# Practical Part

3.1. Boosting

1. The plot shows that as the number of learners in the AdaBoost ensemble increases, both training and test errors decrease rapidly at first and then stabilize. The test error is slightly higher than the training error, indicating effective learning and good generalization without significant overfitting.


Training and Test Errors vs Number of Learners (Noise=0)
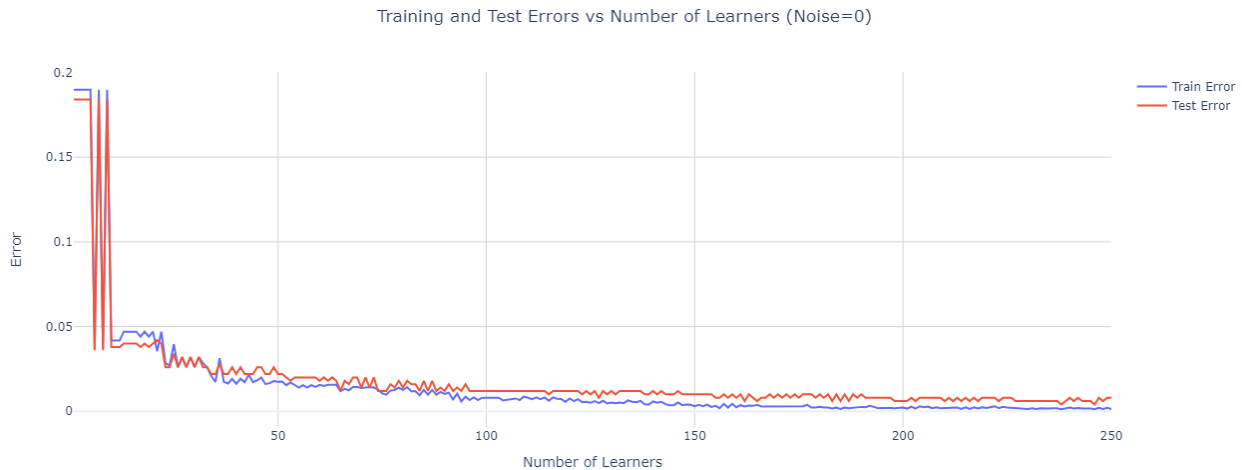
2. The dataset features a blue class in a circular shape surrounded by red points, making it inherently non-linearly separable. With just 5 decision stumps, the ensemble begins to form a basic decision boundary. As the ensemble grows to 50, 100, and 250 decision stumps, the boundary becomes increasingly smooth and accurately delineates the circular blue class. This progression highlights the AdaBoost ensemble's effectiveness in combining weak learners to enhance the model's ability to capture complex class shapes.

Decision Boundaries at Different Iterations (Noise=0)
