440 Final Project

Xiaowei Zhang, 204007085
Xinyuan Chen, 197001668

December 2021

Feature

For this project, we use the feature extraction by UCB.
Digit: Returns a set of pixel features indicating whether each pixel in the provided datum is white (0) or gray/black (1).
Face: Returns a set of pixel features indicating whether each pixel in the provided datum is an edge (1) or no edge (0).

Perceptron

We initialize the weight to be 0, and then update them later. We stop running when it reaches the max iterations. We have a list to keep the weights for each label and pick the label with the highest classifying $f(x_i, w) = w_0 + w_1\phi_1 + ...$ value as our guess. Update the weight if the guess doesn't equal to the exact answer. We train and test the model with $10\%, 20\%, ..., 100\%$ of the training data for both digits and faces.

- Accuracy plot using Perceptron for digits see Figure 1.
- Standard deviation plot using Perceptron for digits see Figure 2.
- Time plot using Perceptron for digits see Figure 3.
- Accuracy plot using Perceptron for faces see Figure 4.
- Standard deviation plot using Perceptron for faces see Figure 5.
- Time plot using Perceptron for faces see Figure 6.

Naive Bayes

We iterate over all the data and calculate the probability $P(Y = y) = \frac{numberofdatawithlabelY=yintraining}{totalnumberoftraining}$. We find the conditional probability for each label, feature, value. For classifying, we take the log of the probability for each label and the label with the max

