MIS710 – Machine Learning in Business - Trimester 1 2024
Assessment Task 2 – Case Study Report and Business Report – Individual

DUE DATE:	**Friday, 31st May 2024, by 8:00pm (AEST)**
PERCENTAGE OF FINAL GRADE:	**60% including Part A (40%) and Part B (20%)**
WORD COUNT:	**Part A Technical Report: Maximum number of words: 2000 words** **Part B Business Report: Maximum number of words: 1000 words**

Description

Purpose

There are two parts in this assignment.

- Part A provides you with opportunities to learn a range of machine learning methods and Python skills (GLO1 & ULO1) and apply your digital literacy to research and develop a machine learning solution (GLO3, GLO5, and ULO2). By completing this task, you will gain knowledge and skills in selecting and applying one or more appropriate machine learning algorithm(s) to develop and evaluate a machine learning solution and interpret the outcomes.
- In Part B, you will report your application of machine learning and make recommendations to the business and management audience. By completing this task, you will gain ability to explain and justify machine learning options and discuss their pros and cons to the business audience.

Context/Scenario

The Great Ocean Banking Group, commonly known as Great Ocean Bank, a regional bank in Victoria, Australia, serves over 1 million customers and offers a comprehensive range of financial services predominantly to residents of Victoria and other states. With diverse offerings, including Home Mortgages, Personal Loans, and various savings and credit card accounts, the bank strives to meet the multifaceted financial needs of its customers. Through its periodic sales campaigns, it engages not only its existing clientele but also reaches out to prospective customers, highlighting the extensive benefits of its banking products and fostering a culture of financial well-being.

Aiming to enhance the effectiveness of these campaigns and to gain deeper insights into customer needs, the bank explores data analytics and Machine Learning opportunities to refine their marketing strategies and improve the overall customer experience, thereby reinforcing the Great Ocean Banking Group's dedication to building lasting relationships and ensuring financial prosperity within the community.

You have been approached by Mr. Gary Peterson, Head of Customer Experience, and Ms. Kathy Hoang, Head of Data Analytics, at the Great Ocean Banking Group - the client. They are interested in exploring the dataset **(GoBank.csv),** which is a subset of data compiled from various databases of the Great Ocean Banking Group. This dataset includes customer information, banking relationships, details of the last contact from a marketing campaign, performances from a previous campaign, and several economic indices, such as the RBA cash rate. Further details are available in **GoBank_metadata.csv**.

Mr. Gary Peterson and Ms. Kathy Hoang seek to better understand their customers and the factors influencing their campaign performances, aiming to predict potential 'Sale' or 'No Sale' outcomes. You are tasked with performing two analytical tasks: uncovering data insights and exploring machine learning opportunities.

Regarding the former, Great Ocean Bank has requested an analysis on the following:

1. How do the demographic details of customers, such as Age and Qualification, among others, influence the success of sales outcomes (Sale Outcome)?
2. Does having any type of account (be it a home mortgage, personal loan, or any other account) influence a customer's decision to open an additional account with Great Ocean Bank during the current campaign, thereby impacting the Sale Outcome?
3. How does the method of our last contact with a customer (whether we called them - Outbound, or they called us - Inbound) influence their decision to open a new account (Sale) or not (No Sale)?
4. Does the outcome of previous campaigns influence the success of subsequent sales efforts (Sale Outcome)?
5. How do economic indicators (RBA Cash Rate, Employment Variation Rate, Consumer Confidence Index) impact Sale Outcomes?
6. Are there other things Great Ocean Bank hasn't considered yet that might affect whether their sales are successful, or do different factors influence each other?

Regarding the latter point—machine learning opportunities—the client is interested in exploring machine learning models. At this stage, they require one predictive machine learning model to predict potential 'Sale' or 'No Sale' outcomes and one clustering model to be developed. Suggestions for future opportunities are also welcome.