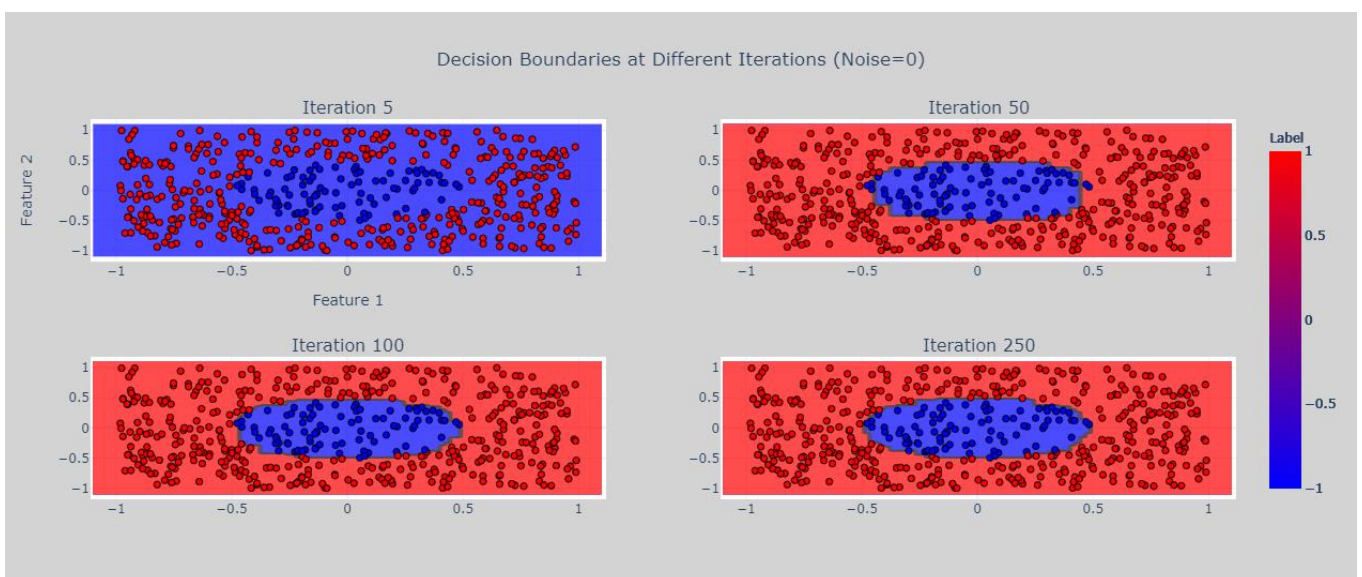
3. The ensemble size that achieved the lowest test error is 238, with an accuracy of 1.00. The decision surface plot for this ensemble size shows a clear boundary that perfectly separates the blue circular class from the surrounding red points, effectively classifying all test set data points correctly. This result highlights the optimal performance of the AdaBoost ensemble with 238 learners in this scenario.
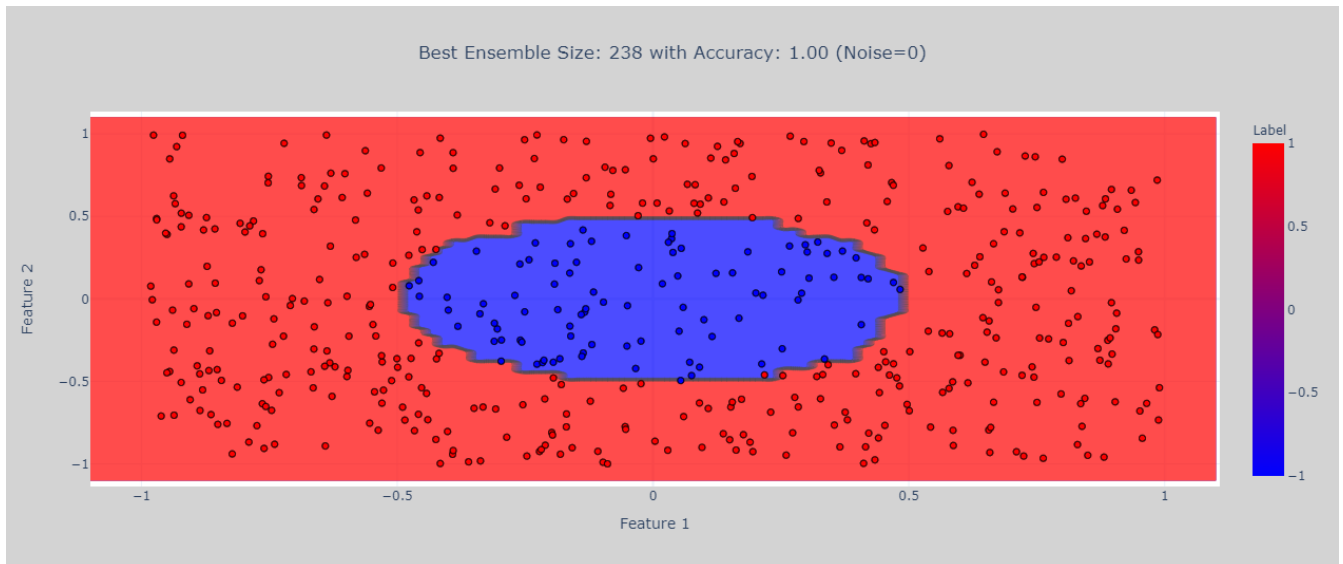
4. The plot shows training samples with point sizes proportional to their weights from the final iteration of the AdaBoost ensemble. The decision surface separates the blue circular class from the red outer region. Samples deep within these regions have small weights, indicating they were easy to classify. The most challenging samples, located near the boundary of the blue circle, have larger weights due to their proximity to the decision boundary. This demonstrates how AdaBoost focuses on difficult samples to improve overall classifier performance.
