FML : Assignment 03 Report

Department of Computer Science

Indian Institute of Technology Bombay

Guided by

Prof. Sunita Sarawagi

Foundation of Machine Learning

CS 725

Prepared By

Bharat Patidar
23M0761

Rohit Singh Yadav
23M0773

Report :

In this Exercise/Assignment we have learnt classification on toy datasets with Naive Bayes Method. There is one datasets which is a 3-class classification having 10 features (X1,X2,.......,X10) .

We have implemented these functions fit( ), getParams( ), predict( ), f1score( ), recall( ) and precision( ) functions.

Each datapoint consists of 10 dimension, which we label as (X1, X2, .., X10) and class label from (0, 1, 2).

1. (X1, X2) are drawn independently from two different univariate G**aussian distributions**.

2. (X3, X4) are random variables drawn independently from two different B**ernoulli Distributions**

3. (X5, X6) are random variables drawn independently from two different **Laplace Distributions**

4. (X7, X8) are random variables drawn independently from two different **Exponential Distribution**

5. (X9, X10) are random variables drawn independently from two different **Multinomial Distributions**

our goal was to calculate Maximum Likelihood estimators for each of these distributions and create a naive Bayes classifier with an appropriate prior to classify these points.

1. Gaussian : (μ, σ2)

2. Bernoulli : p

3. Laplace : (μ, b)

4. Exponential : λ

5. Multinomial : [p1, p2, .., pk]

Our Approach For this Problem :

`fit(self, X, y)`: This method is used for training the Naive Bayes classifier. It calculates the prior probabilities and distribution parameters (mean, variance, etc.) for each class based on the input training data `X` and corresponding labels `y`. Here's a breakdown of what it does for different distribution types:

Gaussian Distribution: It calculates the mean and variance for each feature in each class.

Bernoulli Distribution: It calculates probabilities for binary features.

Laplace Distribution: It calculates parameters for the Laplace distribution.

Exponential Distribution: It calculates the rate parameter (lambda) for exponential distribution.

Multinomial Distribution: It calculates probabilities for discrete features.

`predict(self, X)`: This method is used to make predictions on new data `X` based on the trained model. It calculates the posterior probabilities for each class and selects the class with the highest posterior probability as the predicted class.

`getParams(self)`: This method returns the calculated priors and parameters for all classes in the form of dictionaries. It provides the prior probabilities and distribution parameters for each class, which can be used for evaluation.

These functions are used for model evaluation.

`net_f1score(predictions, true_labels)`: Calculates the F1 score for each class and returns a list of F1 scores. It uses precision and recall to compute the F1 score.

`accuracy(predictions, true_labels)`: Calculates the accuracy of the model by comparing predicted labels with true labels.

Accuracy of Model :

Training Accuracy: 0.8417333333333333

Validation Accuracy: 0.8402666666666667

F1 Scores :

CLASS	Training F1 Score	Validation F1 Score
Y=0	0.82	0.82
Y=1	0.78	0.78
Y=2	0.90	0.90

More precise F1 Scores Vectors :

Training F1 Score: [0.8282335950898095, 0.7849141570503799, 0.9069084026497982]

Validation F1 Score: [0.8290071528183903, 0.7800147450802473, 0.9054523785877445]

Gaussian distributions :

Parameters: Mean (μ) and Variance (σ^2)

The Gaussian distribution is characterized by its mean (μ), which represents the central location of the distribution, and its variance (σ^2), which measures the spread or dispersion of the data.

Field X1 :

CLASS	Mean (μ)	Varience(σ2)
Y=0	2.02094922	9.0519516
Y=1	0.02138694	2.51608912
Y=2	8.02485039	3.5668865

Field X2 :

CLASS	Mean (μ)	Varience(σ2)
Y=0	3.90677339	7.84283490
Y=1	0.85591792	2.3003185
Y=2	-0.02166141	4.00754373

Bernoulli Distributions :

Parameter: p (probability of success)

The Bernoulli distribution models the probability of a binary outcome, typically denoted as success (1) or failure (0). The parameter 'p' represents the probability of success.

CLASS	P(0)	P(1)
Y=0	0.2023	0.104
Y=1	0.5984	0.8018
Y=2	0.9053	0.1947

Laplace Distributions :

Parameters: Mean (μ) and Scale parameter (b)

The Laplace distribution is characterized by its mean (μ), which represents the central location of the distribution, and the scale parameter (b), which controls the spread or width of the distribution. It is often used to model data with heavy tails.

CLASS	Mean(μ)	Scale Parameter(b)
Y=0	0.06056	0.8660
Y=1	0.381827	0.290613
Y=2	0.74811	0.20976

Exponential Distributions :

Parameter: Rate parameter (λ)

The exponential distribution models the time between events in a Poisson process. The rate parameter (λ) represents the average number of events per unit of time, and it is also the reciprocal of the mean ($\lambda = 1/\mu$).

CLASS	Mean(λ)
Y=0	1.97829981
Y=1	2.9841095
Y=2	8.94272

Multinomial Distributions :

Parameters: Probabilities for k outcomes ($p_1, p_2, ..., p_k$)

The multinomial distribution generalizes the binomial distribution to more than two categories or outcomes. It is characterized by a set of probabilities ($p_1, p_2, ..., p_k$), where each p_i represents the probability of outcome i out of k possible outcomes.

CLASS	P1	P2
Y=0	1.9765, -0.97649	3.5107, -2.510
Y=1	2.0463, -1.046	3.982, -2.98
Y=2	1.7326, -0.7325	3.7983, -2.798