Neural Networks

By Paul Abhishek

Table of Contents
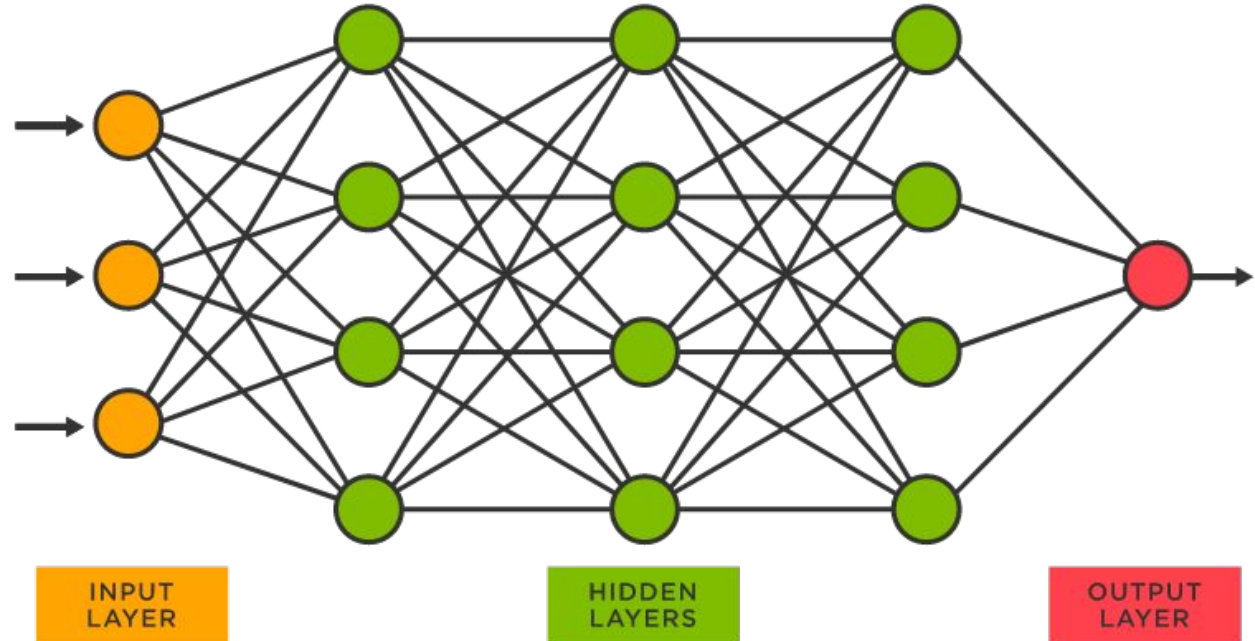
What is a neural network ?

A neural network is **a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain**. It is a type of machine learning process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain.

Importance :

Neural networks can **help computers make intelligent decisions with limited human assistance**. This is because they can learn and model the relationships between input and output data that are nonlinear and complex.
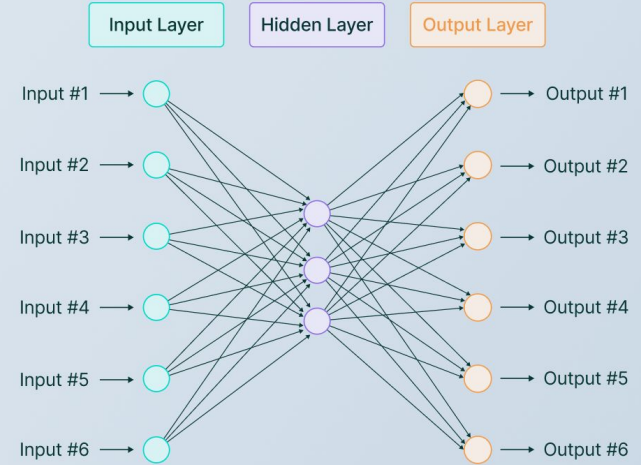

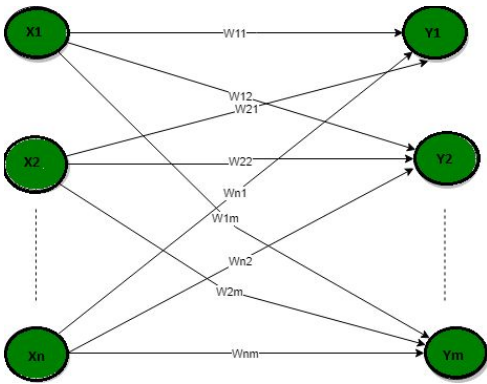INPUT LAYER

HIDDEN LAYERS

OUTPUT LAYER

Architecture

The architecture **provides the working parameters—such as the number, size, and type of layers in a neural network**. Models are one piece of your architecture; a specific instance that trains on a chosen set of data. For example, in a neural net, the trained weights of each node, per the architecture, comprise the model.

There exist five basic types of neuron connection architecture :
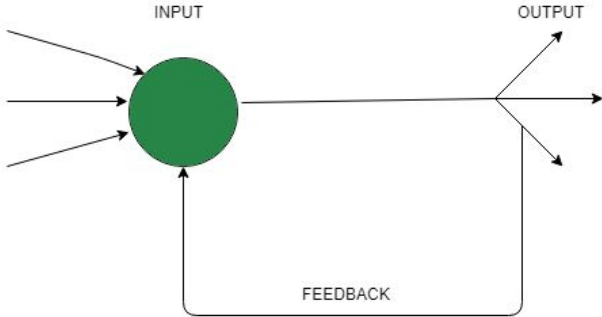
- Single-layer feed-forward network.
- Multilayer feed-forward network.
- Single node with its own feedback.
- Single-layer recurrent network.
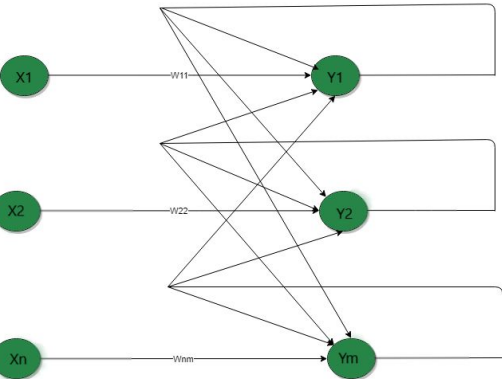- Multilayer recurrent network.
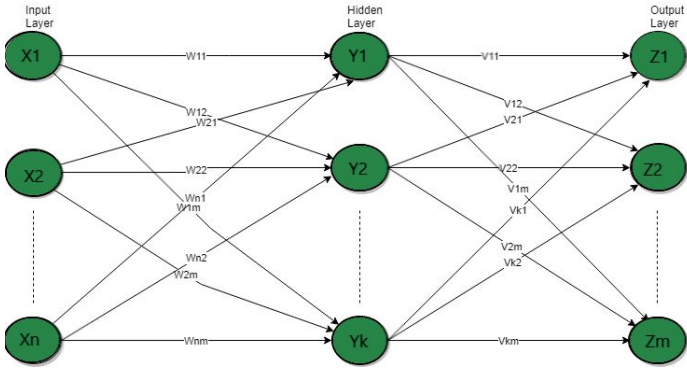
Single-layer feed-forward network

Multi Layer feedforward network

Single node with its own feedback

INPUT

OUTPUT

FEEDBACK

Single-layer recurrent network.

Multi - layer recurrent network
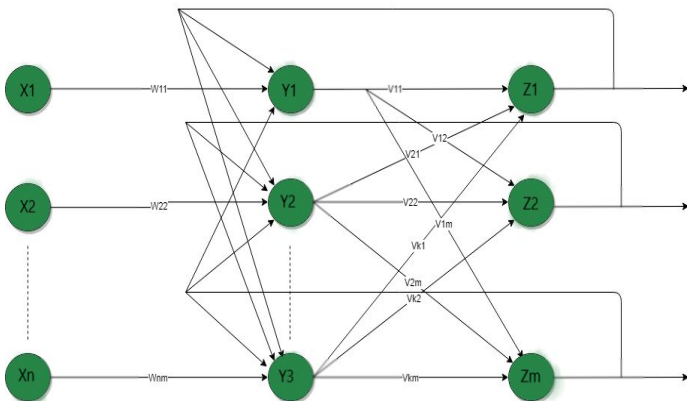
Optimizers

Optimizers are algorithms or methods used to change the attributes of your neural network such as weights and learning rate in order to reduce the losses. How you should change your weights or learning rates of your neural network to reduce the losses is defined by the optimizers you use.

Adam:

Adam optimization is a stochastic gradient descent method that is based on adaptive estimation of first-order and second-order moments.

SGD:

Stochastic gradient descent is an iterative method for optimizing an objective function with suitable smoothness properties. It can be regarded as a stochastic approximation of gradient descent optimization, since it replaces the actual gradient by an estimate thereof.

Epochs

An epoch means **training the neural network with all the training data for one cycle**. In an epoch, we use all of the data exactly once. A forward pass and a backward pass together are counted as one pass: An epoch is made up of one or more batches, where we use a part of the dataset to train the neural network.

What is a Neural Network Activation Function?

An Activation Function decides whether a neuron should be activated or not. This means that it will decide whether the neuron's input to the network is important or not in the process of prediction using simpler mathematical operations.

