Final Project Submission Guidelines and Requirements

Overview:
For the final project in the Web Mining course, students are required to select and implement one algorithm from the list below. The project aims to deepen your understanding of data mining techniques, their practical applications, and analytical insights.

Algorithm Options:
1. Apriori
2. FP-Growth
3. DBSCAN
4. Bayes Classification
5. CAR
6. K-means
7. ID3

Project Examples:

Example 1: Text Categorization using Multinomial Naïve Bayesian Method
Purpose:
Implement text topic categorization to understand the principles and methods of the multinomial naïve Bayesian algorithm. Gain insights into basic preprocessing steps in textual data mining.

Requirements:
a. Show micro-average and macro-average accuracy, precision, recall, and F1-value using 10-fold cross-validation.
b. Use the dataset available at [THUCTC](http://thuctc.thunlp.org/).
c. Choose any programming language, such as C, C++, Java, or Python.
d. Prepare an experimental report including the following sections:
《Data Mining》Final Report
- Title
- Student Number
- Name
1. Purpose
2. Algorithm Description and Analysis
3. Tricks and Details in Implementation
4. Data and Testing Method
5. Results and Their Analysis
6. Conclusion
7. References

Example 2: Implementing the Apriori Algorithm and testing it empirically.
Purpose:
Understand and implement the Apriori algorithm to test its performance and behavior with transactional datasets.

Requirements:
a. Test implementation with any transactional dataset.

b. Show frequent itemsets and strong association rules for varying Min_Sup and Min_Conf values.
c. Analyze time and memory usage under different Min_Sup and Min_Conf configurations.
d. Choose any programming language, such as C, C++, Java, or Python.
e. Prepare an experimental report including the same sections listed in Example 1.

Submission Guidelines:

a. Deliverables:
I. A project report of at least 4000 words adhering to the format provided.
II. A screen-recorded video showing the real-time execution of your project code.

b. Submission Deadline: 10th January 2025

c. Submission Method: Email both the report and the video to 3168744799@qq.com

d. Honesty Code:
I. All submissions must be your original work.
II. Any form of cheating or plagiarism will result in a ***ZERO*** grade.

Grading Criteria:
- Algorithm understanding and analysis.
- Accuracy and clarity of results.
- Quality of the report and adherence to formatting requirements.
- Creativity and effort in implementation.

Good luck with your project!