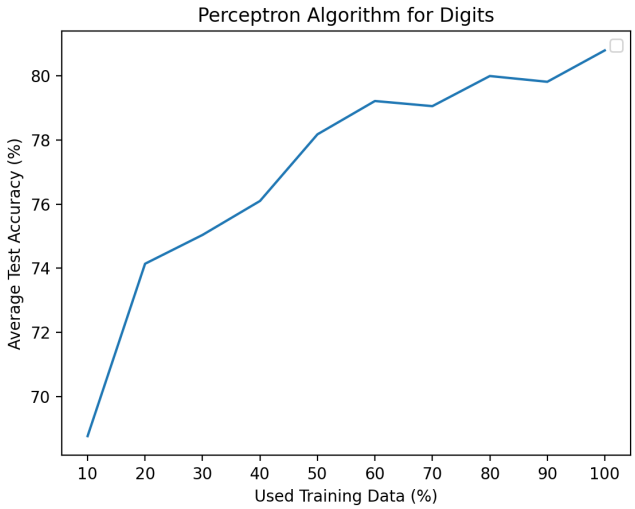

Figure 1: Average Accuracy of Perceptron for digits

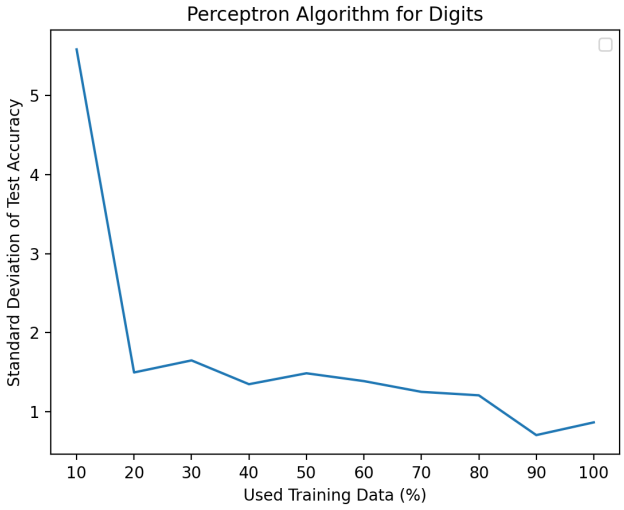

Figure 2: Standard Deviation of Accuracy of Perceptron for digits

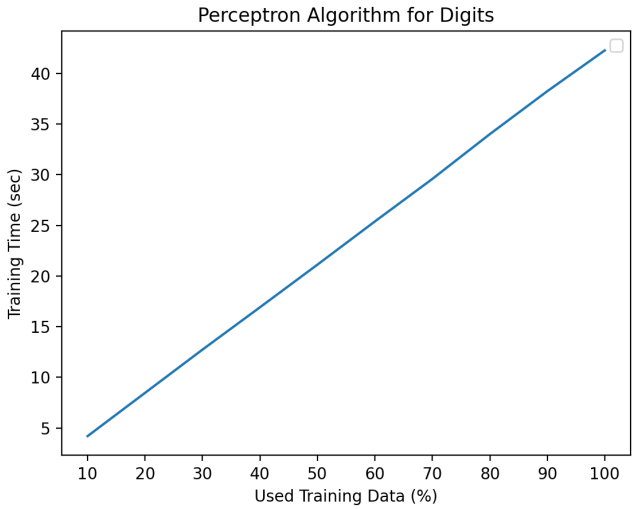

Figure 3: Time of Perceptron for digits

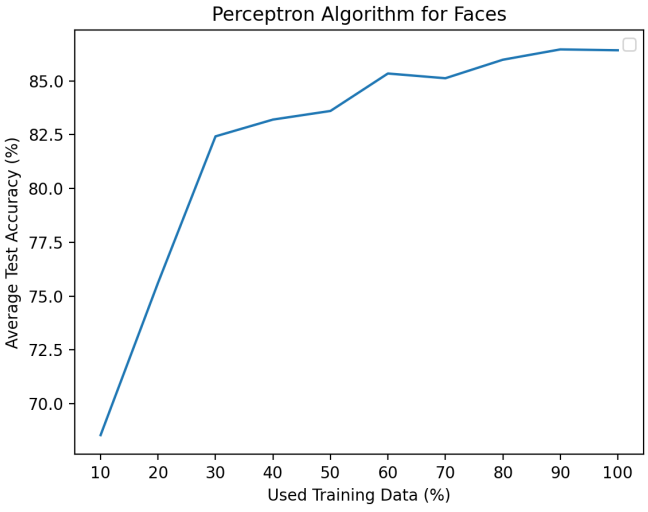

Figure 4: Average Accuracy of Perceptron for faces

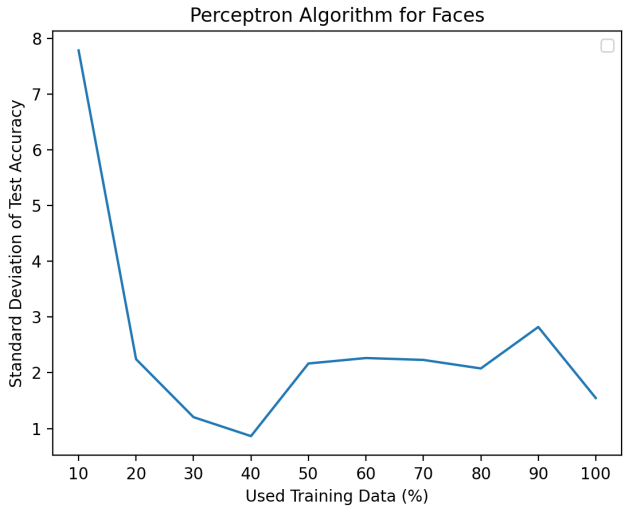

Figure 5: Standard Deviation of Accuracy of Perceptron for faces

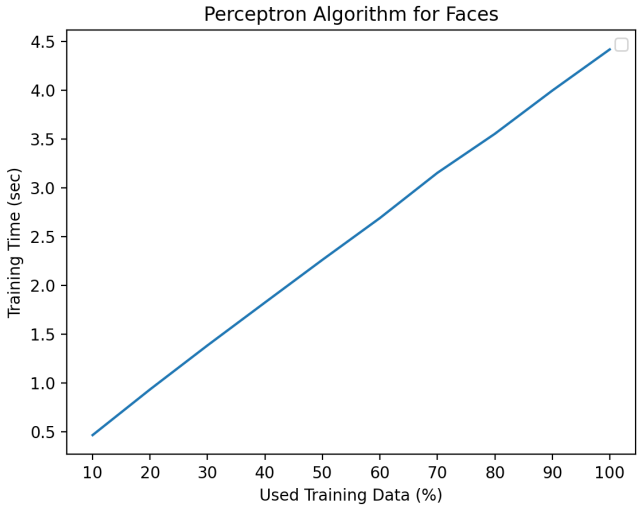

Figure 6: Time of Perceptron for faces

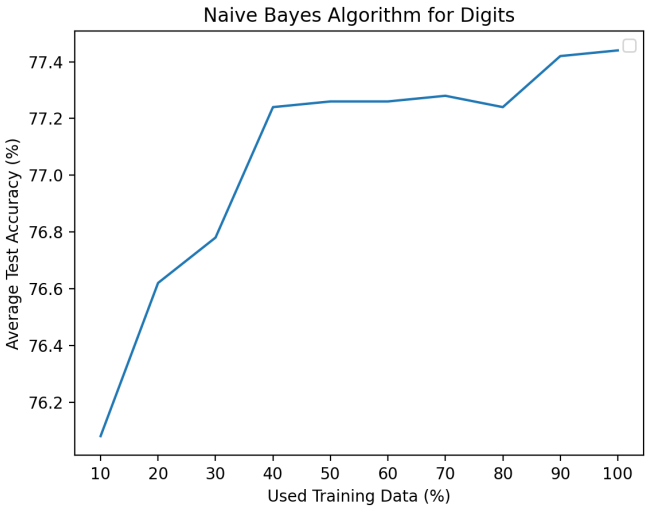

Figure 7: Average Accuracy of Naive Bayes for digits

probability is our guess. We train and test the model with $10\%, 20\%, ..., 100\%$ of the training data for both digits and faces.

- Accuracy plot using Naive Bayes for digits see Figure 7.
- Standard deviation plot using Naive Bayes for digits see Figure 8.
- Time plot using Naive Bayes for digits see Figure 9.
- Accuracy plot using Naive Bayes for faces see Figure 10.
- Standard deviation plot using Naive Bayes for faces see Figure 11.
- Time plot using Naive Bayes for faces see Figure 12.

Log Regression

The logistic regression treat the classification barrier as a linear classifier. Taking the input data, we first do a normalization on the data and compute the negative log probability for each label. By using stochastic gradient descent, we optimize the average cross-entropy loss batch by batch. For hyperparameter tuning, we try several batch sizes, learning rates and regularization weight and find the best ones. We train and test the model with $10\%, 20\%, ..., 100\%$ of the training data for both digits and faces.

- Accuracy plot using Log Regression for digits see Figure 13.

