Experiment 5: Prediction Using Bayesian Classification

Aim:

To predict the **species** of a new Iris sample using **Bayesian classification** and display the **probability of each class**.

Theory:

- **Bayesian classification** (Naive Bayes) is a **probabilistic classifier**.
- Assumes **attribute independence** and computes the probability of each class for a given instance.
- The class with the **highest probability** is selected as the predicted class.
- Useful for **quick and accurate classification** of datasets.

Dataset (iris.arff)

@relation iris

@attribute SepalLength numeric

@attribute SepalWidth numeric

@attribute PetalLength numeric

@attribute PetalWidth numeric

@attribute Species {Setosa, Versicolor, Virginica}

@data

5.1,3.5,1.4,0.2,Setosa

4.9,3.0,1.4,0.2,Setosa

6.2,3.4,5.4,2.3,Virginica

5.9,3.0,5.1,1.8,Virginica

6.0,2.2,4.0,1.0,Versicolor

5.5,2.3,4.0,1.3,Versicolor

...

New Test Sample: SepalLength=6.1, SepalWidth=2.9, PetalLength=4.7, PetalWidth=1.4

Procedure (Using WEKA):

1. Open **WEKA → Explorer**.
2. Click **Open File** → select **iris.arff**.
3. Go to **Classify tab**.
4. Choose **Classifier → bayes → NaiveBayes**.
5. Click **Start** to train the model.
6. Use **Supplied test instance** to predict **Species**.
7. Observe **probability distribution** for each class.

Result (Sample / Expected):

- **Predicted Species:** Versicolor
- **Probability of Each Class:**
    - Setosa: 0.01
    - Versicolor: 0.85
    - Virginica: 0.14

Conclusion:

- Bayesian classification predicted the **species** based on computed probabilities.
- Naive Bayes is effective for **numerical and categorical data**.
- WEKA provides **easy computation of class probabilities** and predictions.