Evaluation of Naive Bayes and Perceptron Classifiers for Digit and Face Recognition

Tomer Butbul 210002678

10/20/24

Naive Bayes Classifier

The Naive Bayes classifier is grounded in Bayes' theorem and assumes conditional independence among features given the class label. The posterior probability for a class C_k is calculated as:

$$P(C_k|X) \propto P(C_k) \prod_{i=1}^{n} P(x_i|C_k),$$

where $P(C_k)$ is the prior probability of class C_k and $P(x_i|C_k)$ is the likelihood of feature x_i given the class.

The implementation involved calculating priors and likelihoods for each class using the training data. Laplace smoothing was applied to handle zero probabilities. During prediction, the posterior probability for each class was computed, and the class with the highest probability was assigned. NumPy was used for efficient vectorized operations, ensuring scalability to large datasets.

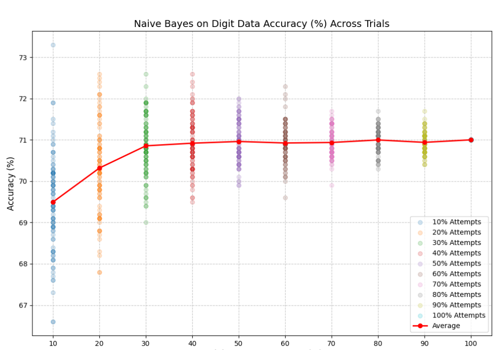

Figure 1

Figure 1 demonstrates how accuracy increases steadily with training data percentage for digit recognition, while Figure 2 shows the linear growth in training time.

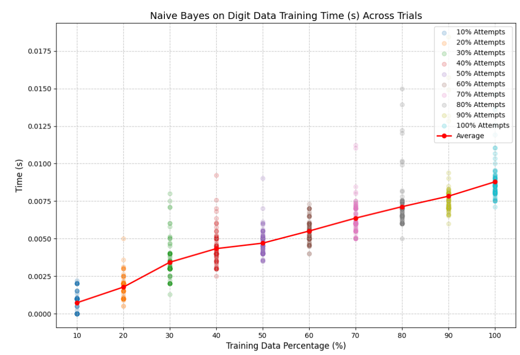

Figure 2

Perceptron Classifier

The perceptron classifier is a linear model that iteratively updates weights and biases based on misclassified examples. The decision rule is:

$$y = \text{sign}(W \cdot X + b),$$

where W represents the weight vector, X is the input features, and b is the bias term.

When a misclassification occurs, the weights and biases are updated as:

$$W \leftarrow W + \eta y_i X_i, \quad b \leftarrow b + \eta y_i,$$

where η is the learning rate. The implementation was designed with adjustable epochs and learning rates to balance convergence speed and accuracy.

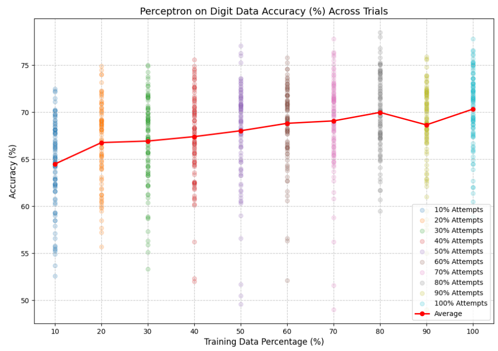

Figure 3

The classifier iteratively improved its decision boundary with each pass over the training data. Figure 3 highlights the perceptron's superior performance in accuracy compared to Naive Bayes, while Figure 4 emphasizes the computational cost.

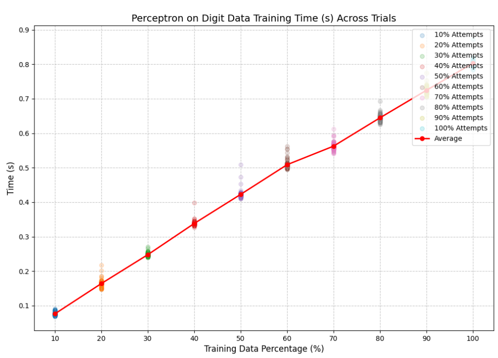

Figure 4

Feature Extraction

For digit recognition, 28×28 images were flattened into binary pixel values, with # pixels assigned higher importance due to their role in defining the digit's structure. For face recognition, 70×60 images were divided into grids, and each grid cell was treated as a binary feature indicating whether it contained markings. This approach preserved spatial relationships and reduced dimensionality.