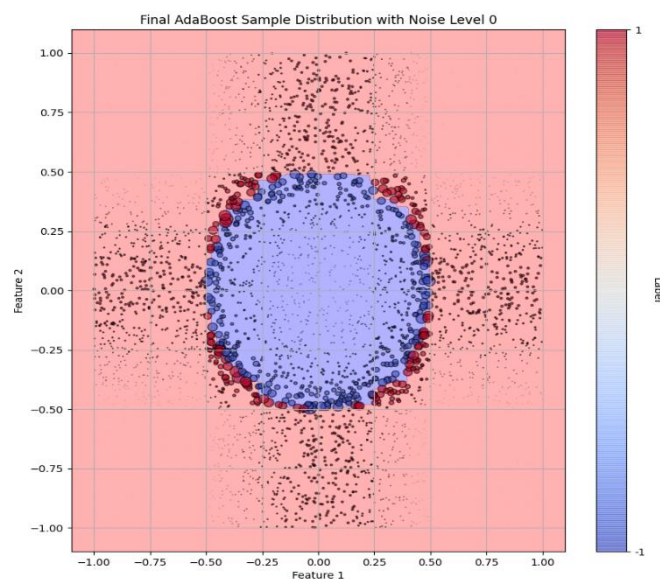
5. The graph shows the training and test errors with a noise level of 0.4, compared to the previous graph with no noise. With a noise level of 0.4, both training and test errors are higher and more fluctuating compared to the no-noise scenario. The test error is generally higher than the training error, indicating overfitting. This reflects the bias-variance tradeoff, where noise increases variance, making it harder for the model to generalize effectively.
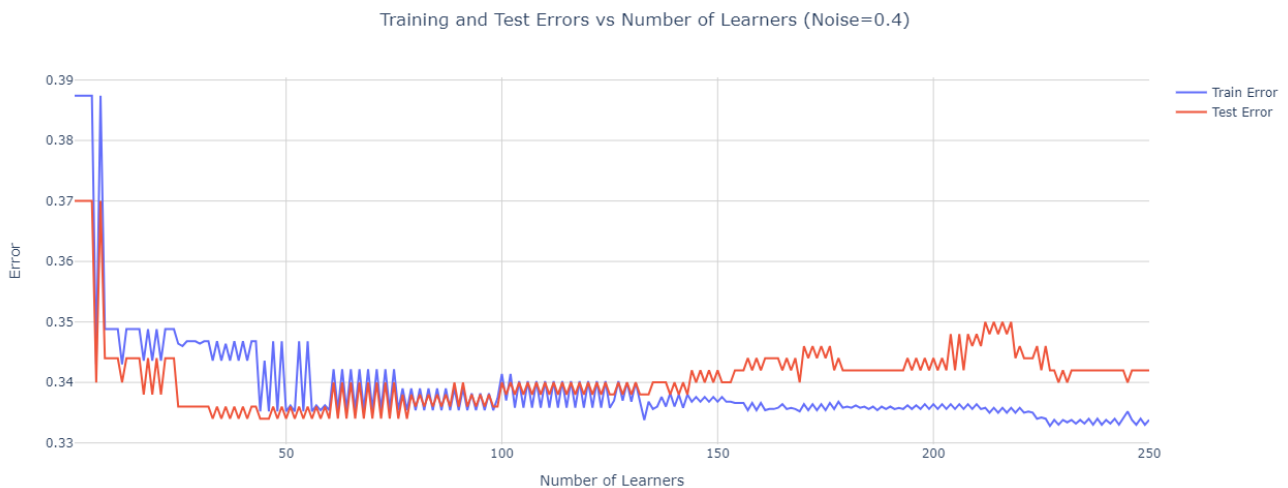
With a noise level of 0.4, the plot shows larger weights for more samples, especially near the decision boundary and within the regions. The increased noise leads to more misclassifications and a less accurate decision boundary, highlighting the bias-variance tradeoff. The model's variance increases, causing overfitting to the noisy data and reducing its generalization performance.