Datasets provided:

- **GoBank.csv**

Data description

- **GoBank_metadata.csv**

You are required to explore the first dataset, *GoBank.csv*, and develop and test machine learning option(s) using Python. You are also required to develop two reports:

- **The first technical report (Part A)** should present your analysis and findings to Kathy Hoang, Head of Data Analytics. This report should detail your approach to exploring the dataset, the machine learning techniques used, and your findings. Your findings should be supported by relevant visualisations and statistical analysis. This report should also develop and compare **two** predictive machine learning models using different algorithms, recommend **one** of these models; develop and evaluate **one** clustering model; inform model deployment, and recommend future engagements with the client. See further details in the Specific Requirements section below.
- Based on your data analysis and machine learning findings, you are then required to develop a **consultancy report (Part B)** for Gary Peterson, Head of Customer Experience. The report should include your response to the client's **six** (6) questions, the proposed machine learning models, and recommendations for use. You should also discuss the limitations of your approach and any potential areas for future improvements. See further details in the Specific Requirements section below.

See further details in the section Specific Requirements below.

The dataset, acquired from a public source, has been modified and pre-processed for the assessment purpose.

Specific Requirements

You are required to:

- Develop your business and data understandings.
- Prepare and explore the provided dataset, cleanse and pre-process data as needed. Undertake an exploratory data analysis (EDA) to respond to the client's six questions.
- Undertake supervised machine learning model development, evaluation, and selection. Two predictive models should be developed, tested, and compared.
- Undertake unsupervised machine learning using clustering analytics.
- Develop two reports:
  - **The first technical report (Part A)** should present your EDA (Exploratory Data Analysis) and machine learning findings to Kathy Hoang, Head of Data Analytics.
  - **The second consultancy report (Part B)**, for Gary Peterson, Head of Customer Experience, should present, in the business domain language, responses to the six specific requests about data, insights from clustering analytics, and **one** predictive machine learning model.
- Format and present your report professionally. Two sample report templates are provided under Assessment Resources.
- Correctly use the APA7 style of referencing, and include in-text citations when quoting, referring to, summarising, or paraphrasing from any source:

  https://www.deakin.edu.au/students/studying/study-support/referencing

Deliverables:

Part A. Case Study Report

Part A.1 Machine Learning Solution

- A cover page (**not** included in the word count) that includes:
  - Report Title
  - Unit code and name
  - Student name and student ID
- A table of contents (**not** included in the word count)
- An executive summary of max. 200 words is required (**included** in the word count).
- The report should include:
  1. Introduction:
     - Objective: the business problem to be addressed in its business context, and the value proposition of the project.
  2. Approach:
     - Overview of the machine learning approach, including machine learning types and problem(s), and prediction target(s).
  3. Data preparation and Exploratory Data Analysis (EDA):
     - Data sources, data size, types, quality, cleansing and pre-processing, and any observations.
     - EDA: statistical analysis and visualisation.
     - Key insights gained from EDA to inform feature selection and data splitting.

4. Model development and evaluation:
    - Supervised Machine Learning:
        - Two predictive models and performance metrics.
        - Model comparison based on your selection criteria.
    - Unsupervised Machine Learning:
        - Clustering analytics results and justification of the number of clusters.
5. Solution recommendation:
    - Interpretation and discussion of results obtained from the validation and comparison.
    - Solution recommendation – what predictive model is to offer to the client.
    - Future engagements with the client.
6. Technical recommendations:
    - Summary of the development and testing environment, such as software libraries, the programming language and computing environment used.
    - To inform model deployment, provide your machine process diagram and data pre-processing.
    - Suggestions for maintenance of accuracy and relevance over time (based on your research).

- References (**not** included in the word count)
- Optional appendices (**not** included in the word count – not subject to assessment), such as additional technical details, supplementary figures and tables.

Part A.2 Files

- A python notebook with detailed comments to guide the deployment team

Part B. Business report

- A cover page (**not** included in the word count) that includes:
    - Report Title
    - Unit code and name
    - Student name and student ID
- A table of contents (**not** included in the word count)
- An executive summary of max. 100 words is required (**included** in the word count).
- The report should include:
    1. Introduction:
        - Business understanding of the project using the Business Analysis Core Concept Model (BACCM) framework[1] .
    2. Insights from Exploratory Data Analysis (EDA):
        - Answers to the Client's six (6) questions.
        - Additional insights, such as comments on data quality or observations beyond the client's six questions and possible insights gained from clustering analytics.