The grid-based method improved face classification accuracy, as seen in Figures 5 and 6.

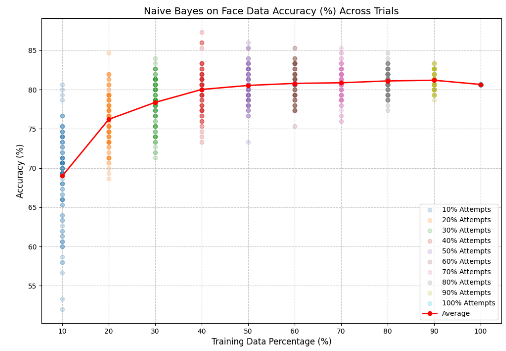

Figure 5

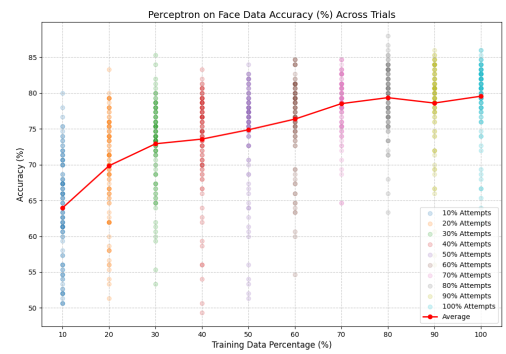

Figure 6

Comparison of Classifiers

The Naive Bayes classifier demonstrated consistent performance with minimal computational overhead, making it suitable for quick evaluations and datasets with simpler decision boundaries. However, its accuracy plateaued as training data increased, highlighting its limitations in capturing complex patterns. In contrast, the Perceptron consistently outperformed Naive Bayes in accuracy for both digit and face datasets. With smaller training percentages, the Perceptron demonstrated significant improvements over Naive Bayes. As training percentages increased, the Perceptron maintained its lead, achieving peak accuracies of 76% for digits and 85% for faces.

The trade-off came in computational cost. Figures 2 and 4 highlight the faster training times of Naive Bayes compared to the Perceptron. This trade-off makes Naive Bayes suitable for applications requiring quick training, while the Perceptron is better suited for scenarios where accuracy is paramount.

Results

The results for accuracy and training time were evaluated across multiple training data percentages, ranging from 10% to 100%, with 100 trials per percentage. The average results were computed and plotted alongside all individual trials.

The spread of accuracy results across trials decreased with larger training percentages, reflecting an increase in precision. This trend highlights the importance of robust training datasets for generalizing effectively to unseen data.

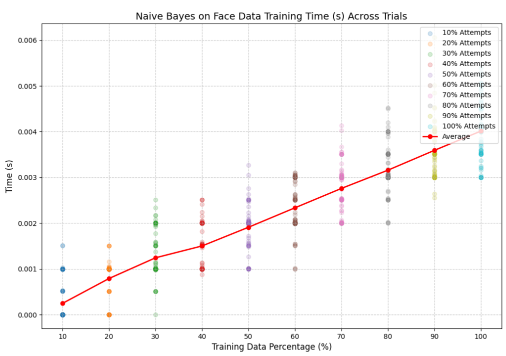

Figure 7

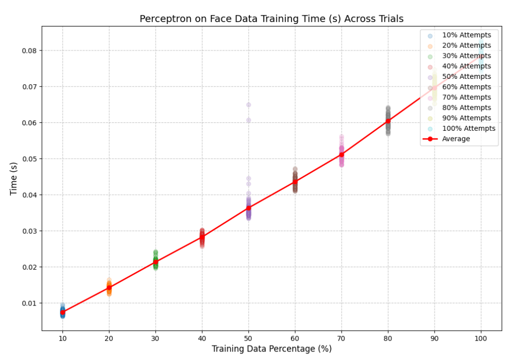

Figure 8