Python_Lesson4: Python Programming

Lesson Overview:
In this lesson we will introduce classification.
b. Classification algorithm
c. Scikit learn
d. Advanced concept related to machine learning algorithm like overfitting, underfitting, cross validation, evaluation for clustering methods

Use Case Description:
k-nearest neighbor classifier

Programming elements:
Classification

Source Code:
https://umkc.box.com/s/g9fc03pkton7hptl8wbj3k5ro5ik7mqr

Data Set:

URL: https://app.box.com/file/468476100076
Dataset description: https://www.kaggle.com/uciml/glass/downloads/glass-classification.zip/1
The name of target Column is **Type**

In class programming:
1. find the correlation between supervised (target column) and sex column for the use case in calss. Do you think we should keep this feature?
2. Implementing Naïve Bayes method using scikit-learn library
Use dataset available in https://app.box.com/file/468476100076
Use **train_test_split** to create training and testing part
Evaluate the model on testing part

3. Implement linear SVM method using scikit library
Use the same dataset above
Use **train_test_split** to create training and testing part
Evaluate the model on testing part
Which algorithm you got better accuracy? Can you justify why?

4. use the SVM with RBF kernel on the same dataset. How the result changed?

ICP Submission Guidelines (for In Class students):
1. ICP Submission is in pairs of two students.
2. Once completed, must be presented to TA or Instructor before the completion of the class
3. Submission after class is considered as a late submission. (Check the late submission policy in the syllabus)
4. ICP Code with brief explanation should be pushed to GitHub. Submit GitHub link through the Feedback Form: https://docs.google.com/forms/d/e/1FAIpQLSdmJkDgBMxr4qv73c9y5k1jtky44-sMmOI1v1jFtNEbUJ6H9A/viewform

Online Submission Guidelines (for Online students):
1. Submit your source code and documentation to GitHub and represent the work through wiki page properly (submit your screenshots as well. The screenshot should have both the code and the output)

2. Comment your code appropriately
3. Video Submission (2 – 3 min video showing the demo of the ICP, with brief voice over on the code explanation)
4. Submission after class is considered as a late submission. (Check the late submission policy in the syllabus)
5. Use the following Google link to submit your ICP # (GitHub wiki page link for ICP #):
https://docs.google.com/forms/d/e/1FAIpQLSdmJkDgBMxr4qv73c9y5k1jtky44-sMmOI1v1jFtNEbUJ6H9A/viewform

Evaluation Criteria:

1. Completeness of Features
2. Code Quality (https://en.wikipedia.org/wiki/Best_coding_practices)
3. Time
4. Feedback Submission

Note: *Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL: https://catalog.umkc.edu/special-notices/academic-honesty/*