Relu :

ReLU. The ReLU function is another non-linear activation function that has gained popularity in the deep learning domain. ReLU stands for Rectified Linear Unit. The main advantage of using the ReLU function over other activation functions is that it does not activate all the neurons at the same time.
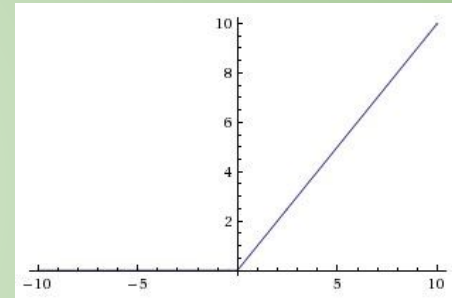
Sigmoid:

The Sigmoid function performs the role of an activation function in machine learning which is used to add non-linearity in a machine learning model. Basically, the function determines which value to pass as output and what not to pass as output.
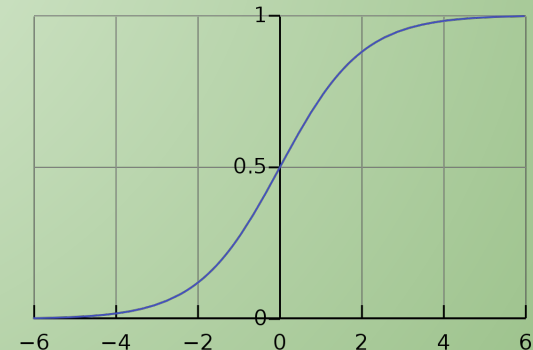
Auto Mpg

By Paul Abhishek

Background

MPG stands for miles per gallon and is used to show how far your car can travel for every gallon (or 4.55 litres) of fuel it uses. For example, if you own a car that returns 50mpg and its fuel tank only has one gallon of petrol or diesel in it, you'll drive 50 miles before the car runs out of fuel.

Objective

To determine the Miles Per Gallon (MPG) by the given inputs.

Path

The data set has a continuous outcome which varies from x to x . Preprocessing of the Data and Exploratory Data Analysis are done using python, pandas functions. Neural Network regression algorithm is applied on various train test ratios and run through Different architectures to obtain the best model which gives the lowest MAE.

Data and Data quality check

➜ The dataset consists of 142 rows and 20 columns

➜ 6 null values were found

➜ The attributes to be noticed are mpg,cylinders,displacement, horsepower, weight, acceleration.

➜ Object column (HorsePower) is converted into integer type.

➜ All the attributes are numeric Type.

➜ No outliers were found.

➜ Outcome is continuous

DataFrame

index	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin	car name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino
5	15.0	8	429.0	198	4341	10.0	70	1	ford galaxie 500
6	14.0	8	454.0	220	4354	9.0	70	1	chevrolet impala
7	14.0	8	440.0	215	4312	8.5	70	1	plymouth fury iii
8	14.0	8	455.0	225	4425	10.0	70	1	pontiac catalina
9	15.0	8	390.0	190	3850	8.5	70	1	amc ambassador dpl
10	15.0	8	383.0	170	3563	10.0	70	1	dodge challenger se
11	14.0	8	340.0	160	3609	8.0	70	1	plymouth 'cuda 340
12	15.0	8	400.0	150	3761	9.5	70	1	chevrolet monte carlo
13	14.0	8	455.0	225	3086	10.0	70	1	buick estate wagon (sw)
14	24.0	4	113.0	95	2372	15.0	70	3	toyota corona mark ii
15	22.0	6	198.0	95	2833	15.5	70	1	plymouth duster
16	18.0	6	199.0	97	2774	15.5	70	1	amc hornet
17	21.0	6	200.0	85	2587	16.0	70	1	ford maverick
18	27.0	4	97.0	88	2130	14.5	70	3	datsun pl510
19	26.0	4	97.0	46	1835	20.5	70	2	volkswagen 1131 deluxe sedan
20	25.0	4	110.0	87	2672	17.5	70	2	peugeot 504
21	24.0	4	107.0	90	2430	14.5	70	2	audi 100 ls
22	25.0	4	104.0	95	2375	17.5	70	2	saab 99e
23	26.0	4	121.0	113	2234	12.5	70	2	bmw 2002
24	21.0	6	199.0	90	2648	15.0	70	1	amc gremlin

Modified DataFrame

index	mpg	cylinders	displacement	horsepower	weight	acceleration
0	18.0	8	307.0	130	3504	12.0
1	15.0	8	350.0	165	3693	11.5
2	18.0	8	318.0	150	3436	11.0
3	16.0	8	304.0	150	3433	12.0
4	17.0	8	302.0	140	3449	10.5
5	15.0	8	429.0	198	4341	10.0
6	14.0	8	454.0	220	4354	9.0
7	14.0	8	440.0	215	4312	8.5
8	14.0	8	455.0	225	4425	10.0
9	15.0	8	390.0	190	3850	8.5
10	15.0	8	383.0	170	3563	10.0
11	14.0	8	340.0	160	3609	8.0
12	15.0	8	400.0	150	3761	9.5
13	14.0	8	455.0	225	3086	10.0
14	24.0	4	113.0	95	2372	15.0
15	22.0	6	198.0	95	2833	15.5
16	18.0	6	199.0	97	2774	15.5
17	21.0	6	200.0	85	2587	16.0
18	27.0	4	97.0	88	2130	14.5
19	26.0	4	97.0	46	1835	20.5
20	25.0	4	110.0	87	2672	17.5
21	24.0	4	107.0	90	2430	14.5
22	25.0	4	104.0	95	2375	17.5
23	26.0	4	121.0	113	2234	12.5
24	21.0	6	199.0	90	2648	15.0

PairPlots

Pair Plots are a really simple (one-line-of-code simple!) way to visualize relationships between each variable. It produces a matrix of relationships between each variable in your data for an instant examination of our data. It can also be a great jumping off point for determining types of regression analysis to use.
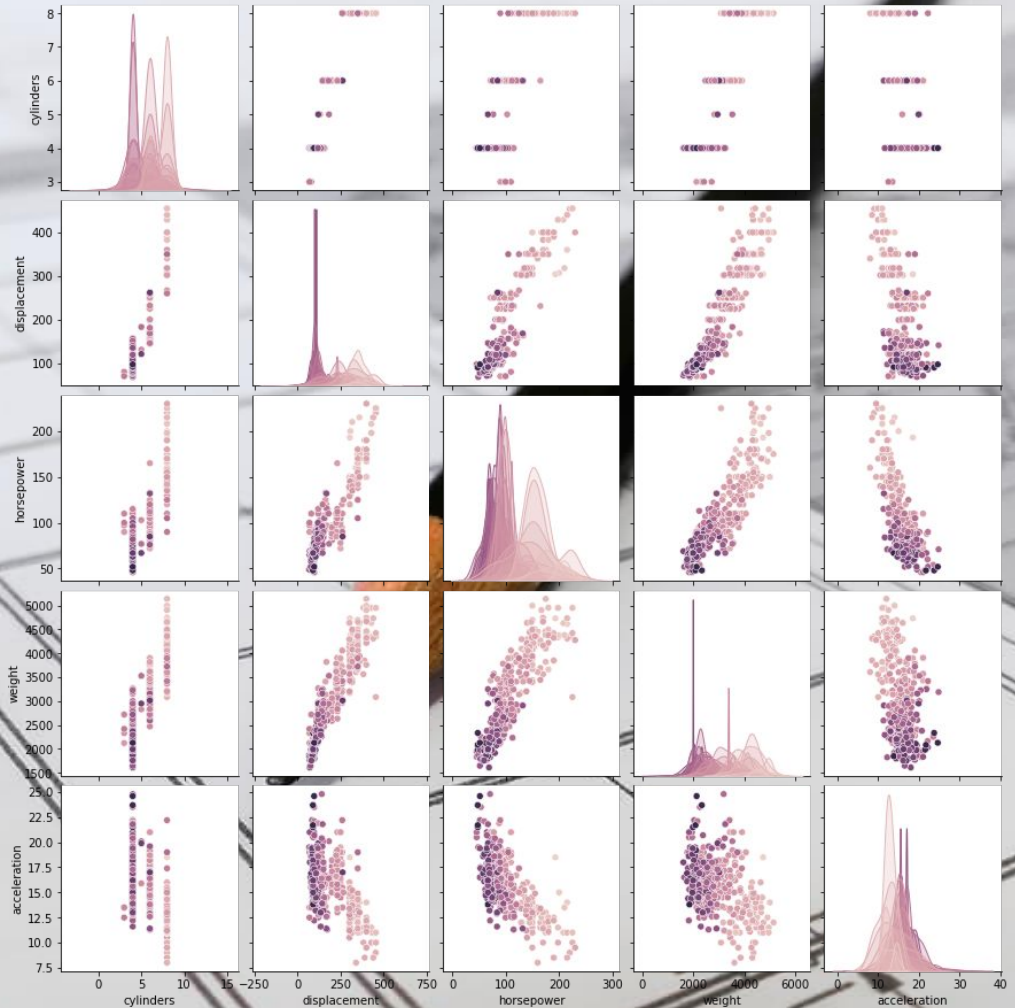
Heatmap

