

ECS 170 - Programming Assignment 3

Aaron Allen / Sean Johnson

March 15, 2017

1. We use a feed-forward neural network with a single hidden layer and stochastic gradient descent. A feed-forward neural network was chosen due to its simplicity and ability to approximate any function. Our network is partially connected, due to the fact that our input layer does not directly connect to the output layer (only by proxy through the hidden layer), as is common with feed forward network designs. We chose to use only a single hidden layer to minimize potential overfitting that could occur with a more complex hidden layer design. We use 20 nodes to roughly approximate the number of key features differentiable through the network. Such features could include jawline, nose-size, hair, etc. Though somewhat arbitrary, we feel that a size of 20 would allow for key facial features to be represented without a high potential for overfitting as would occur with a large number of nodes that could pertain to set-specific anomalies.
2. We did not have time to complete this portion of the project.
3. See AASJ.predictions
4. Our network operates as follows. It takes each .txt file as input, and uses forward propagation to come up with a "hypothesis" for the sex of the subject. More specifically: the input layer is fed into the hidden layer (the values of the input layer are multiplied by the corresponding weights in the hidden layer, which is then stored in the hidden layer nodes), used to represent the features our neural network needs to differentiate between male and female faces. Our network comes up with a prediction by summing the results of the activation functions across all nodes in the hidden layer. Each of our 20 nodes roughly corresponds to some prevalent feature across the image set, including the background and specific facial features. The error of our prediction is then determined by finding the difference between predicted and actual results, which is used in back propagation to update the weights for the next round. Over the course of many rounds, our network will more correctly weight each of our predicted features of the image set.