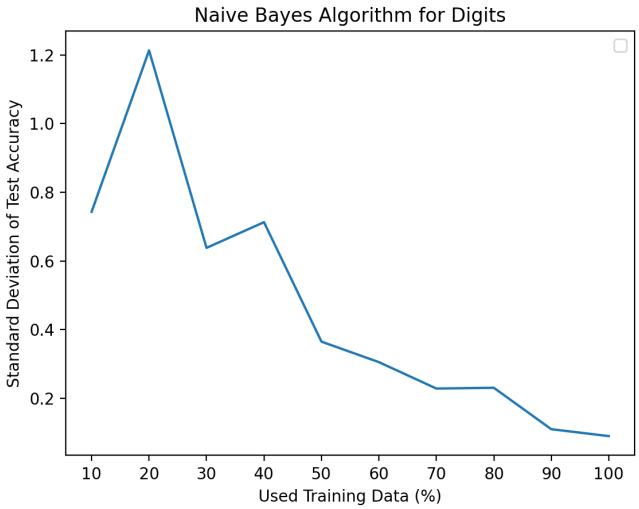

Figure 8: Standard Deviation of Accuracy of Naive Bayes for digits

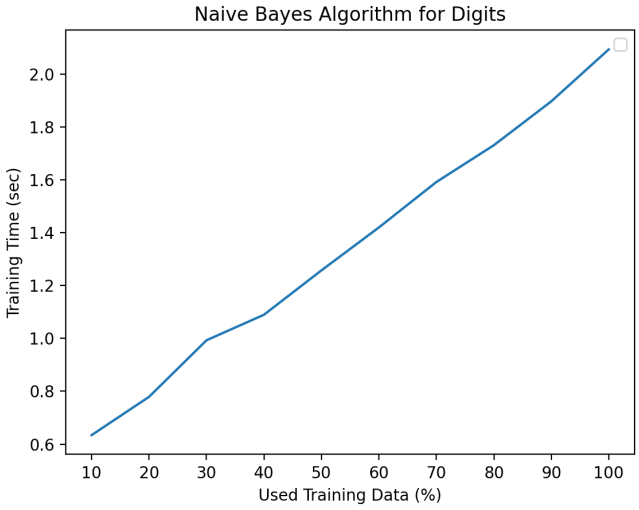

Figure 9: Time of Naive Bayes for digits

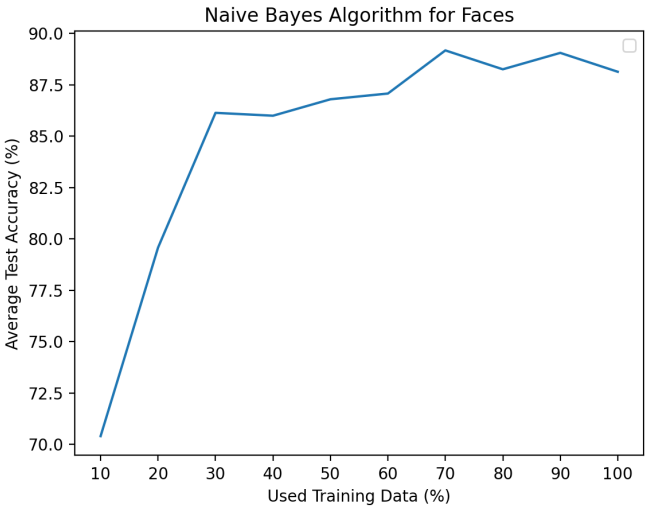

Figure 10: Average Accuracy of Naive Bayes for faces

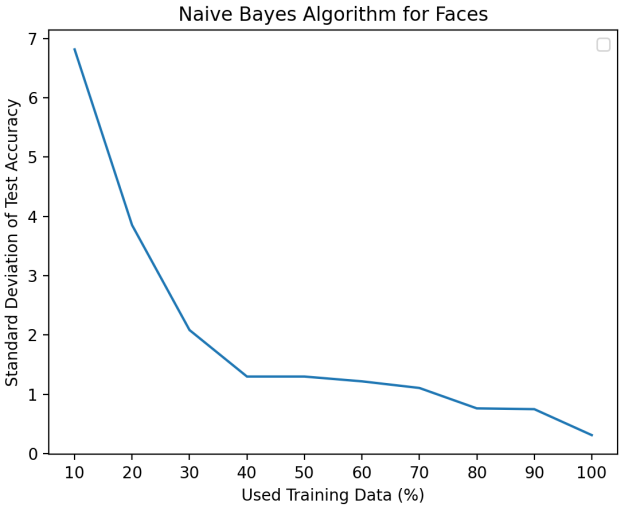

Figure 11: Standard Deviation of Accuracy of Naive Bayes for faces

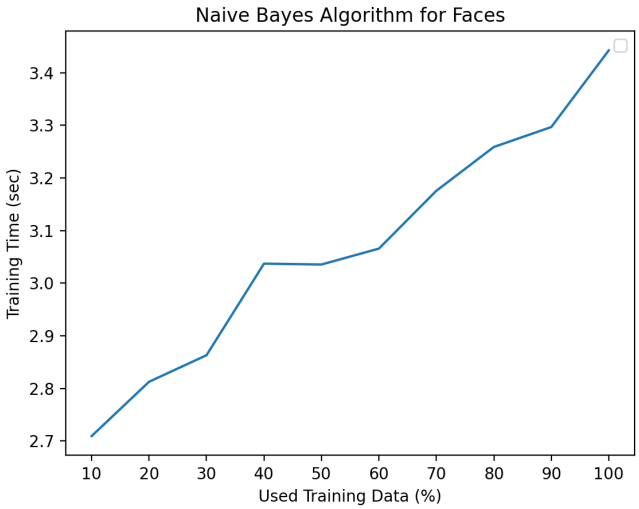

Figure 12: Time of Naive Bayes for faces

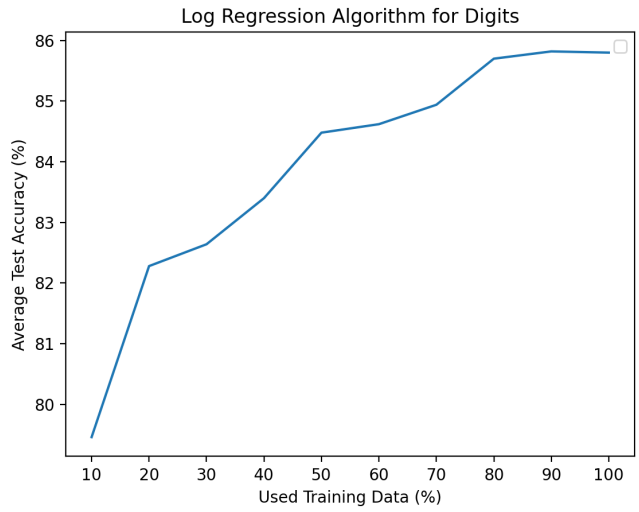