By definition, Heat Maps are graphical representations of data that utilize color-coded systems. The primary purpose of Heat Maps is to better visualize the volume of locations/events within a dataset and assist in directing viewers towards areas on data visualizations that matter most.In this dataset most of the attributes are highly correlated with each other.
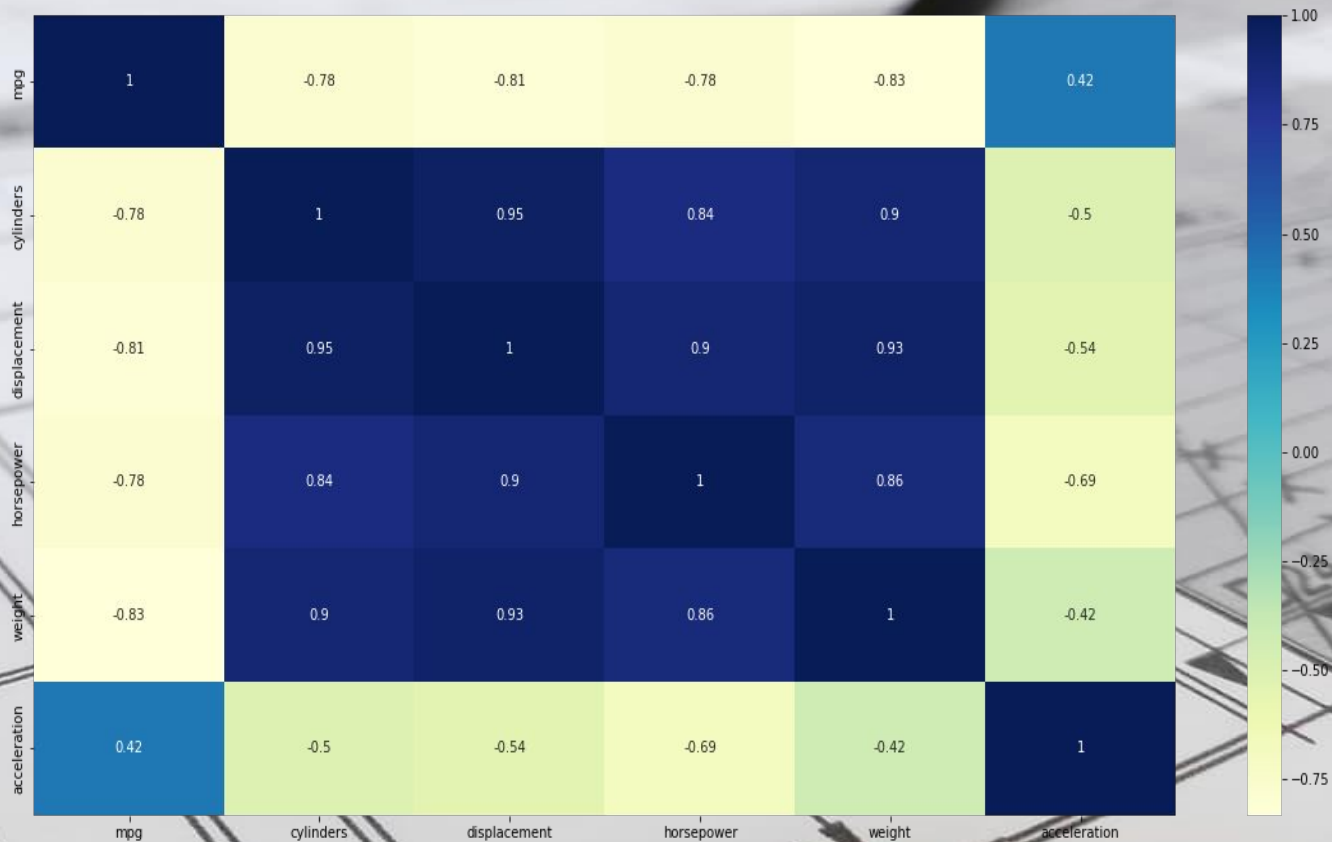
Combinations used while experimenting

Test Size - Train size:

- 20 - 80
- 25 - 75
- 30 - 70
- 40 - 80

Architectures:

- 10-5-3-1
- 10-6-3-2-1
- 6-3-2-1
- 6-3-1

Epochs:

- 100
- 200
- 300
- 400

Optimizers:

- Adam
- SGD
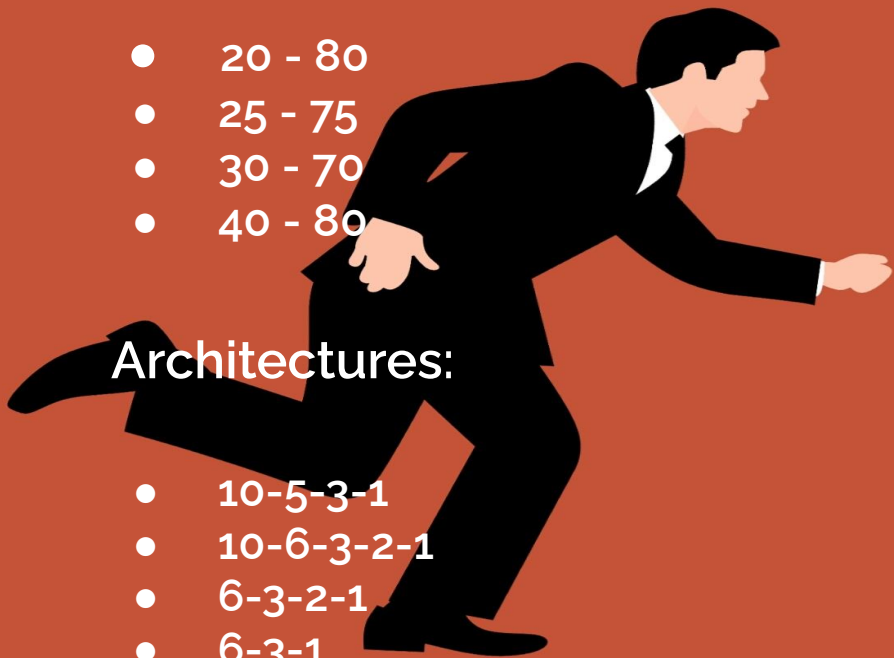
train test Proportion	Architecture	Optimizer	Epochs	MAE
20-80	10-5-3-1	Adam	100	5.6974
20-80	10-5-3-1	Adam	200	4.3886
20-80	10-5-3-1	Adam	300	5.1916
20-80	10-5-3-1	Adam	400	4.1584
20-80	10-5-3-1	SGD	100	nan
20-80	10-5-3-1	SGD	200	nan
20-80	10-5-3-1	SGD	300	nan
20-80	10-5-3-1	SGD	400	nan
20-80	10-6-3-2-1	Adam	100	6.1529
20-80	10-6-3-2-1	Adam	200	4.6579
20-80	10-6-3-2-1	Adam	300	4.1389
20-80	10-6-3-2-1	Adam	400	4.7107
20-80	10-6-3-2-1	SGD	100	nan
20-80	10-6-3-2-1	SGD	200	nan
20-80	10-6-3-2-1	SGD	300	nan
20-80	10-6-3-2-1	SGD	400	nan
20-80	6-3-2-1	Adam	100	5.1082
20-80	6-3-2-1	Adam	200	4.4675

20-80	6-3-2-1	Adam	300	6.457
20-80	6-3-2-1	Adam	400	4.5726
20-80	6-3-2-1	SGD	100	nan
20-80	6-3-2-1	SGD	200	nan
20-80	6-3-2-1	SGD	300	nan
20-80	6-3-2-1	SGD	400	nan
20-80	6-3--1	Adam	100	11.5905
20-80	6-3--1	Adam	200	4.4426
20-80	6-3--1	Adam	300	6.6574
20-80	6-3--1	Adam	400	4.2219
20-80	6-3--1	SGD	100	nan
20-80	6-3--1	SGD	200	nan
20-80	6-3--1	SGD	300	nan
20-80	6-3--1	SGD	400	nan
25-75	10-5-3-1	Adam	100	5.8217
25-75	10-5-3-1	Adam	200	5.4238
25-75	10-5-3-1	Adam	300	6.3053
25-75	10-5-3-1	Adam	400	5.6485
25-75	10-5-3-1	SGD	100	nan

25-75	10-5-3-1	SGD	200	nan
25-75	10-5-3-1	SGD	300	nan
25-75	10-5-3-1	SGD	400	nan
25-75	10-6-3-2-1	Adam	100	5.568
25-75	10-6-3-2-1	Adam	200	6.6522
25-75	10-6-3-2-1	Adam	300	4.424
25-75	10-6-3-2-1	Adam	400	6.0586
25-75	10-6-3-2-1	SGD	100	nan
25-75	10-6-3-2-1	SGD	200	nan
25-75	10-6-3-2-1	SGD	300	nan
25-75	10-6-3-2-1	SGD	400	nan
25-75	6-3-2-1	Adam	100	6.1636
25-75	6-3-2-1	Adam	200	5.226
25-75	6-3-2-1	Adam	300	4.0248
25-75	6-3-2-1	Adam	400	6.4666
25-75	6-3-2-1	SGD	100	nan
25-75	6-3-2-1	SGD	200	nan
25-75	6-3-2-1	SGD	300	nan
25-75	6-3-2-1	SGD	400	nan

