Deep Learning – Project: Part2

Submitters:
Netanel Hugi, 203553490.
Shimon Hagag, 311367536.

We built a model of logistic regression that classifies the presence of heart disease in the patient.

The model distinguishes between 2 options:
0 – Absence of heart disease.
1 – Presence of heart disease.

The data set we worked on contains files with patient details from 4 different hospitals, consisting of 75 features.
We used only one data file (of Cleveland), and the 13 following features:
1. Age: in years.
2. Sex: 0=female, 1=male.
3. Chest pain type:
   - Value 1: typical angina.
   - Value 2: atypical angina.
   - Value 3: non-anginal pain.
   - Value 4: asymptomatic.
4. Resting blood pressure: in mm Hg.
5. Serum cholesterol: in mg/dl.
6. Fasting blood sugar > 120 mg/dl: 1 = true, 0 = false.
7. Resting electrocardiographic results:
   - Value 0: normal.
   - Value 1: having ST-T wave abnormality.
   - Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria.
8. Maximum heart rate achieved.
9. Exercise induced angina: 1 = yes, 0 = no.
10. ST depression induced by exercise relative to rest.
11. The slope of the peak exercise ST segment:
   - Value 1: upsloping.
   - Value 2: flat.
   - Value 3: downsloping.
12. Number of major vessels: 0/1/2/3.
13. Thal: 3 = normal, 6 = fixed defect, 7 = reversable defect.

(Link to full description: https://archive.ics.uci.edu/ml/datasets/Heart+Disease)

The file of Cleveland database contains data of 303 different patients.
From this data we used 70% for train data, and the rest to testing our model.
The data was divided between test and train randomly.

Our best result is:

```
Alpah:  0.0001
Iterations:  100000

Total patients:  303
Train:  212
Test:  91

Results: 83 / 91
+-------------------+-------------------------+-------------------------+
|                   | Classified as Positive:39 | Classified as Negative:52 |
+===================+=========================+=========================+
| Really Positive:41 | True Positive: 36       | False Negative: 5       |
+-------------------+-------------------------+-------------------------+
| Really Negative:50 | False Positive: 3       | True Negative: 47       |
+-------------------+-------------------------+-------------------------+

Accuracy:  0.912
Recall:    0.878
Precision: 0.923
F-Measure: 0.900
```