

EXPERIMENT 8

Weka Explorer

PreprocessClassifyClusterAssociateSelect attributesVisualize

Classifier

ChooseNaiveBayes

Test options

☐ Use training set

☐ Supplied test set

☒ Cross-validation

☐ Percentage split

Folds10

%66

More options...

(Nom) class

StartStop

Result list (right-click for options)

08:46:00 - bayesNaiveBayes

Classifier output

==== Run information ====

Scheme: weka.classifiers.bayes.NaiveBayes

Relation: iris

Instances: 150

Attributes: 5

sepal.length

sepal.width

petal.length

petal.width

class

Test mode: 10-fold cross-validation

==== Classifier model (full training set) ====

Naive Bayes Classifier

Class

AttributeIris-setosaIris-versicolourIris-virginica

(0.33)(0.33)(0.33)

====

sepal.length

mean4.99135.93796.5795

std. dev.0.3550.50420.4353

weight sum505050

precision0.10590.10590.1059

sepal.width

mean3.40152.76072.9629

std. dev.0.39250.30380.3008

weight sum505050

precision0.10910.10910.1091

petal.length

mean1.46944.24525.5516

std. dev.0.17820.47120.5529

weight sum505050

precision0.14050.14050.1405

petal.width

mean0.27431.30972.0343

std. dev.0.10960.19150.2646

weight sum505050

precision0.11430.11430.1143

Time taken to build model: 0.01 seconds

==== Stratified cross-validation ====

==== Summary ====

Correctly Classified Instances14496%

Incorrectly Classified Instances64%

Kappa statistic0.94

Mean absolute error0.0342

Root mean squared error0.155

Relative absolute error7.4597%

Root relative squared error32.8794%

Total Number of Instances150

==== Detailed Accuracy By Class ====

TP RateFP RatePrecisionRecallF-MeasureMCCROC AreaPRC AreaClass

1.0000.0001.0001.0001.0001.0001.0001.000Iris-setosa

0.9600.0400.9230.9600.9410.9110.9920.993Iris-versicolour

0.9200.0800.9680.9200.9390.9100.9920.994Iris-virginica

Weighted Avg.0.9600.0200.9600.9600.9600.9400.9940.993

==== Confusion Matrix ====

a b c <- classified as

50 0 0 | a = Iris-setosa

0 49 2 | b = Iris-versicolour

0 4 46 | c = Iris-virginica

Status

OK

Log

x0

Student Name : Roll No. :

Experiment No. : Date :

Experiment-8

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

$$\text{In PDF} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right).$$

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

- Steps:-

- ① Open weka and go to explorer panel.
- ② Load dataset (IRL 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 classifies IRL dataset by predicting species based on feature probabilities & class wise statistics.

Viva-Questions:-

Q1 Enumerate significance of Gaussian probability density function (GPDF) for NB 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 Smooths decision boundaries

Q2 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