25-75	6-3--1	Adam	100	5.9229
25-75	6-3-1	Adam	200	4.8127
25-75	6-3--1	Adam	300	4.77
25-75	6-3--1	Adam	400	10.548
25-75	6-3--1	SGD	100	nan
25-75	6-3--1	SGD	200	nan
25-75	6-3-1	SGD	300	nan
25-75	6-3-1	SGD	400	nan
30-70	10-5-3-1	Adam	100	5.8217
30-70	10-5-3-1	Adam	200	4.7885
30-70	10-5-3-1	Adam	300	4.5742
30-70	10-5-3-1	Adam	400	5.065
30-70	10-5-3-1	SGD	100	nan
30-70	10-5-3-1	SGD	200	nan
30-70	10-5-3-1	SGD	300	nan
30-70	10-5-3-1	SGD	400	nan
30-70	10-6-3-2-1	Adam	100	5.2788
30-70	10-6-3-2-1	Adam	200	4.3113
30-70	10-6-3-2-1	Adam	300	5.2044

30-70	10-6-3-2-1	Adam	400	6.0586
30-70	10-6-3-2-1	SGD	100	nan
30-70	10-6-3-2-1	SGD	200	nan
30-70	10-6-3-2-1	SGD	300	nan
30-70	10-6-3-2-1	SGD	400	nan
30-70	6-3-2-1	Adam	100	5.9337
30-70	6-3-2-1	Adam	200	4.6061
30-70	6-3-2-1	Adam	300	4.8186
30-70	6-3-2-1	Adam	400	4.2779
30-70	6-3-2-1	SGD	100	nan
30-70	6-3-2-1	SGD	200	nan
30-70	6-3-2-1	SGD	300	nan
30-70	6-3-2-1	SGD	400	nan
30-70	6-3--1	Adam	100	7.4012
30-70	6-3--1	Adam	200	12.2685
30-70	6-3--1	Adam	300	6.2578
30-70	6-3--1	Adam	400	5.4085
30-70	6-3--1	SGD	100	nan
30-70	6-3--1	SGD	200	nan

30-70	6-3--1	SGD	300	nan
30-70	6-3-1	SGD	400	nan
35-75	10-5-3-1	Adam	100	6.8723
35-75	10-5-3-1	Adam	200	4.5116
35-75	10-5-3-1	Adam	300	4.7849
35-75	10-5-3-1	Adam	400	4.3661
35-75	10-5-3-1	SGD	100	nan
35-75	10-5-3-1	SGD	200	nan
35-75	10-5-3-1	SGD	300	nan
35-75	10-5-3-1	SGD	400	nan
35-75	10-6-3-2-1	Adam	100	7.5295
35-75	10-6-3-2-1	Adam	200	6.1945
35-75	10-6-3-2-1	Adam	300	4.6792
35-75	10-6-3-2-1	Adam	400	4.3261
35-75	10-6-3-2-1	SGD	100	nan
35-75	10-6-3-2-1	SGD	200	nan
35-75	10-6-3-2-1	SGD	300	nan
35-75	10-6-3-2-1	SGD	400	nan
35-75	6-3-2-1	Adam	100	6.0759

35-75	6-3-2-1	Adam	200	4.727
35-75	6-3-2-1	Adam	300	6.109
35-75	6-3-2-1	Adam	400	4.397
35-75	6-3-2-1	SGD	100	nan
35-75	6-3-2-1	SGD	200	nan
35-75	6-3-2-1	SGD	300	nan
35-75	6-3-2-1	SGD	400	nan
35-75	6-3--1	Adam	100	6.0983
35-75	6-3--1	Adam	200	7.2459
35-75	6-3--1	Adam	300	4.9364
35-75	6-3--1	Adam	400	18.2616
35-75	6-3--1	SGD	100	nan
35-75	6-3--1	SGD	200	nan
35-75	6-3--1	SGD	300	nan
35-75	6-3--1	SGD	400	nan
40-60	10-5-3-1	Adam	100	8.3044
40-60	10-5-3-1	Adam	200	5.0884
40-60	10-5-3-1	Adam	300	5.0592
40-60	10-5-3-1	Adam	400	6.8075

40-60	6-3-2-1	SGD	400	nan
40-60	6-3--1	Adam	100	5.4525
40-60	6-3--1	Adam	200	37.3416
40-60	6-3-1	Adam	300	5.9473
40-60	6-3-1	Adam	400	8.0715
40-60	6-3--1	SGD	100	nan
40-60	6-3--1	SGD	200	nan
40-60	6-3--1	SGD	300	nan
40-60	6-3--1	SGD	400	nan

Best combination for lowest MAE

Test size : 30%

Architecture : 6-3-2-1

Optimizer : Adam

Epochs : 400

MAE : 4.2279
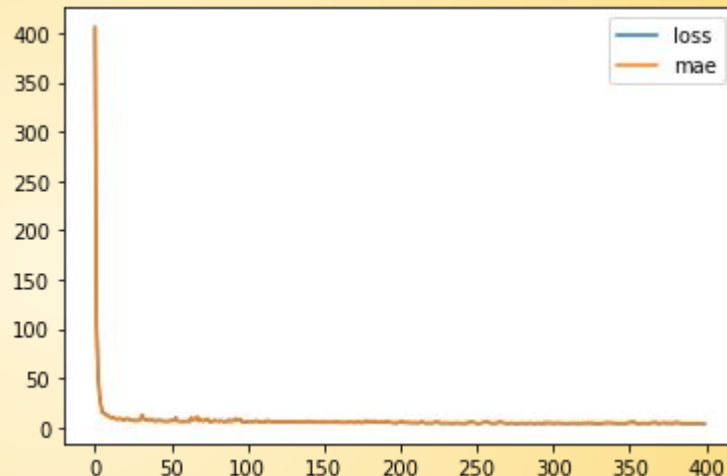
30-70	6-3-2-1	Adam	400	4.2779

Worst case scenario

Test size : 40%

Architecture : 6-3-1

Optimizer : Adam

Epochs : 200

MAE : 37.3416
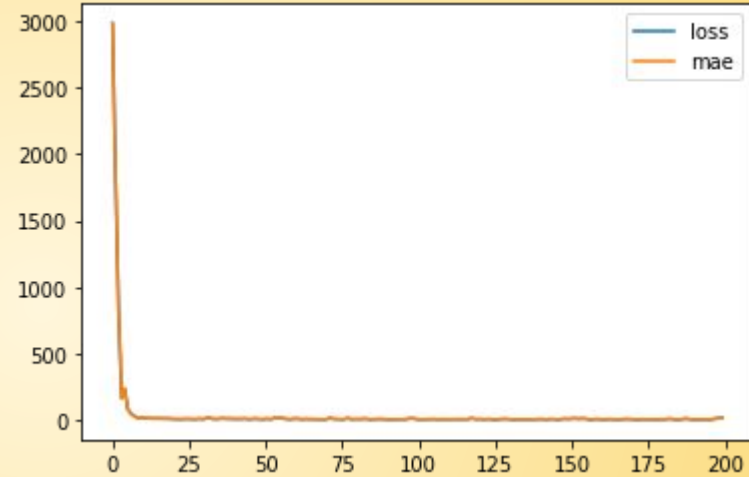
40-60	6-3--1	Adam	200	37.3416

Conclusion

➢ The results form the table clearly states that the best optimizer out there is Adam. The reason it is best because at most of the times it gives higher accuracy when compared to SGD.
➢ Nan values were showing up when the optimizer was SGD.
➢ So, it is suggested to use Adam optimizer over SGD.

Best Results :

Test size   : 30%

Architecture : 6-3-2-1

Optimizer   : Adam

Epochs   :   400   : 400

MAE   : 4.2279

Worst Results :

Test size   : 40%

Architecture : 6-3-1

Optimizer   : Adam

Epochs   :   400   : 400

MAE   : 37.346

HEPATITIS

- By Paul Abhishek

Background

Hepatitis means inflammation of the liver. The liver is a vital organ that processes nutrients, filters the blood, and fights infections. When the liver is inflamed or damaged, its function can be affected. Heavy alcohol use, toxins, some medications, and certain medical conditions can cause hepatitis.

Objective

The main objective of the dataset is to predict the chance of survival of a person whose suffering from hepatitis.

Path

The dataset is is a binary outcome dataset. It is a yes/no dataset. Preprocessing of the Data and Exploratory Data Analysis is applied. Neural Network in the category of the classification is used to determine the accuracy of the data set.

Data and Data quality check

➔   The dataset consists of 142 rows and 20 columns

➔   Every attribute is important in determining the output.

➔   The outcome is binary. (live/dead)

➔   No missing values were found in the dataset.

