

EXPERIMENT 8

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose **NaiveBayes**

Test options

- Use training set
- Supplied test set Set...
- Cross-validation Folds 10
- Percentage split % 66
- [More options...](#)

(Nom) class

Start Stop

Result list (right-click for options)

08:46:00 - bayes.NaiveBayes

```
==== Run information ====
Scheme: weka.classifiers.bayes.NaiveBayes
Relation: iris
Instances: 150
Attributes: 5
sepallength
sepalwidth
petallength
petalwidth
class
Test mode: 10-fold cross-validation
==== Classifier model (full training set) ====
Naive Bayes Classifier

Class
Attribute Iris-setosa Iris-versicolor Iris-virginica
mean (0.33) (0.33) (0.33)
=====
sepallength
mean 4.9913 5.9379 6.5795
std. dev. 0.355 0.5042 0.6353
weight sum 50 50 50
precision 0.1059 0.1059 0.1059

sepalwidth
mean 3.4015 2.7687 2.9629
std. dev. 0.3925 0.3038 0.3088
weight sum 50 50 50
precision 0.1091 0.1091 0.1091

petallength
mean 1.4694 2.4252 5.5516
std. dev. 0.1782 0.4712 0.5529
weight sum 50 50 50
precision 0.1405 0.1405 0.1405

petalwidth
mean 0.2743 1.3097 2.0343
std. dev. 0.1096 0.1915 0.2646
weight sum 50 50 50
precision 0.1143 0.1143 0.1143

Time taken to build model: 0.01 seconds
==== Stratified cross-validation ====
==== Summary ====
Correctly Classified Instances 144 96 %
Incorrectly Classified Instances 6 4 %
Kappa statistic 0.94
Mean absolute error 0.0342
Root mean squared error 0.155
Relative absolute error 7.6997 %
Root relative squared error 32.8754 %
Total Number of Instances 150
==== Detailed Accuracy By Class ====

```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	Iris-setosa
0.960	0.040	0.923	0.960	0.941	0.911	0.992	0.983	0.983	Iris-versicolor
0.920	0.020	0.958	0.920	0.939	0.910	0.992	0.986	0.986	Iris-virginica
Weighted Avg.	0.960	0.020	0.960	0.960	0.940	0.994	0.989	0.989	

```
==== Confusion Matrix ====
a b c <-- classified as
50 0 0 | a = Iris-setosa
0 49 2 | b = Iris-versicolor
0 4 46 | c = Iris-virginica
```

Status OK

Log x0

Student Name: Roll No.:

Experiment No.: Date: 7

Experiment:-8

- Objective:- Naive Bayes algorithm implementation on continuous dataset (IRIS dataset) using WEKA explorer.

$$\text{Impdf} = \frac{1}{\sqrt{2\pi r^2}} \exp \left(-\frac{1}{2} \left(\frac{z-u}{r} \right)^2 \right).$$

- Theory :- Naive Bayes algorithm is a probabilistic classifier based on Bayes theorem assuming independence b/w features.

UIET
UNIVERSE

- Steps:-

- ① Open weka and go to explorer panel.
- ② Load dataset (IRIS dataset)
- ③ Click on classify Tab.
- ④ Choose Naive Bayes from bayes in classifier panel.
- ⑤ Set test option.
- ⑥ Click the Start Button.

- Results:- Naive Bayes classifier in weka successfully classifier, IRIS dataset by predicting species based on feature probabilities & class wise statistics.

Student Name: Roll No.

Experiment No.: Date:

#

Viva-questions-

v1

Explain the significance of Gaussian probability density function (GPDF) for NR algo. application to continuous dataset.

→

GPDF is significant for Naive Bayes in continuous dataset because:-

- ① Handles continuous data
- ② Assumes Normal distribution
- ③ Computes class probabilities
- ④ Smooth decision boundaries

v2

How can the model performance be evaluated for any classification algorithm?

→ ①

Accuracy → Overall correctness of prediction

②

Precision → correct positive prediction out of total predictive positives

③

F1 score → Harmonic mean of precision & recall.

④

Cross Validation → evaluates scalability & generalization