Figure 13: Average Accuracy of Log Regression for digits

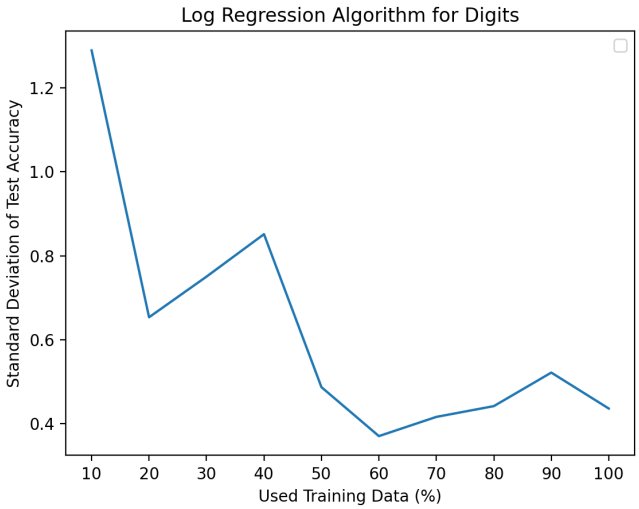

Figure 14: Standard Deviation of Accuracy of Log Regression for digits

- Standard deviation plot using Log Regression for digits see Figure 14.
- Time plot using Log Regression for digits see Figure 15.
- Accuracy plot using Log Regression for faces see Figure 16.
- Standard deviation plot using Log Regression for faces see Figure 17.
- Time plot using Log Regression for faces see Figure 18.

