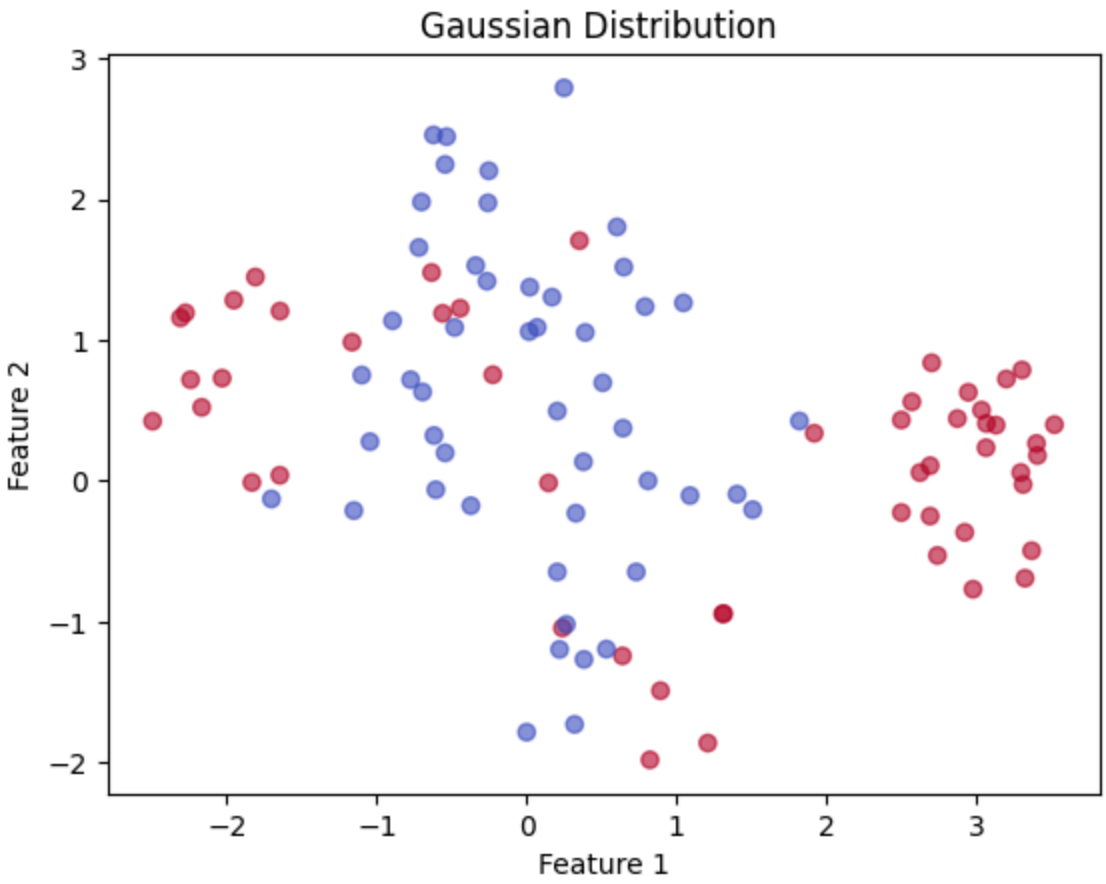
Guna Vengalasetti

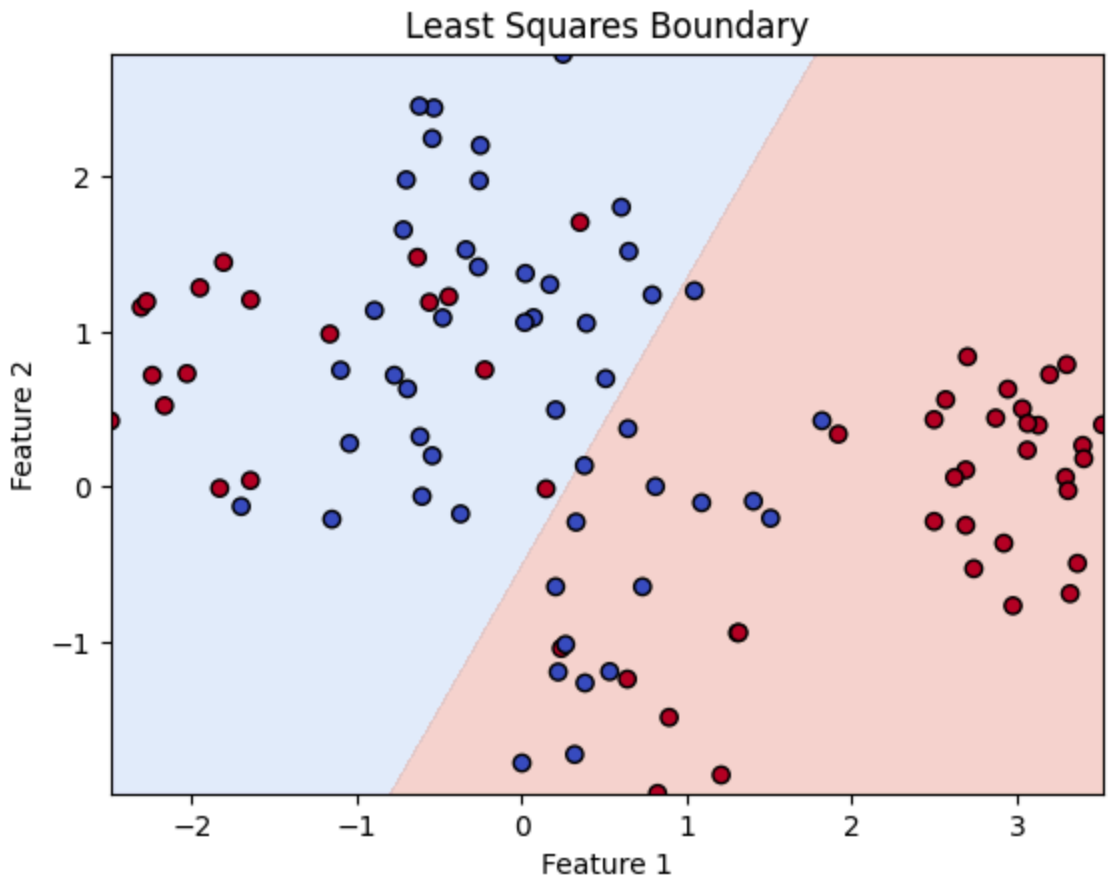
CMPE 255 Section 02

September 22 2024

This process shows how a linear regression model can classify data and measure the accuracy of it. First, two distinct clusters of data points are generated, each representing a different class. The linear regression model is trained to separate these classes by finding a boundary between them. The grid of values shows this boundary on a plot, showing how the model divides the data into two groups.

Next, the model is tested on the original dataset to see how well it predicts class labels. The false positive rate measures how often the model incorrectly classifies points from class 0 as class 1, and the false negative rate shows how often class 1 points are wrongly labeled as class 0. These rates are important because they reveal where the model makes mistakes, helping us understand its strengths and weaknesses. Overall, this process helps evaluate how well a machine learning model can classify data into different categories.


Gaussian Distribution
Feature 2
Feature 1
3
2
1
0
−1
−2
−2
−1
0
1
2
3



False Positive Rate: 34.69%

False Negative Rate: 35.29%

The False Positive Rate (FPR) is 34.69%, meaning 34.69% of class 0 points were wrongly classified as class 1. It's calculated as:

FPR= (False Positives)/(False Positives+True Negatives) ×100

The False Negative Rate (FNR) is 35.29%, meaning 35.29% of class 1 points were incorrectly classified as class 0. It's calculated as:

FNR=False Negatives/(False Negatives+True Positives) ×100

These rates show that the model is making many mistakes and needs improvement to better classify the data.