input Description:

First attribute is reserved for the output and the other attribute are taken as an input. There are total 22 attributes.

Outputs Interpretation:

Hidden layer transforms an input in such a way that the output layer can use. So, according to my output, for getting good results, we should use more hidden layers.

Architecture Performance:

Architecture	Accuracy	TP	TN	FP	FN
1.	.9590	400	374	16	23
2.	.9891	414	389	1	9
3.	1.0	427	386	0	0

Possible Ways of Improving the result:

When we increase the number of hidden layer the accuracy of the architecture get improved, so it is possible to improve the output by increasing the hidden layers not being less than number of outputs.

Summary:

Most of the mushroom are classified correctly as poisonous or edible.

References:

https://github.com/abramhindle/theanets-tutorial

http://archive.ics.uci.edu/ml/datasets/Mushroom