➔   Output attribute : 'Class'

➔   Binary Outcome

DataFrame

index	class	age	sex	steroid	antivirals	fatigue	malaise	anorexia	liver_big	liver_firm	spleen_palable	spiders	ascites	varices	bilirubin	alk_phosphate	sgot	albumin	protime	histology
0	2	30	2	1	2	2	2	2	1	2	2	2	2	2	1.0	85	18	4.0	61	1
1	2	50	1	1	2	1	2	2	1	2	2	2	2	2	0.9	135	42	3.5	61	1
2	2	78	1	2	2	1	2	2	2	2	2	2	2	2	0.7	96	32	4.0	61	1
3	2	34	1	2	2	2	2	2	2	2	2	2	2	2	1.0	105	200	4.0	61	1
4	2	34	1	2	2	2	2	2	2	2	2	2	2	2	0.9	95	28	4.0	75	1
5	1	51	1	1	2	1	2	1	2	2	1	1	2	2	1.42	105	85	3.81	61	1
6	2	23	1	2	2	2	2	2	2	2	2	2	2	2	1.0	105	85	3.81	61	1
7	2	39	1	2	2	1	2	2	2	1	2	2	2	2	0.7	105	48	4.4	61	1
8	2	30	1	2	2	2	2	2	2	2	2	2	2	2	1.0	105	120	3.9	61	1
9	2	39	1	1	1	2	2	2	1	1	2	2	2	2	1.3	78	30	4.4	85	1
10	2	32	1	2	1	1	2	2	2	1	2	1	2	2	1.0	59	249	3.7	54	1
11	2	41	1	2	1	1	2	2	2	1	2	2	2	2	0.9	81	60	3.9	52	1
12	2	30	1	2	2	1	2	2	2	1	2	2	2	2	2.2	57	144	4.9	78	1
13	2	47	1	1	1	2	2	2	2	2	2	2	2	2	1.42	105	60	3.81	61	1
14	2	38	1	1	2	1	1	1	2	2	2	2	1	2	2.0	72	89	2.9	46	1
15	2	66	1	2	2	1	2	2	2	2	2	2	2	2	1.2	102	53	4.3	61	1
16	2	40	1	1	2	1	2	2	2	1	2	2	2	2	0.6	62	166	4.0	63	1
17	2	38	1	2	2	2	2	2	2	2	2	2	2	2	0.7	53	42	4.1	85	2
18	2	38	1	1	1	2	2	2	1	1	2	2	2	2	0.7	70	28	4.2	62	1
19	2	22	2	2	1	1	2	2	2	2	2	2	2	2	0.9	48	20	4.2	64	1
20	2	27	1	2	2	1	1	1	1	1	1	1	2	2	1.2	133	98	4.1	39	1
21	2	31	1	2	2	2	2	2	2	2	2	2	2	2	1.0	85	20	4.0	100	1
22	2	42	1	2	2	2	2	2	2	2	2	2	2	2	0.9	60	63	4.7	47	1
23	2	25	2	1	1	2	2	2	2	2	2	2	2	2	0.4	45	18	4.3	70	1
24	2	27	1	1	2	1	1	2	2	2	2	2	2	2	0.8	95	46	3.8	100	1

Modified DataFrame

| index	class	age	sex	steroid	antivirals	fatigue	malaise	anorexia	liver_big	liver_firm	spleen_palable	spiders	ascites	varices	histology	bilirubin	alk_phosphate	sgot	albumin	protime
0	1	30	1	0	1	1	1	1	0	1	1	1	1	1	0	1.0	85	18	4.0	61
1	1	50	0	0	1	0	1	1	0	1	1	1	1	1	0	0.9	135	42	3.5	61
2	1	78	0	1	1	0	1	1	1	1	1	1	1	1	0	0.7	96	32	4.0	61
3	1	34	0	1	1	1	1	1	1	1	1	1	1	1	0	1.0	105	200	4.0	61
4	1	34	0	1	1	1	1	1	1	1	1	1	1	1	0	0.9	95	28	4.0	75
5	0	51	0	0	1	0	1	0	1	1	0	0	1	1	0	1.42	105	85	3.81	61
6	1	23	0	1	1	1	1	1	1	1	1	1	1	1	0	1.0	105	85	3.81	61
7	1	39	0	1	1	0	1	1	1	0	1	1	1	1	0	0.7	105	48	4.4	61
8	1	30	0	1	1	1	1	1	1	1	1	1	1	1	0	1.0	105	120	3.9	61
9	1	39	0	0	0	1	1	0	0	1	1	1	1	1	0	1.3	78	30	4.4	85
10	1	32	0	1	0	0	1	1	0	1	1	0	1	1	0	1.0	59	249	3.7	54
11	1	41	0	1	0	0	1	1	0	1	1	1	1	1	0	0.9	81	60	3.9	52
12	1	30	0	1	1	0	1	1	0	1	1	1	1	1	0	2.2	57	144	4.9	78
13	1	47	0	0	0	1	1	1	1	1	1	1	1	1	0	1.42	105	60	3.81	61
14	1	38	0	0	1	0	0	1	1	1	1	0	1	1	0	2.0	72	89	2.9	46
15	1	66	0	1	1	0	1	1	1	1	1	1	1	1	0	1.2	102	53	4.3	61
16	1	40	0	0	1	0	1	1	0	1	1	1	1	1	0	0.6	62	166	4.0	63
17	1	38	0	1	1	1	1	1	1	1	1	1	1	1	1	0.7	53	42	4.1	85
18	1	38	0	0	0	1	1	1	0	0	1	1	1	1	0	0.7	70	28	4.2	62
19	1	22	1	1	0	0	1	1	1	1	1	1	1	1	0	0.9	48	20	4.2	64
20	1	27	0	1	1	0	0	0	0	0	0	0	1	1	0	1.2	133	98	4.1	39
21	1	31	0	1	1	1	1	1	1	1	1	1	1	1	0	1.0	85	20	4.0	100
22	1	42	0	1	1	1	1	1	1	1	1	1	1	1	0	0.9	60	63	4.7	47
23	1	25	1	0	0	1	1	1	1	1	1	1	1	1	0	0.4	45	18	4.3	70
24	1	27	0	0	1	0	0	1	1	1	1	1	1	1	0	0.8	95	46	3.8	100

