PREDICTIVE BUSINESS ANALYTICS PROJECT

Task Type

This project is weighted as **20%** of the assessed work for the course. The project is a **group** work.

COURSEWORK SUBMISSION GENERAL INFORMATION

Academic Integrity Statement

You must adhere to the university regulations on academic conduct. Formal inquiry proceedings will be instigated if there is any suspicion of plagiarism or any other form of misconduct in your work. Students must **NOT** collude with other groups of students or plagiarize their work. We practice zero tolerance towards plagiarism, and we use Turnitin to evaluate the similarity index. Your similarity index score must not exceed 20%. Your tasks must be your own work. Unless the use of Artificial Intelligence (AI) is permitted in your assessment task, using AI to complete your project is a form of plagiarism.

Nature of the submission required

A softcopy of your assignment in PDF version should be submitted to lecturer through eLearn@USM. The zip package must be named according to the following notation:

Diagrams may be used where they are helpful to support your arguments or description. If they are not your own work, the source must be referenced. Please help us to handle and mark your work efficiently.

Documentation guidelines

Groups are required to submit a **SOFTCOPY** of the report and ensure that it uses the following formatted styles: 1) Font type: **ARIAL**, 2) Font size: **11 pt**., 3) Line spacing: **Single spacing**, 4) Page layouts: **Justify**, and 5) Min - Max page number: **30 - 40 pages** (exclude cover page, TOC, TOF, abbreviation and Appendix). Please also include a proper cover page **(Appendix-B)**. Please make sure you have proper format alignment for all paragraphs, following standard writing style and use **American Psychological Association (APA)** for citation. Please include a **HEADER** with the following information: **Course code, Academic Year, Project, and Group Number** (example of header: **CDS513 Project Group 01**). Please zip your files into one folder. The zip package is named CDS513_2023-2024_Project_G01.

PREDICTIVE BUSINESS ANALYTICS PROJECT

Project Description and Instructions

Define your project problem and describe it in terms of its real BUSINESS world's application.

For data sets, you may choose your own data set either from an online source or one of your own. Note that you need to seek for two related datasets (discrete and continuous datasets).

Based on your own selection of datasets, do some experiments, and perform the analysis on the following approaches. You may perform various experiments for each approach (below) in order to reveal the pattern and insight of your work.

- *Market Basket Analysis*
- *Recommender Systems*
- *Regression Analysis (no time-series allowed)*

To understand the requirements of the project further, the links below may help you.

The **first link is divided into 3 parts**, where it discusses on complete business model and discuss a step-by-step process of how analytics is set up in a new business, how is it used in daily processes and some of the advanced analytics techniques which a business can use to make meaningful segmentation and prediction to optimize its marketing and sales campaign. It's a good startup for you to understand what kind of data it should be, how to relate the data and understand it further to come up with a good storyline for the project.

Part 1: https://www.analyticsvidhya.com/blog/2014/02/learn-analytics-business-case/
Part 2: https://www.analyticsvidhya.com/blog/2014/03/learn-analytics-business-case-study-part-ii/
Part 3: https://www.analyticsvidhya.com/blog/2014/03/learn-analytics-business-case-study-part-iii/

The **second link** is more on general discussion, on the steps that need to be done to tackle the business analytics problem.
https://www.bigtime.net/blogs/forecasting-revenue-predictive-analytics/

Perform in-depth observation and analysis based on your experiments and results. Research and adopt suitable visualization techniques to represent the results. Write the conclusions and future research work. Provide sample output screen shots and explain each of them.

PREDICTIVE BUSINESS ANALYTICS PROJECT

Report Format

You are required to submit your final report with the following report format:

- A project report must be prepared using Microsoft Word/PDF, font type Arial, size 12, single space.
- A cover page should contain course name (including semester and year), project title, group number, matrix no and name.
- Group contribution page that indicates name of the members and task handled.
- Sample Table of Contents:
    Abstract
    1.0   Project Background *(~2 pages)*
        1.1   Background of the problem domain
        1.2   Issue and Problem Statement
        1.3   Objective/s of the project
    2.0   Literature Review *(~6 pages)*
        [Simple literature review on your domain problem and approaches chosen]
    3.0   Methodology *(~10 pages)*
        3.1   Project Framework
        3.2   Dataset and Preparation
        3.3   Approach, algorithms etc. [Theoretical description, implementation details]
    4.0   Experiment and Analysis *(~10 pages)*
        4.1   Experimental Setup
        4.2   Result Analysis
    5.0   Conclusion *(~1 page)*
    References *(~1 page)*
    Appendix
        Project source code
        Sample output screen shots
        [any others relevant]

Project Presentation

Each group is required to present their project by using ONE-page POSTER (portrait or landscape), font type Arial, size 11. Each group is given 15 + 5 minutes to present. The latter 5 minutes is for questions and answers. All members are required to present. So, distribute your work accordingly.

The poster should include (not limited to):

- The background of the problem and issues handled.
- Describe and analyze the data obtained.
- The approaches, the experiments and result analysis.
- Show the implementation of the solution and the results of your experiments.
- Findings and conclusion.
- Each group member should present.

