When we see a person, we can instantly categorize them as being female or male. In this project you will create a neural network that will classify a portrait image as being either that of a male or female and then determine on what basis these decisions are being made.

Implement a feed forward neural network that uses SIGMOID units. The model should be learnt using a stochastic gradient descent back-propagation scheme using a momentum term (you may choose to implement weight decay if you wish). The neural network will need to perform both TRAIN and TEST functions. The network will learn to recognize the gender (male or female) of a facial image.

Use the training data set provided on the course web site. Each picture consists of 128 x 120 pixels with a range of intensities that represent a level of grayness.

**Evaluating Performance**

1. Fully describe the network architecture and why it was chosen. How many input/hidden/output layer nodes? Fully connected or partially connected? etc. Present your code. (15 points)
2. Randomly divide the training data into a five folds. Perform 5-fold cross-validation. Repeat this 10 times. List the mean and standard deviation of the training and test accuracies for each of the 10 experiments. Report your approximation of the learner bias, variance and MSE, explaining how you obtained your results. (25 points)
3. Try your neural network on the test set available on the course home page. This test set contains pictures similar to the training set, but also other pictures of non-graduate males and female students. For each instance, report in machine readable format, the prediction and confidence of your neural network. One prediction \{MALE, FEMALE\} per line and the confidence in the order given. Explain how you derived your confidence estimate (20 points).

**Understanding the Network Behavior**

4. For each hidden layer node try to determine what latent concept it represents. Explain your methodology and justify your conclusions by presenting empirical analysis. For example, you may wish to explore what parts of the face (lips, eyes, hair) are most used for predicting if a face is male or female. (20 points)

**Ensemble of Neural Networks**

5. Implement ADABoost or Bagging for neural networks. Repeat the experiments for question 2 be sure to use the same cross fold divisions. Did boosting or bagging improve performance? Did it reduce/increase the variance and/or bias of the learner for this problem? What additional correct predictions or errors were made when compared with question 2 (20 points).

For the interested student

Try using TANH units
Experiment with differing topologies with respect to:
- Number of hidden layer nodes and hidden layers.
- Fully versus partial connection of neurons.
- Ordering of the training instances. What changes effect the overall accuracy?