People died vs survived
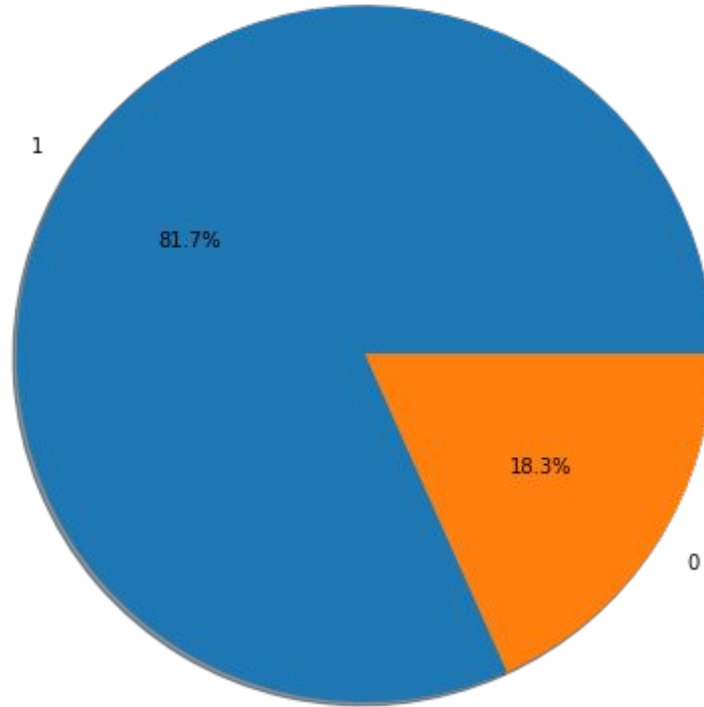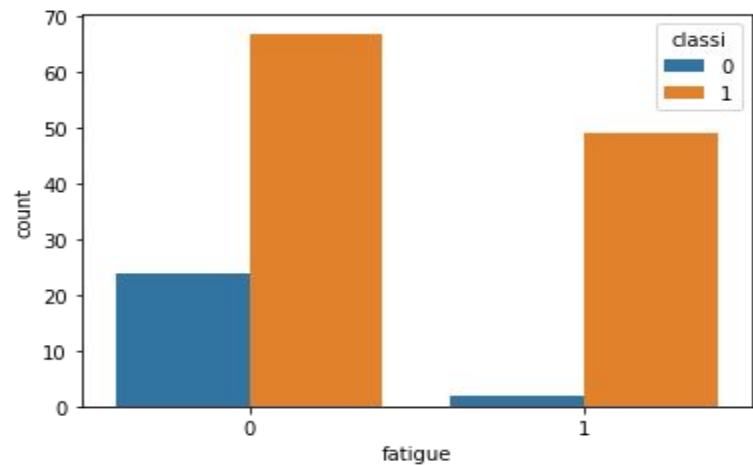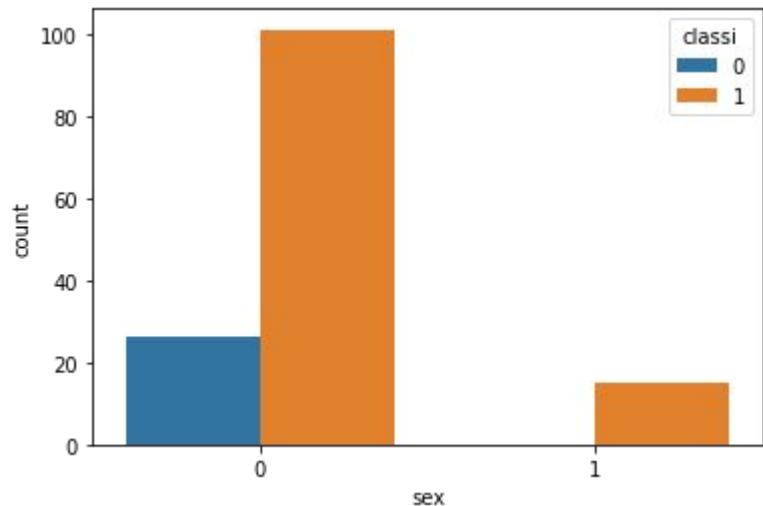
Count plots
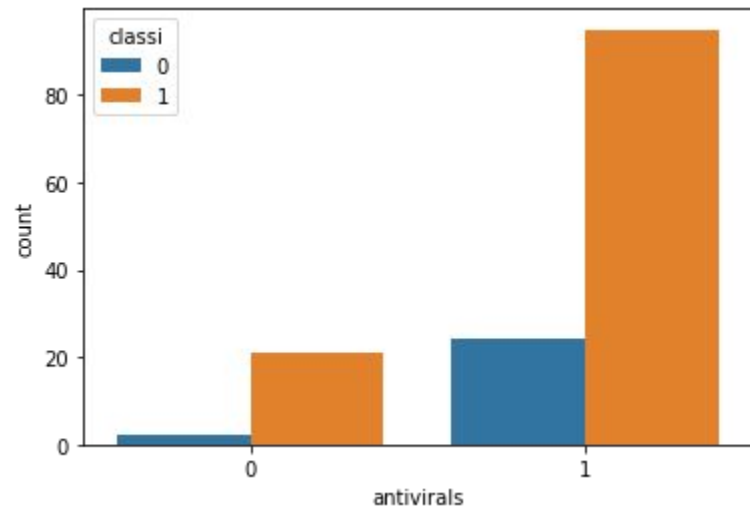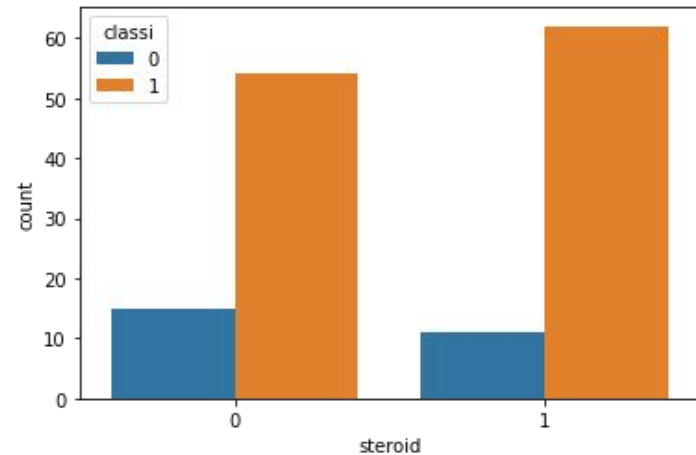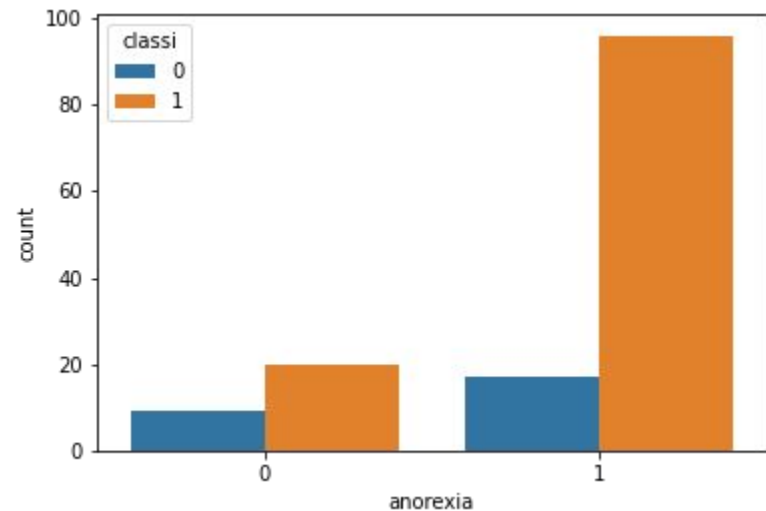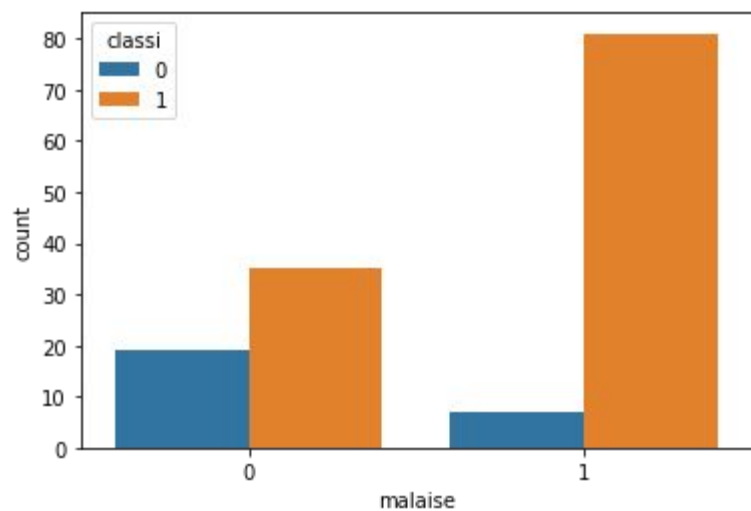
Count plots
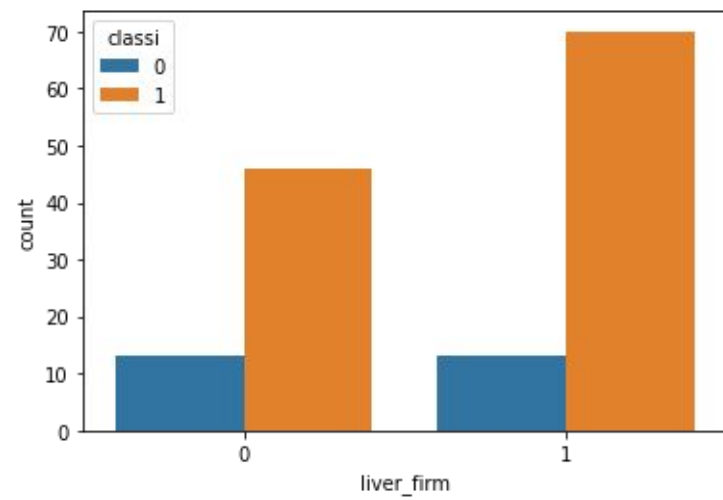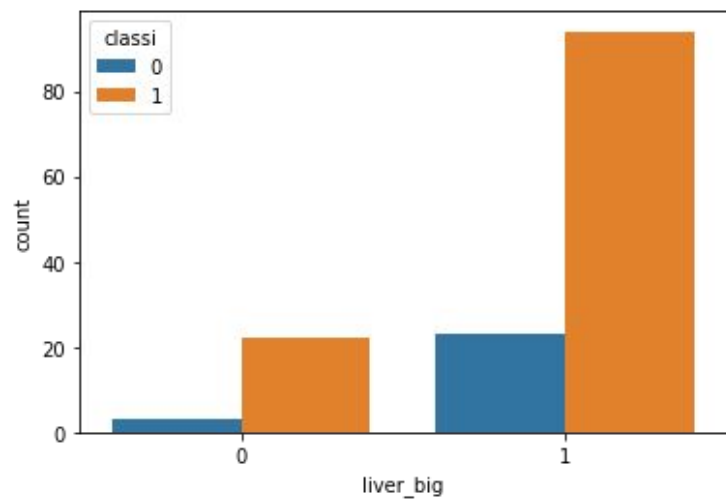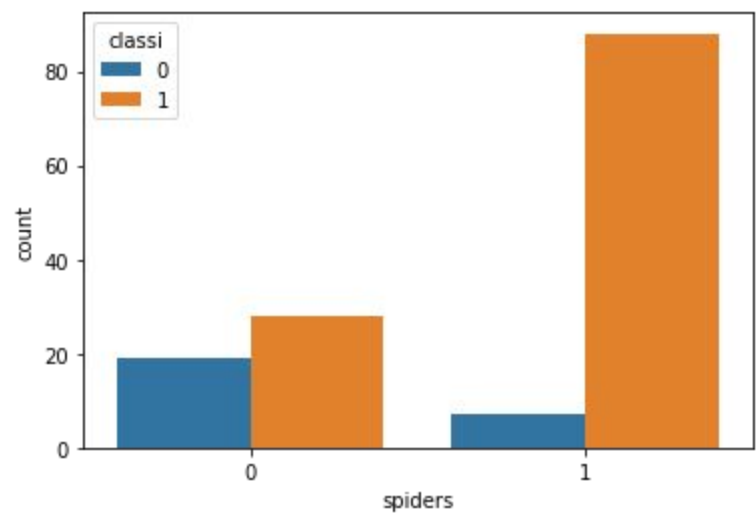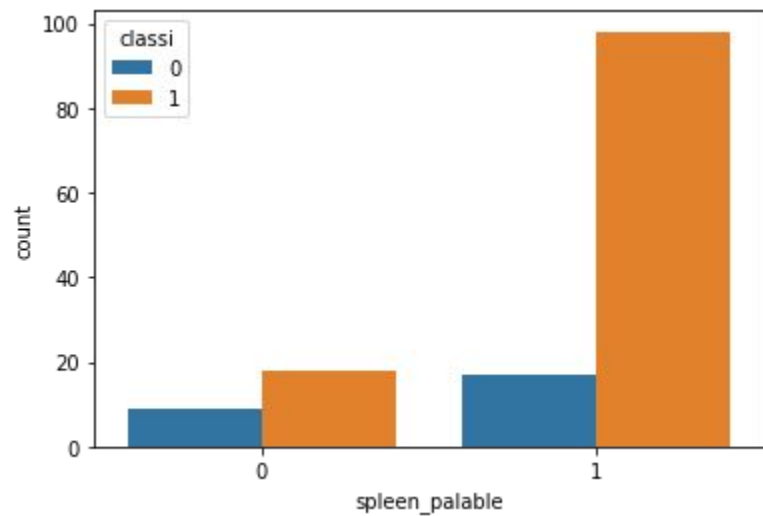
Count plots
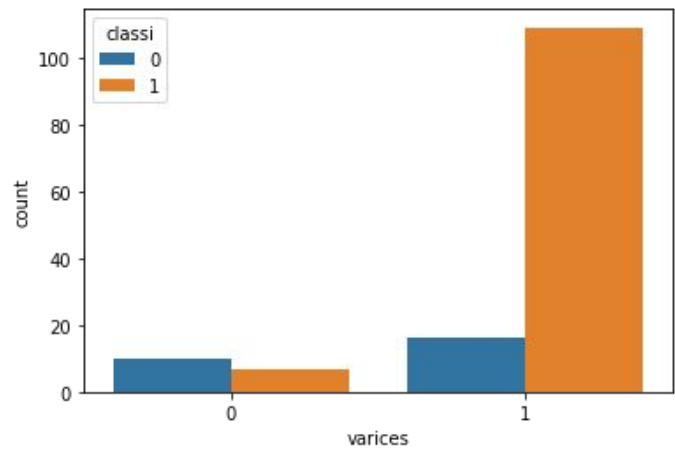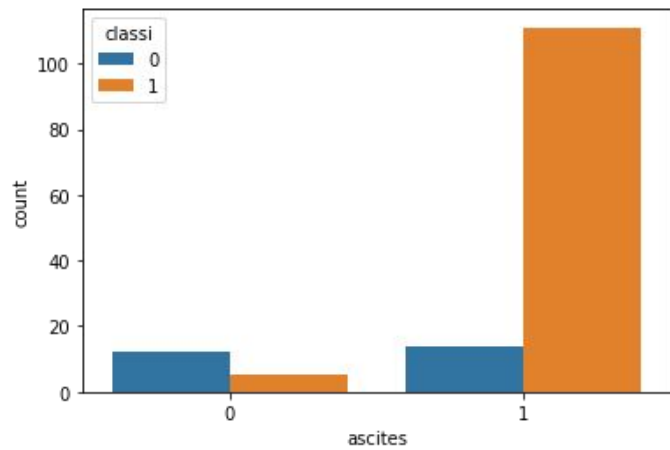
Take away from Graphs