[1] *Business Analysis Core Concept Model (BACCM) https://www.iiba.org/business-analysis-blogs/6-steps-to-applying-the-baccm/ see also Topic 1*

3. Proposed machine learning solution:
    - The selected machine learning model.
    - Interpretation of its performance and discussion of pros and cons.
4. Recommendations and conclusions:
    - Recommendations of business applications.
    - Potential benefits to stakeholders and how they relate to the value proposition.
    - Implications such as changes to business processes and decision making and possible impacts.
    - Recommendations for further improvements.

- References (**not** included in the word count)
- Optional appendices (**not** included in the word count – not subject to assessment), such as supplementary figures and tables.

Important Notes

- The final submission should be presented professionally. The reports should use clear, concise, and relevant language to communicate the content relevant to the target audiences.
- You should undertake research and use various tools to solve the business problem. In the end, you must exercise and understand the Python code yourself for your own learning purposes, develop and present your business understandings and solution to the client(s). Cite and reference any sources you use.

Student Toolkits

A set of toolkits was prepared by experienced Deakin students to help you learn the generic skills required in the Business & Law professions: https://d2l.deakin.edu.au/d2l/home/93063

You will find the following tool kits to be useful:

- Communication Skills - especially Writing Skills: https://d2l.deakin.edu.au/d2l/le/content/93063/viewContent/6086619/View

- Use APA7 style of referencing and include in-text citations: https://www.deakin.edu.au/students/studying/study-support/referencing

Learning Outcomes

This task allows you to demonstrate your achievement towards the Unit Learning Outcomes (ULOs) which have been aligned to the Deakin Graduate Learning Outcomes (GLOs). Deakin GLOs describe the knowledge and capabilities graduates acquire and can demonstrate on completion of their course. This assessment task is an important tool in determining your achievement of the ULOs. If you do not demonstrate achievement of the ULOs you will not be successful in this unit. You are advised to familiarise yourself with these ULOs and GLOs as they will inform you on what you are expected to demonstrate for successful completion of this unit.

The learning outcomes that are aligned to this assessment task are:

Unit Learning Outcomes (ULOs)		Graduate Learning Outcomes (GLOs)
ULO1	Analyse and frame business challenges using machine learning concepts, techniques, and the machine learning model development lifecycle.	GLO1: Discipline-specific knowledge and capabilities
ULO2	Select and apply appropriate machine learning techniques to solve business problems and evaluate the machine learning model performance.	GLO3: Digital literacy GLO5: Problem solving
ULO3	Explain the application of machine learning and interpret the outcomes to the various stakeholders	GLO2: Communication

Submission

You must submit your assignment in the Assignment Dropbox in the unit CloudDeakin site on or before the due date. The submission must include two files:

- Two (2) report documents in **PDF for marking**. Name your documents using the following syntax: **<your surname_your first name_your Deakin student ID number_[unitcodeA1].pdf**. For example, 'MIS710A1_Jones_Barry_123456789_MIS710A2 ReportA.pdf' and 'MIS710A1_Jones_Barry_123456789_MIS710A2 ReportB.pdf'
- Two (2) report documents **in Word as back-ups**. Name your documents using the following syntax: **<your surname_your first name_your Deakin student ID number_[unitcodeA1].doc (or '.docx')**. For example, 'MIS710A1_Jones_Barry_123456789_MIS710A2 ReportA.doc' and 'MIS710A1_Jones_Barry_123456789_MIS710A2 ReportB.doc'

- One (1) Python notebook

Submitting a hard copy of this assignment is not required. You must keep a backup copy of every assignment you submit until the marked assignment has been returned to you. In the unlikely event that one of your assignments is misplaced you will need to submit your backup copy.

