CS725 - Foundations Of Machine Learning

Homework - 3

Debdoot Roy Chowdhury 23M0765
Pranab Kumar Paul 23M0800

Contents

1 Approach

In this Homework, we have implemented a Naive Bayes Classifier on a toy dataset for the classification task. We have implemented the Naive Bayes algorithm in Python with libraries: Panda, Numpy, and Matplotlib. The approach we have taken to build the model for the classification task according to the Naive Bayes algorithm is described in the following.

The dataset that is given to us has 10 features labeled as $(X_1, X_2, ..., X_{10})$ and one output label Y. The output label Y has three labels which are 0,1 and 2. The approach for building this model is given below:

- Parameters are estimated for each distribution using maximum likelihood estimation.
- For each distribution the highest probability class is chosen as the predicted label for each datapoint.
- The output label that occurs maximum among the predicted labels of all the distribution is chosen as the final predicted output label.

2 Parameters

The parameters for each of the class labels with respect to various distributions are listed below:

1. **Gaussian:** The parameters for each of the class labels w.r.t Gaussian distribution is shown below.

$ClassLabels(Y)$	$MeanX_1$	$MeanX_2$	$VarX_1$	$VarX_2$
0	2.02	3.90	9.05	78.42
1	0.02	0.85	25.16	230.03
2	8.02	-0.02	35.67	4.00

Table 1: $MeanX_i$ means the mean parameter (μ) for the column X_i and similarly $VarX_i$ means variance parameter (σ^2) for the column X_i.

2. **Bernoulli:** The parameters for each of the class labels w.r.t Bernoulli Distribution is shown below.

$ClassLabels(Y)$	PX_3	PX_4
0	0.20	0.10
1	0.59	0.80
2	0.90	0.19

Table 2: PX_i means the Bernoulli parameter (p) for column X_i.

3. **Laplace:** The parameters for each of the class labels w.r.t Laplace Distribution is shown below.

$ClassLabels(Y)$	$MeanX_5$	$MeanX_6$	$ScaleX_5$	$ScaleX_6$
0	0.06	0.87	1.98	5.98
1	0.38	0.29	0.99	5.99
2	0.75	0.21	3.00	3.06

Table 3: $MeanX_i$ means the mean parameter (μ) for the column X_i and similarly $ScaleX_i$ means scale parameter (b) for the column X_i.

4. **Exponential:** The parameters for each of the class labels w.r.t Exponential Distribution is shown below.

$ClassLabels(Y)$	$LambdaX_1$	$LambdaX_2$
0	1.98	3.93
1	2.98	7.98
2	8.94	14.68

Table 4: $LambdaX_i$ is the parameter (λ) for the column X_i.

5. **Multinomial:** The parameters for each of the class labels w.r.t Multinomial Distribution is shown below.

$ClassLabels(Y)$	X_9					X_10							
	P0	P1	P2	P3	P4	P0	P1	P2	P3	P4	P5	P6	P7
0	0.20	0.20	0.20	0.19	0.19	0.121	0.123	0.125	0.127	0.127	0.127	0.124	0.123
1	0.09	0.19	0.40	0.15	0.14	0.100	0.050	0.050	0.199	0.152	0.148	0.200	0.096
2	0.20	0.29	0.10	0.34	0.05	0.197	0.048	0.048	0.105	0.155	0.153	0.098	0.195

Table 5: P_i is the multinomial parameter for value i in column X_i.

3 Result

1. **Training Accuracy:** 0.8071333333333334
2. **Validation Accuracy:** 0.8073

4 F1-Score

1. **Training F1 Score:** [0.7781034225820874, 0.7691524560956986, 0.8729232594936708]
2. **Validation F1 Score:** [0.7779874838581504, 0.7701085018593041, 0.8725475660983445]