➢ People with antivirals have a very high chance of surviving.
➢ People with less fatigue have a high chance of surviving.
➢ People with anorexia have a high chance of surviving
➢ In the graph steroids don't play a major role in detecting the chances of survival

Combinations used while experimenting

Test Size - Train size:

- 20 - 80
- 25 - 75
- 30 - 70
- 40 - 80

Epochs:

- 100
- 200
- 300
- 400

Architectures:

- 10-5-3-1
- 10-6-3-2-1
- 6-3-2-1
- 6-3-1

Optimizers:

- Adam
- SGD

Experimentation

train test Proportion	Architecture	Optimizer	Epochs	Accuracy
20-80	10-5-3-1	Adam	100	0.8966
20-80	10-5-3-1	Adam	200	0.8966
20-80	10-5-3-1	Adam	300	0.8966
20-80	10-5-3-1	Adam	400	0.8966
20-80	10-5-3-1	SGD	100	0.8966
20-80	10-5-3-1	SGD	200	0.8966
20-80	10-5-3-1	SGD	300	0.8966
20-80	10-5-3-1	SGD	400	0.8966
20-80	10-6-3-2-1	Adam	100	0.8966
20-80	10-6-3-2-1	Adam	200	0.8966
20-80	10-6-3-2-1	Adam	300	0.8966
20-80	10-6-3-2-1	Adam	400	0.8966
20-80	10-6-3-2-1	SGD	100	0.8966

20-80	10-6-3-2-1	SGD	200	0.8966
20-80	10-6-3-2-1	SGD	300	0.8966
20-80	10-6-3-2-1	SGD	400	0.8966
20-80	6-3-2-1	Adam	100	0.8966
20-80	6-3-2-1	Adam	200	0.8621
20-80	6-3-2-1	Adam	300	0.8621
20-80	6-3-2-1	Adam	400	0.8621
20-80	6-3-2-1	SGD	100	0.8966
20-80	6-3-2-1	SGD	200	0.8966
20-80	6-3-2-1	SGD	300	0.8966
20-80	6-3-2-1	SGD	400	0.8966
20-80	6-3--1	Adam	100	0.8966
20-80	6-3--1	Adam	200	0.8966
20-80	6-3--1	Adam	300	0.8966

20-80	6-3--1	Adam	400	0.8966
20-80	6-3--1	SGD	100	0.8966
20-80	6-3--1	SGD	200	0.8966
20-80	6-3--1	SGD	300	0.8966
20-80	6-3--1	SGD	400	0.8966
25-75	10-5-3-1	Adam	100	0.8889
25-75	10-5-3-1	Adam	200	0.8889
25-75	10-5-3-1	Adam	300	0.8889
25-75	10-5-3-1	Adam	400	0.8889
25-75	10-5-3-1	SGD	100	0.8889
25-75	10-5-3-1	SGD	200	0.8889
25-75	10-5-3-1	SGD	300	0.8889
25-75	10-5-3-1	SGD	400	0.8889
25-75	10-6-3-2-1	Adam	100	0.8889

25-75	10-6-3-2-1	Adam	200	0.8889
25-75	10-6-3-2-1	Adam	300	0.8889
25-75	10-6-3-2-1	Adam	400	0.8889
25-75	10-6-3-2-1	SGD	100	0.8889
25-75	10-6-3-2-1	SGD	200	0.8889
25-75	10-6-3-2-1	SGD	300	0.8889
25-75	10-6-3-2-1	SGD	400	0.8889
25-75	6-3-2-1	Adam	100	0.8056
25-75	6-3-2-1	Adam	200	0.8611
25-75	6-3-2-1	Adam	300	0.8333
25-75	6-3-2-1	Adam	400	0.8611
25-75	6-3-2-1	SGD	100	0.8889
25-75	6-3-2-1	SGD	200	0.8889
25-75	6-3-2-1	SGD	300	0.8889

25-75	6-3-2-1	SGD	400	0.8889
25-75	6-3--1	Adam	100	0.8889
25-75	6-3--1	Adam	200	0.8889
25-75	6-3--1	Adam	300	0.8889
25-75	6-3--1	Adam	400	0.8889
25-75	6-3--1	SGD	100	0.8889
25-75	6-3--1	SGD	200	0.8889
25-75	6-3--1	SGD	300	0.8889
25-75	6-3--1	SGD	400	0.8889
30-70	10-5-3-1	Adam	100	0.8837
30-70	10-5-3-1	Adam	200	0.8837
30-70	10-5-3-1	Adam	300	0.8837
30-70	10-5-3-1	Adam	400	0.8837
30-70	10-5-3-1	SGD	100	0.8837

30-70	10-5-3-1	SGD	200	0.8837
30-70	10-5-3-1	SGD	300	0.8837
30-70	10-5-3-1	SGD	400	0.8837
30-70	10-6-3-2-1	Adam	100	0.8837
30-70	10-6-3-2-1	Adam	200	0.8837
30-70	10-6-3-2-1	Adam	300	0.8837
30-70	10-6-3-2-1	Adam	400	0.8837
30-70	10-6-3-2-1	SGD	100	0.8837
30-70	10-6-3-2-1	SGD	200	0.8837
30-70	10-6-3-2-1	SGD	300	0.8837
30-70	10-6-3-2-1	SGD	400	0.8837
30-70	6-3-2-1	Adam	100	0.814
30-70	6-3-2-1	Adam	200	0.8837
30-70	6-3-2-1	Adam	300	0.8372