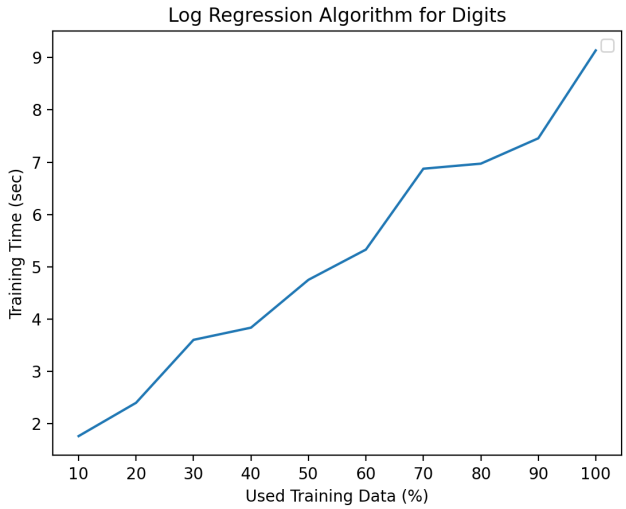

Figure 15: Time of Log Regression for digits

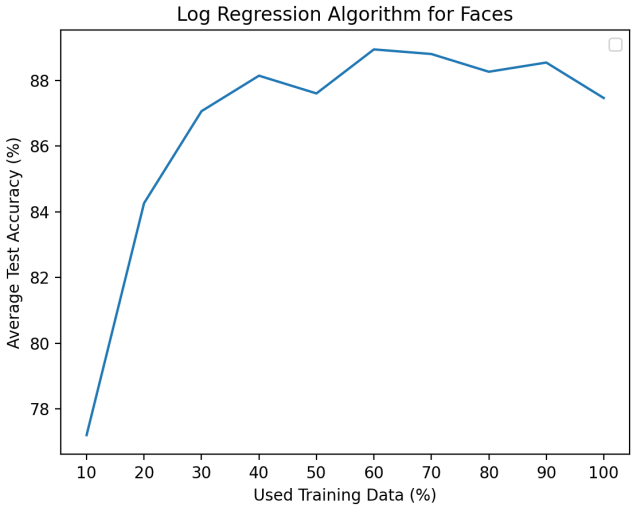

Figure 16: Average Accuracy of Log Regression for faces

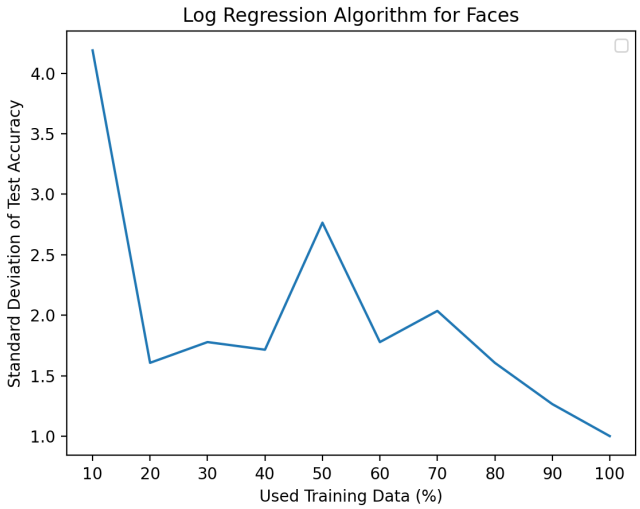

Figure 17: Standard Deviation of Accuracy of Log Regression for faces

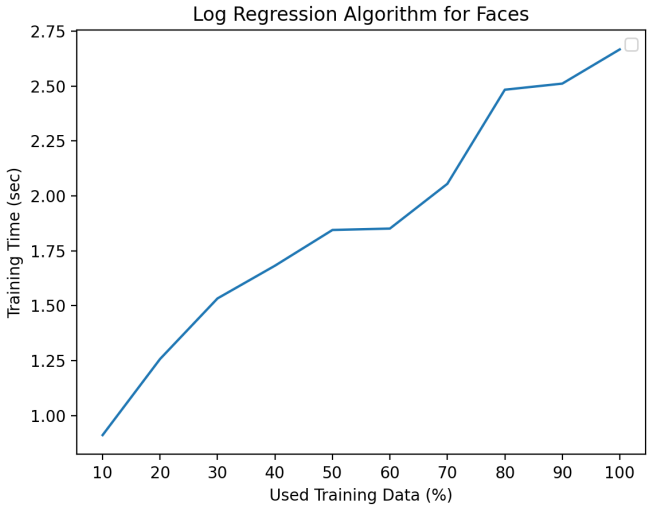

Figure 18: Time of Log Regression for faces