Any work you submit may be checked by electronic or other means for the purposes of detecting collusion and/or plagiarism and for authenticating work.

When you submit an assignment through your CloudDeakin unit site, you will receive an email to your Deakin email address confirming that it has been submitted. You should check that you can see your assignment in the Submissions view of the Assignment Dropbox folder after upload and check for, and keep, the email receipt for the submission.

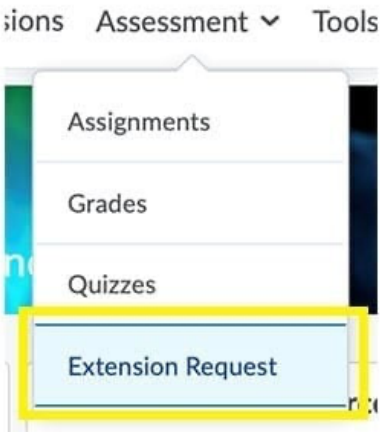
Marking and feedback

The marking rubric indicates the assessment criteria for this task. It is available in the CloudDeakin unit site in the Assessment folder, under Assessment Resources. Criteria act as a boundary around the task and help specify what assessors are looking for in your submission. The criteria are drawn from the ULOs and align with the GLOs. You should familiarise yourself with the assessment criteria before completing and submitting this task.

Students who submit their work by the due date will receive their marks and feedback on CloudDeakin 15 working days after the submission date.

Extensions

Extensions can only be granted for exceptional and/or unavoidable circumstances outside of your control. Requests for extensions must be made by 12 noon on the submission date using the online Extension Request form under the Assessment tab on the unit CloudDeakin site. All requests for extensions should be supported by appropriate evidence (e.g., a medical certificate in the case of ill health).



Applications for extensions after 12 noon on the submission date require University level special consideration and these applications must be must be submitted via StudentConnect in your DeakinSync site.

Late submission penalties

If you submit an assessment task after the due date without an approved extension or special consideration, 5% will be deducted from the available marks (60%) for each day after the due date up to seven days*. Work submitted more than seven days after the due date will not be marked and will receive 0% for the task. The Unit Chair may refuse to accept a late submission where it is unreasonable or impracticable to assess the task after the due date. *'Day' means calendar day for electronic submissions and working day for paper submissions.

An example of how the calculation of the late penalty based on an assignment being due on a Thursday at 8:00pm is as follows:

- 1 day late: submitted after Thursday 11:59pm and before Friday 11:59pm– 5% penalty.
- 2 days late: submitted after Friday 11:59pm and before Saturday 11:59pm – 10% penalty.
- 3 days late: submitted after Saturday 11:59pm and before Sunday 11:59pm – 15% penalty.
- 4 days late: submitted after Sunday 11:59pm and before Monday 11:59pm – 20% penalty.
- 5 days late: submitted after Monday 11:59pm and before Tuesday 11:59pm – 25% penalty.
- 6 days late: submitted after Tuesday 11:59pm and before Wednesday 11:59pm – 30% penalty.
- 7 days late: submitted after Wednesday 11:59pm and before Thursday 11:59pm – 35% penalty.

The Dropbox closes the Thursday after 11:59pm AEST/AEDT time.

Support

The Division of Student Life provides a range of **Study Support** resources and services, available throughout the academic year, including **Writing Mentor** and **Maths Mentor** online drop ins and the SmartThinking 24 hour writing feedback service at **this link**. If you would prefer some more in depth and tailored support, **make an appointment online with a Language and Learning Adviser**.

Referencing and Academic Integrity

Deakin takes academic integrity very seriously. It is important that you (and if a group task, your group) complete your own work in every assessment task Any material used in this assignment that is not your original work must be acknowledged as such and appropriately referenced. You can find information about referencing (and avoiding breaching academic integrity) and other study support resources at the following website: http://www.deakin.edu.au/students/study-support

Your rights and responsibilities as a student

As a student you have both rights and responsibilities. Please refer to the document ***Your rights and responsibilities as a student*** in the Unit Guide & Information section in the Content area in the CloudDeakin unit site.