30-70	6-3-2-1	Adam	400	0.8837
30-70	6-3-2-1	SGD	100	0.8837
30-70	6-3-2-1	SGD	200	0.8837
30-70	6-3-2-1	SGD	300	0.8837
30-70	6-3-2-1	SGD	400	0.8837
30-70	6-3--1	Adam	100	0.8372
30-70	6-3--1	Adam	200	0.8837
30-70	6-3--1	Adam	300	0.8837
30-70	6-3--1	Adam	400	0.8837
30-70	6-3--1	SGD	100	0.8837
30-70	6-3--1	SGD	200	0.8837
30-70	6-3--1	SGD	300	0.8837
30-70	6-3--1	SGD	400	0.8837
35-75	10-5-3-1	Adam	100	0.86

35-75	10-5-3-1	Adam	200	0.9
35-75	10-5-3-1	Adam	300	0.78
35-75	10-5-3-1	Adam	400	0.8
35-75	10-5-3-1	SGD	100	0.86
35-75	10-5-3-1	SGD	200	0.8
35-75	10-5-3-1	SGD	300	0.86
35-75	10-5-3-1	SGD	400	0.86
35-75	10-6-3-2-1	Adam	100	0.86
35-75	10-6-3-2-1	Adam	200	0.86
35-75	10-6-3-2-1	Adam	300	0.86
35-75	10-6-3-2-1	Adam	400	0.86
35-75	10-6-3-2-1	SGD	100	0.86
35-75	10-6-3-2-1	SGD	200	0.86
35-75	10-6-3-2-1	SGD	300	0.86

35-75	10-6-3-2-1	SGD	400	0.86
35-75	6-3-2-1	Adam	100	0.82
35-75	6-3-2-1	Adam	200	0.82
35-75	6-3-2-1	Adam	300	0.82
35-75	6-3-2-1	Adam	400	0.54
35-75	6-3-2-1	SGD	100	0.86
35-75	6-3-2-1	SGD	200	0.86
35-75	6-3-2-1	SGD	300	0.86
35-75	6-3-2-1	SGD	400	0.86
35-75	6-3--1	Adam	100	0.84
35-75	6-3--1	Adam	200	0.86
35-75	6-3--1	Adam	300	0.86
35-75	6-3--1	Adam	400	0.88
35-75	6-3--1	SGD	100	0.86

35-75	6-3--1	SGD	200	0.86
35-75	6-3--1	SGD	300	0.86
35-75	6-3--1	SGD	400	0.86
40-60	10-5-3-1	Adam	100	0.8596
40-60	10-5-3-1	Adam	200	0.8596
40-60	10-5-3-1	Adam	300	0.8596
40-60	10-5-3-1	Adam	400	0.8596
40-60	10-5-3-1	SGD	100	0.8596
40-60	10-5-3-1	SGD	200	0.8596
40-60	10-5-3-1	SGD	300	0.8596
40-60	10-5-3-1	SGD	400	0.8596
40-60	10-6-3-2-1	Adam	100	0.8596
40-60	10-6-3-2-1	Adam	200	0.8596
40-60	10-6-3-2-1	Adam	300	0.8596

40-60	10-6-3-2-1	Adam	400	0.8596
40-60	10-6-3-2-1	SGD	100	0.8596
40-60	10-6-3-2-1	SGD	200	0.8596
40-60	10-6-3-2-1	SGD	300	0.8596
40-60	10-6-3-2-1	SGD	400	0.8596
40-60	6-3-2-1	Adam	100	0.8421
40-60	6-3-2-1	Adam	200	0.7895
40-60	6-3-2-1	Adam	300	0.807
40-60	6-3-2-1	Adam	400	0.7895
40-60	6-3-2-1	SGD	100	0.8596
40-60	6-3-2-1	SGD	200	0.8596
40-60	6-3-2-1	SGD	300	0.8596
40-60	6-3-2-1	SGD	400	0.8596
40-60	6-3--1	Adam	100	0.8421

40-60	6-3--1	Adam	200	0.8596
40-60	6-3--1	Adam	300	0.8596
40-60	6-3--1	Adam	400	0.8596
40-60	6-3--1	SGD	100	0.8596
40-60	6-3--1	SGD	200	0.8596
40-60	6-3--1	SGD	300	0.8596
40-60	6-3--1	SGD	400	0.8596

Best combination for highest accuracy

Test size   : 35 %

Architecture : 10-5-3-1

Optimizer   : Adam

Epochs   : 200
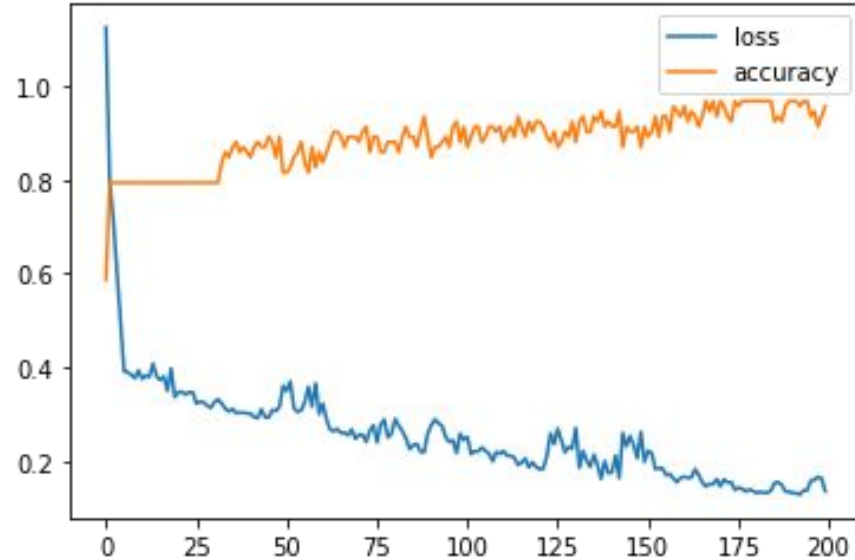
Accuracy   : 90 %

| 35-75 | 10-5-3-1 | Adam | 200 | 0.9 |

Same combination but the Optimizer is SGD

Test size   : 35%

Architecture : 10-5-3-1

Optimizer   : Adam

Epochs   : 200

Accuracy   : 80 %
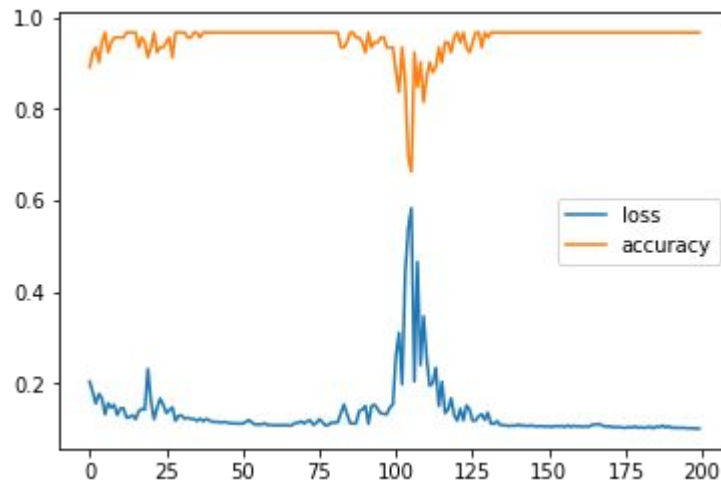
| 35-75 | 10-5-3-1 | SGD | 200 | 0.8 |

Worst Case Scenario

Test size   : 35%

Architecture : 10-5-3-1

Optimizer   : Adam

Epochs   : 300

Accuracy   : 70%
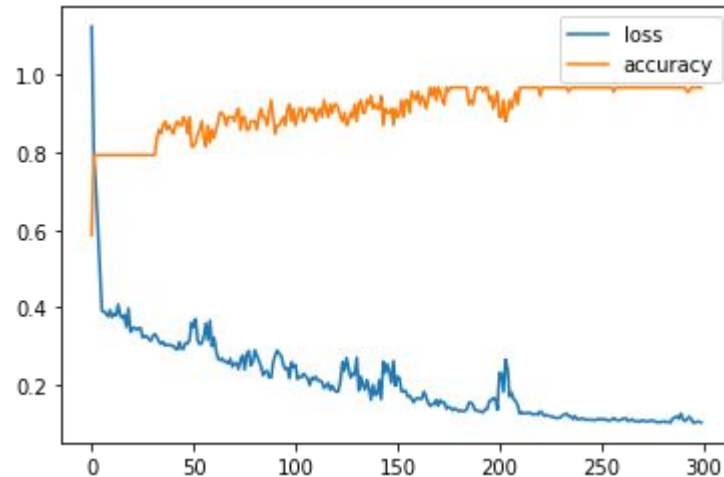
| 35-75 | 10-5-3-1 | Adam | 300 | 0.78 |

Conclusion

➢ The results form the table clearly states that the best optimizer out there is Adam. The reason it is best because at most of the times it gives higher accuracy when compared to SGD.

➢ So, it is suggested to use Adam optimizer over SGD

The best possible combination :

Test size   : 35 %

Architecture   : 10-5-3-1

optimizer   : Adam

Epochs   : 200

Accuracy   : 90 %