FAST-NUCES

Course Code: AI-2002
Course Name: Artificial Intelligence
Assignment 3
Due Date: March 27

1 Problem Statement

This assignment consists of two classification problems:

1. Implementing the K-Nearest Neighbors (KNN) classifier from scratch and comparing the results with the Scikit-learn implementation.
2. Implementing the Naïve Bayes classifier from scratch and comparing it with the Scikit-learn implementation.

2 Question 1: K-Nearest Neighbors (KNN)

2.1 Dataset

We have the following dataset for classification based on brightness and saturation:

Brightness	Saturation	Hue	Contrast	Class
40	20	150	30	Red
50	50	180	40	Blue
60	90	200	50	Blue
10	25	120	25	Red
70	70	220	60	Blue
60	10	140	35	Red
25	80	190	45	Blue

Table 1: KNN Training Dataset

2.2 KNN Implementation from Scratch

Using Euclidean distance formula:

$$d = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \tag{1}$$

Write the code to calculate distances, sort, and classify based on the nearest neighbors.

2.3 KNN Using Scikit-learn

Using the built-in KNN classifier from Scikit-learn, fit the model and compare the results.

3 Question 2: Naïve Bayes Classifier

3.1 Dataset

Given email dataset with categorical features:

Contains "Buy"	Contains "Win"	Email Length	Special Characters	Class
Yes	No	Short	Few	Not Spam
No	Yes	Long	Many	Spam
Yes	Yes	Long	Many	Spam
No	No	Short	Few	Not Spam
Yes	No	Long	Many	Spam

Table 2: Naïve Bayes Training Dataset

3.2 Naïve Bayes Implementation from Scratch

Using Bayes' Theorem:

$$P(Class|Features) = \frac{P(Features|Class)P(Class)}{P(Features)} \tag{2}$$

Write the code to compute probabilities and classify new entries.

3.3 Naïve Bayes Using Scikit-learn

Using Scikit-learn's built-in Naïve Bayes classifier, fit the model and compare the results.

4 Comparison and Conclusion

Compare the accuracy and results of both classifiers implemented manually and using Scikit-learn. Additionally, visualize the performance comparison of both models using graphs.

5 Submission

Students are required to submit the following:

- A detailed report explaining the implementation and comparison, including graphs comparing the performance of both the Scikit-learn model and the model from scratch.
- A Jupyter Notebook containing the implementation of both classifiers.