CDS 503: Machine Learning
Academic Session: Semester 1, 2022-2023

School of Computer Sciences, USM, Penang

ASSIGNMENT 2

Task Type

Assignment 2 is a group assignment of 2 per group. The members are **not** necessarily same as project group members.

Assignment Description

1. Find **any problem** that can be solved using **Apriori-based algorithms**.
2. Explain the problem and how you can do an association analysis using the Apriori algorithm based on the steps below:
    - Preparing the data for the apriori algorithm
    - Setting the parameters for the algorithm and running the algorithm
    - Inspecting the frequent items
    - Generate association rules and analyze the results using confidence, support, and lift.
3. Develop the Apriori algorithm **WITHOUT** using any toolbox. However, you can use any library to helpyou complete the association analysis.
4. You have to code your Apriori algorithm in Phyton.

Report Requirement and Format

- A report must be prepared using Microsoft Word, font type Arial, size 12 in a single line, between 8 – 12 pages. Every chapter should start with a **new** page (Chapters 1.0 to 5.0, and references).
- A **cover page** should contain the course name (including semester and year), assignment title, name, and matrix no.
- Table of Contents

    1.0 Problem Background
    2.0 Data Understanding & Integration
    3.0 Pre-processing options
    4.0 Experiment and Analysis using Apriori Algorithm
    • Setting the parameters for the algorithm and running the algorithm.
    • Inspect frequent items.
    • Generate association rules and analyze the results using confidence, support, and lift.
    5.0 Conclusion

    References

    Note: You may create additional subsections as deemed necessary.

CDS 503: Machine Learning
Academic Session: Semester 1, 2022-2023

School of Computer Sciences, USM, Penang

ASSIGNMENT 2

Report Submission Instruction

- Submit a soft copy (PDF to eLearn@USM).
- The PDF file must be named according to the following notation: ***CDS503_Assignment 2_Name1_ Name2.***

Assignment Evaluation

This assignment will be graded (A to F scale).

IMPORTANT: Students who copied or plagiarized other's work or let their work be copied or plagiarized will be given an F grade. The student may be barred from sitting for final exam and reported to the university's disciplinary board.

Assignment Due Date: Refer to the course planner.

Grading Rubric – Assignment 1

Course Learning Outcome (CLO):

- C03: Design relevant machine learning solutions for typical real-world problems.

Rubrics

Component	2-1 (Poor)	5-3 (Average)	8-6 (Good)	10-9 (Excellent)	Weigh t
Intro & Problem Background	Introduction and problem background are **poorly** explained.	Introduction and problem background are **fairly** explained.	Introduction and problem background are **adequately** explained.	Introduction and problem background are **clearly** explained.	20%
Data Understanding & Integration	Data Understanding & Integration are **poorly** explained.	Data Understanding & Integration are **fairly** explained.	Data Understanding & Integration are **adequately** explained.	Data Understanding & Integration are **clearly** explained.	15%

CDS 503: Machine Learning
Academic Session: Semester 1, 2022-2023

School of Computer Sciences, USM, Penang

ASSIGNMENT 2

Pre-processing options	Pre-processing options are **minimally** discussed and justified.	Pre-processing options are **fairly** discussed and justified.	Pre-processing options are **adequately** discussed and justified.	Pre-processing options are **clearly** discussed and justified.	10%
Experiment Result & Analysis	The experiment is **poorly** performed according to the plan. The results and analysis are **poorly** discussed.	The experiment is **minimally** performed according to the plan. The results and analysis are **minimally** discussed.	The experiment is **fairly** performed according to the plan. The results and analysis are **fairly** discussed.	The experiment is **clearly** performed according to the plan. The results and analysis are **clearly** discussed.	45%
Conclusion & References	The conclusion is **absent,** and no references provided.	The conclusion is of **simplistic** summary and **few** references are provided.	The conclusion is a **partially** complete summary and **adequately** references are given.	The conclusion contains a **comprehensive** summary and good references are provided.	5%
Report Formatting	Some writings are inaccurate and unclear. Follow the format given and somewhat organized.	Some writings are inaccurate and unclear. Follow the format given and somewhat organized.	Most writings are accurate, clear and concise. Somewhat follow the format and organized.	Most writings are accurate, clear and concise language used throughput. Report follows the format given and is properly arranged and well-organized.	5%