Exploring Classification using Naive Bayes

Approach:

For the Exploring classification using Naïve Bayes, firstly for a given class labels we had to find the unique class labels and also count of labels, then we calculated the prior probability over each class. For each unique label, we found the estimated parameters for each distribution i.e., Gaussian(μ, σ^2), Bernoulli(p), Laplace(μ,b), Exponential(λ), and Multinomial Distributions. Then we divided the 10 features into 2 for each distribution and calculated the LL(log-likelihood) for each distribution for 2 features over the labels and priors, after calculating the individual distribution LL we found the total log-likelihood and then argmax over the posterior probability to get the predictions.

Estimated Parameters:

1. Priors

Class label	Prior
0	0.33
1	0.33
2	0.33

2. Gaussian Distribution :

Class label	Mean(μ_x1, μ_x2)	Variance(σ_x1,σ_x2)
0	[2.02094922, 3.90677339]	[9.0519516 , 78.42834902]
1	[0.02138694, 0.85591792]	[25.16089116, 230.03185742]
2	[8.02485039, -0.02166141]	[35.66886557, 4.00754373]

3. Bernoulli Distribution:

Class label	Probability of success(p_x3,p_x4)
0	[0.2023, 0.104]
1	[0.5984, 0.8018]
2	[0.9053, 0.1947]

4. Laplace Distribution:

Class label	Mean(μ_x5,μ_x6)	(b_x5,b_x6)
0	[0.06056981, 0.86609291]	[2.81382881, 8.43575503]
1	[0.38182717, 0.29061336]	[1.42005384, 8.43941067]
2	[0.7481149 , 0.20976785]	[4.23410835, 4.27751756]

5. Exponential Distribution

Class label	Lambda(λ_7,λ_8)
0	[1.97829982, 3.93542478]
1	[2.98410954, 7.98003545]
2	[0.31325933, 0.68672259]

6. Multinomial Distribution

Class label	Mean(μ)
0	[0.36020046, 0.63978132]
1	[0.33944863, 0.66053478]
2	[0.31325933, 0.68672259]

Accuracy and F1 score:

Training Accuracy : 0.84276 = 84.27%

Validation Accuracy: 0.84396 = 84.39%

Training F1 Score: [0.8260065192001065, 0.8362091299567933, 0.8709372742436247]

Validation F1 Score: [0.828240066978956, 0.8330647257104032, 0.8752764263105668]

Result:

Calling predict : [9.0519516 78.42834902]

Accuracy on test dataset : 0.8439666 = 84.39%