DEPARTMENT OF COMPUTER SCIENCE

ARTIFICIAL INTELLIGENCE: TEST

DURATION:1HR 30 MINS

EXAMINER: T D KAVU

INSTRUCTIONS:

1. Answer all questions.
2. This is an **open book** practical test.
3. Create a **Google Colab file** and put all your experiments(code) in that file, submit a link for the file on LMS, make sure you have set the share options to (**anyone with the link can edit the file**).

Question One [50 marks]

Topic: Unsupervised Learning

(a) Application of Association rule mining.

A given a dataset ***transactions.txt*** which consists of text that looks as follows:

```
1 3 4
1 2 3 5
2 3 5
2 5
1 2 3 6
```

In the file, blanks separate items (identified by integers) and new lines separate transactions. For example, the above illustration contains information about a total of 5 transactions and its second transaction consists of 4 items.

(a) Using the **apriori** algorithm and the dataset given, generate the association rules which have minimum support value of 0.157 and minimum confidence value of 0.9. [20 Marks]

(b) Print the rules as a DataFrame sorted by the number of items that they contain in decreasing order. [10 Marks]

(c) Print the rules as a DataFrame sorted by the confidence value in decreasing order. [10 Marks]

(d) Print the rules as a DataFrame sorted by the lift value in decreasing order. [10 Marks]

Total Marks [50 Marks]

Question Two [30 marks]

Topic: Ensemble Learning

Given a Dataset CSV File (pima-indians-diabetes.csv) perform the following operations:

a) Scale the **proper** independent variables using a proper scaling algorithm. [5]

b) Split the data in a proportion of 80% training and 20% testing, again make sure the target variable is **evenly** distribute between the training and testing set. [5]

c) Build, train and evaluate (using a **classification report**) an AdaboostClassifier. [5]

d) Build, train and evaluate (using a **classification report**) an GradienBoostClassifier. [5]

e) Build, train and evaluate (using a **classification report)** a hard VotingClassifier composed of (decision trees, support vector classifier and naive bayes classifier). [5]

f) Which among the three classifiers is the best according to the classification report (put your answer as a *comment* in the last cell). [5]

Total Marks **[30 Marks]**

THE END