Project 2 — Face and Digit Classification

Fernando Gonzalez, Pranav Prakash, and Wanyun Liu

{*fdg17, pp618, wl432*}*@scarletmail.rutgers.edu*

August 9, 2020

1 Classification Algorithms

1.1 Naive Bayes

The Naive Bayes algorithm classifies images by keeping track of two sets of data. First are the *prior probabilities*, which are determined by:

$$\text{Prior}(y) = \Pr(Y = y) = \frac{\text{number of images with label = y}}{\text{total number of images}} \tag{1}$$

Our goal in using the Naive Bayes classifier is to compute the probability that a given image has a certain label, given that a set of features is observed. To compute these conditional probabilities, we introduce Bayes' Rule:

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)} \tag{2}$$

Here, A should refer to a *class*. The image classes are the set of all labels that can be given to an image. When classifying digits, these are the actual digits $\{0, 1, \ldots, 9\}$. For faces, the classes are "*not-face*" and "*face*", where "*not-face*" is represented by 0 or False, and "*face*" is represented by 1 or True. We let Y, y refer to classes, and X, x refer to features, where capitals are random variables. From (1), we have:

$$\Pr(Y = y \mid X_i = x_i) = \frac{\Pr(X_i = x_i \mid Y = y)\Pr(Y = y)}{\Pr(X_i = x_i)} \tag{3}$$

Observing from (1) that $\Pr(Y = y)$ is the prior probability $\text{Prior}(y)$, we have:

$$\Pr(Y = y \mid X_i = x_i) = \frac{\Pr(X_i = x_i \mid Y = y)\text{Prior}(y)}{\Pr(X_i = x_i)} \tag{4}$$

1.2 Perceptron

2 Implementation

2.1 Features

2.2 Results

2.2.1 Naive Bayes

2.2.2 Perceptron

2.3 Obstacles