Deadline: 28th March 2022

Task:

Iris dataset is mostly employed to understand multi classification problem. The data set contains 3 classes (Iris Setosa, Iris Versicolour, and Iris Virginica) of 50 instances each, where each class refers to a type of iris plant. Four features are used to describe instances of each class. The four features are as follows:

1. sepal length in cm
2. sepal width in cm
3. petal length in cm
4. petal width in cm

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

The task requires you to apply K Nearest Neighbour (KNN) algorithm on iris dataset and perform an in-depth analysis. Following general steps are to be followed:

1. Load data.
   ```
   from sklearn import datasets
   iris = datasets.load_iris()
   ```

2. Divide into features and labels and encode labels using LabelEncoder().

   ```
   from sklearn.preprocessing import LabelEncoder
   ```

3. Split data into training and test data sets

   ```
   from sklearn.cross_validation import train_test_split
   ```

4. Fit KNN classifier on the training set.

   ```
   from sklearn.neighbors import KNeighborsClassifier
   ```

```
classifier.fit(X_train, y_train)
```

5. Make predictions on the test data.

```
y_pred = classifier.predict(X_test)
```

6. Evaluate model using confusion matrix.

```
cm = confusion_matrix(y_test, y_pred)
```

7. Calculate model accuracy.

Prepare a comprehensive report with **detailed analysis** on how the performance varies with respect to:

   a. Different values of K
   b. Different distance functions
   c. Different training and test split ratios
   d. Different label encoding techniques

Please note that implementation is not the objective of this assignment and you are free to borrow source code (with acknowledgement). The assignment will be assessed on the depth of analysis in your report and not just the source code.

Submission Requirements

Submit your assignment in the form of a technical report. The report must be formatted according to **IEEE two column conference paper template**[1]. Attach the source code of your implementation as an annexure to the report. You are free to seek help from any sources with proper citation/acknowledgment.

Note: This is an individual assignment. No marks will be awarded for PLAGIARIZED work.

---

[1] https://www.ieee.org/conferences/publishing/templates.html