

MIS710 – Machine Learning in Business - Trimester 2 2023

Assessment Task 2 – Case Study Report and Business Report – Individual

WORD COUNT:

Part A Technical Report: Maximum number of words: 1000 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

For your information:

- **FoodieBay** is an Indian multinational restaurant aggregator that operates a digital platform for food ordering, delivery services, and customer reviews and ratings across many cities in various countries. The platform offers information on partner restaurants, including menus, online ordering and table booking options, the average cost for two, and user reviews and ratings.
- **FoodsAnalytics** is an independent data analytics consulting service, specialised in analytics and consulting services for restaurant reviews and food delivery platforms. They provide insights into customer experience, health, food quality, and safety.

You've been approached by FoodsAnalytics and provided with a dataset from FoodieBay, one of their clients. Ms Sue Keller, Head of Data Analytics at FoodsAnalytics, is eager to explore the FoodieBay dataset, a subset of data generated from the FoodieBay platform, concerning their partner restaurants in India and user ratings. She wants to better understand the factors that influence restaurant ratings (the *rate* column in the dataset) for two analytical tasks: data insights and machine learning opportunities, as follows.

Regarding the former, the client has requested the following:

1. How many restaurants offer table booking, and what impact does it have on ratings (*rate*)?
2. How many restaurants offer online ordering, and what impact does it have on ratings (*rate*)?
3. What is the effect of the average cost for two (in INR) and average customer review ranking (*ave_review_ranking* in the dataset) on overall ratings (*rate*)?

4. How do ratings (*rate*) and average cost for two vary among the restaurant types?
5. Any additional insights about the restaurants, cities, and cuisines; for example, but not necessarily, what are the best- or worst- performing restaurants, or which city has the greatest number of restaurants?

Regarding the latter, the client is interested in machine learning opportunities to predict restaurant ratings (*rate*) or categorise groups of similar restaurants. At this stage, one machine learning model is required. Future opportunities can be suggested.

Datasets provided:

- *FoodieBay.csv*

Data description

- *FoodieBay_metadata.csv*

You are required to explore the first dataset, *FoodieBay.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 Ms. Sue Keller, Head of Data Analytics at *FoodsAnalytics*. 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 recommend a machine learning 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 Mr. Anil Kumar, Head of Business Development, *FoodieBay*. The report should include your response to the client's five (5) questions, the proposed machine learning model, 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.

The source of the dataset will be provided after the assignment has been returned. The dataset has been pre-processed for the assessment purpose. FoodieBay is name used for the case study.

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).
- 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 Ms Sue Keller, Head of Data Analytics.

- **The second consultancy report (Part B)**, for Mr Anil Kumar, Head of Business Development at FoodieBay, should present responses to the specific requests about data and a machine learning model.
- Format and present your report professionally. Two sample report templates are provided under Assessment Resources. You can use your own template.
- 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 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) framework¹.
 2. Insights from Exploratory Data Analysis (EDA):
 - Answers to the Client's five (5) questions.
 - Additional insights, such as comments on data quality or observations beyond the client's questions.
 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)

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

- 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.

Learning Outcomes

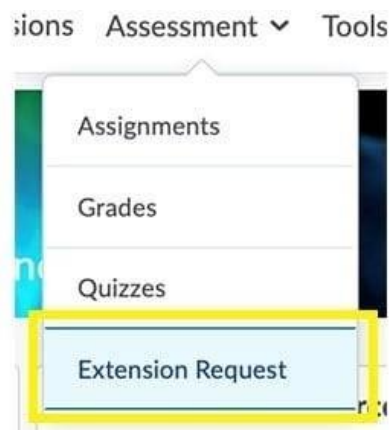
The learning outcomes that are aligned to this assessment task are:

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

The submission must include two files:

- Two (2) report documents. 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



Applications for extensions after 12 noon on the submission date require University level [special consideration](#) and these applications 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.