Navid Al Nadvi, SID 912309925
Jason Zhang, SID 999230758
Team: Gender_Assumer

Question 1: Fully describe the network architecture and why it was chosen. How many input/hidden/output layer nodes? Fully connected or partially connected?

Answer: We have implemented a fully connected neural network that has 15,360 (128*120) neurons in the input layer, 2 hidden layers consisting 5 neurons, and one output neuron. We were able to achieve best accuracy using this architecture. As we train the network on the Male images the output value converges toward 0.5 and on the Female images the values converges towards 1. The decision plane is not well distinguished enough to be full classify the set. Some inputs the values are too close to be well classified. However, the trend is acknowledgeable (Male image inputs the outputs converging to 0.5 and on female image inputs the output values diverges from 1.

Questions 2: Cross Validations result

Amount of Male: 25 Amount of Female: 30
Amount of Male: 32 Amount of Female: 23
Amount of Male: 29 Amount of Female: 26
Amount of Male: 29 Amount of Female: 26
Amount of Male: 26 Amount of Female: 29

MEAN (Training)	MEAN (Testing)	Training STDEV	Test STDEV
87	82	7.3	3.1
91	82	1.064	6.695
90.8	86	5.638	4.172
87.4	86.8	0.708	5.508
74.8	88.2	0.547	5.263
82.6	85.8	1.095	3.240
92.2	88.6	4.182	4.712
88.6	83	1.055	4.635
87	82	1.140	4.561
86	85	1.140	4.321

Questions 3: Results on the test:

As it was explained the in Question 1 our decision plane is not separable enough to be fully classified. If you run the code you can notice the values for our Male inputs converges to our specified threshold value (0.5) and Female inputs diverges away.

General
Is Male | Testing File Output: 0.6277807235519992
Is Male | Testing File Output: 0.6337280957713897
Is Male | Testing File Output: 0.636308661565661
Is Male | Testing File Output: 0.6379836431687411
Is Male | Testing File Output: 0.639712983934013
Is Male | Testing File Output: 0.6408796190289471
Is Female | Testing File Output: 0.6409800390768084
Is Male | Testing File Output: 0.6409651395153672
Is Female | Testing File Output: 0.642226652950371
Is Female | Testing File Output: 0.643047417780537
Is Female | Testing File Output: 0.643592260729379
Is Male | Testing File Output: 0.6430674626409338
Is Female | Testing File Output: 0.6436648029176782
Is Male | Testing File Output: 0.643032229411756
Is Male | Testing File Output: 0.642684641227018
Is Male | Testing File Output: 0.6417178079061424
Is Male | Testing File Output: 0.641325904814433
Is Male | Testing File Output: 0.6406122319213274
Is Male | Testing File Output: 0.6404865238909568
Is Male | Testing File Output: 0.6412866919947879
Is Female | Testing File Output: 0.6418926267419243
Is Female | Testing File Output: 0.6421507347888108
Is Female | Testing File Output: 0.6431202253782787
Is Male | Testing File Output: 0.6429079780197613
Is Male | Testing File Output: 0.64279182413614
Is Female | Testing File Output: 0.6428732150737366
Is Male | Testing File Output: 0.641591034455728
Is Male | Testing File Output: 0.6408447748263851
Is Female | Testing File Output: 0.6415306233110105
Is Female | Testing File Output: 0.642015787649028
Is Male | Testing File Output: 0.6416620675193413
Is Male | Testing File Output: 0.641394314744263
Is Male | Testing File Output: 0.642397767150312
Is Female | Testing File Output: 0.6435842994624141

Is Female | Testing File Output: 0.6438940330056585
Is Male | Testing File Output: 0.6428558406475462
Is Male | Testing File Output: 0.6424855668032168
Is Female | Testing File Output: 0.6436637611865952
Is Female | Testing File Output: 0.6443261135792986
Is Male | Testing File Output: 0.643735777902627
Amount of Male: 34 Amount of Female: 16

Question 4:

Decision making process: Each neuron in the input layer gets multiplied by a random weight when going to each node in the hidden layer. Each hidden layer node sums up the incoming inputs multiplied by the random weight. Each node in the output layer calculate the sigmoid of the summation and feed the results of the sigmoid function forward. In the output layer the incoming sigmoid values from the previous layer is again summed and sigmoid function is calculated in the output layer node. The error is calculated against the expected result and error is propagated backwards to the input layer. Each training the derivative of the ERROR value in terms of weights of each corresponding nodes will be calculated and minimized at each run. The network calculates the gradient descent of the error to find the minima and better predict the label.