PREDICTIVE BUSINESS ANALYTICS PROJECT

Submission Instructions

- Submit your assignment in soft copy using Microsoft Word **AND** .pdf, font type Arial, size 11. Submission link will be provided soon.
- A **cover page** should contain course name (including semester and year), Project title, student names and matrix numbers.
- Soft copy:
    - i) Final report in Microsoft Word
    - ii) Final report in .pdf file
    - iii) Dataset file(s)
    - iv) Programming file(s)
- The .zip/.rar package must be named according to the following notation: **CDS513_2023-2024_Project_GroupNumber.** For example, the zip package is named *CDS513_2023-2024_Project_G01.*

Important Dates

The project due date is **Revision Week:**
(1) Report submission (**due 7 July 2024**, Sunday, 5.00 p.m.)
(2) One-page poster for presentation (**due 9 July 2024**, Tuesday, 5.00 p.m.)
(3) Project presentation date (**10 July 2024 (online)** – slots to choose from will be available on e-learn from 2 July 2024, Tuesday).

Grading

The project contributes 20% of your overall grade. The marks distribution of the project is as follows:

- Project report: **15%**
- Project presentation: **5%**

Project Evaluation

This assignment will be graded (1 - 10 scale for each component in grading rubric).

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

CDS513: Predictive Business Analytics
Academic Session: Semester 2, 2023-2024
School of Computer Sciences, USM, Penang

PREDICTIVE BUSINESS ANALYTICS PROJECT

Project Grading Rubric

Course Learning Outcome (CLO):

- CLO2 Design strategies relevant to predictive business analytics using appropriate technologies and tools.
- CLO4 Propose a business model by incorporating predictive business analytic approaches.
- CLO5 Evaluate predictive business analytics using recent tools.

Grading Rubric – Project Report (15%):

Component	1-2 (Poor)	3-5 (Average)	6-8 (Good)	9-10 (Excellent)	Weight
Abstract & Background	It is **unclear** what is being defined.	Introduction and background are **not** linked very well. It is **not clearly** stated.	Introduction and background are linked **adequately** well. It is adequate.	Introduction and background are linked very well. It is **clearly** stated.	5%
Issues & Problem Statement	It is **unclear** what is being defined.	Problem statement and issue is **not** clear and demonstrates minimal knowledge or depth of the project.	Problem statement and issue is **adequate** and demonstrates a good understanding of the project.	Problem statement and issue is **clear** and concise and demonstrates a deep understanding of the project.	10%
Objectives & Motivation	It is **unclear** what is being defined. Motivations are **not** stated.	Objectives are **not clear,** and motivation is **vaguely** stated.	Objectives are **adequate,** and motivation are **fairly** stated	Goals and motivation are **clearly** stated.	10%
Data Description	Dataset description is **absent**.	Dataset description is **minimal**.	Dataset description is **adequately** complete.	Dataset description is **complete** and **comprehensive**.	10%
Scope, Limits & Ref	Scope is **absent** and no references provided.	Scope is **vaguely** defined, and **few** references are provided.	Scope is **clearly** defined, and **adequately** references are given.	Scope is very **clearly** defined, and good references are provided.	5%

5

PREDICTIVE BUSINESS ANALYTICS PROJECT

Method & Modelling	The data preprocessing, exploration and cleaning are **not** explained. The data modelling techniques are **poorly** presented.	The data preprocessing, exploration and cleaning are **minimally** explained. The data modelling techniques are **minimally** explained.	The data preprocessing, exploration and cleaning are **fairly** explained. The data modelling techniques are **fairly** explained, i.e., parameters, fine-tuning and compared in terms of performance metrics.	The data preprocessing, exploration and cleaning are **clearly** explained. The data modelling techniques are **clearly** explained, i.e., parameters, fine-tuning and compared in terms of performance metrics.	20%
Experiment Result & Analysis	The model is **poorly** presented, and discussion of the model is **absent**. The results are **poorly** discussed and did not fit the research objectives. Insights and contributions are **poorly** discussed or absent.	The best-suited model is **minimally** discussed and justified. The results are **minimally** discussed and tailored to the research objectives. Insights from the analysis are **vague.**	The best-suited model is **fairly** discussed and justified. The results are **fairly** discussed and tailored to the research objectives. Insights from the analysis are **less evident,** and contribution is **fairly** discussed.	The best-suited model is **clearly** discussed and justified. The results are **clearly** discussed and tailored to the research objectives. Insights from the analysis evident and contribution are discussed and well-explained.	20%
Conclusion & Future Research	The conclusion is **absent.**	The conclusion is of **simplistic** summary.	The conclusion is a **fairly** complete summary.	The conclusion contains a **comprehensive** summary.	6%
Screen Shots & Appendix	Screenshots are **absent,** and info is **not** properly arranged in Appendix.	Screenshots are **simplistic,** and related info is **vaguely** stated in Appendix.	Screenshots are **moderately** arranged, and related info is stated **clearly** in Appendix	Screenshots are **well-arranged,** and related info is stated **clearly** in Appendix.	6%
Report Format & Neatness	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.	8%

PREDICTIVE BUSINESS ANALYTICS PROJECT

Grading Rubric – Poster & Presentation (5%):

Component	1-2 (Poor)	3-5 (Average)	6-8 (Good)	9-10 (Excellent)	Weight
Poster Clarity & Attractiveness	Poster is poorly arranged. Presentations are poorly arranged with too little components.	Poster is adequately arranged. Presentations are adequately arranged with few components.	Poster is fairly arranged. Presentations are fairly arranged with some components.	Poster is clearly arranged. Enhances presentation with effective arrangements of all components.	30%
Presentation Delivery	Students(s) lacks confidence and hard to understand what was spoken.	Students(s) confidence is fair. Speaks fairly clear and holds the attention of the listener fairly.	Student(s) confidence is good. Speaks clearly and eloquently. Emphasize important ideas and hold the listener's attention.	Student(s) confidence is noteworthy. Speaks very clearly and eloquently. Emphasize important ideas and hold the listener's attention greatly.	70%