BIRMINGHAM CITY
University

Coursework Assignment Brief
Postgraduate

Academic Year 2021-22

Module Title:	Machine Learning
Module Code:	CMP7288
Assessment Title:	Coursework
Assessment Identifier:	Coursework **CWRK001** Weighting: 100
School:	School of Computing and Digital Technology
Module Co-ordinator:	Besher Alhalabi
Hand in deadline date:	12/01/2023
Return of Feedback date and format	20 working days from date of submission (see Moodle for details)
Re-assessment hand in deadline date:	12pm Mid-day on Monday 26th July 2023
Support available for students required to submit a re-assessment:	Formative feedback opportunities and timetabled support sessions will be available after the hand in date.
NOTE:	At the first assessment attempt, the full range of marks is available. At the re-assessment attempt the mark is capped and the maximum mark that can be achieved is 50%.
Assessment Summary	**In this assessment the student will** apply modern machine learning techniques to solve real-world analytics problems. Each student will submit a comprehensive report detailing the proposed work, results and conclusions. The report will be submitted as **one** deliverable in the form of a written report. The standard of academic writing should be excellent. (Maximum words: **4000 words**, excluding tables, figures and references).

IMPORTANT STATEMENTS

Standard Postgraduate Regulations

Your studies will be governed by the BCU Academic Regulations on Assessment, Progression and Awards. Copies of regulations can be found at   https://icity.bcu.ac.uk/Academic-Services/Information-for-Students/Academic-Regulations-2018-19

For courses accredited by professional bodies such as the IET (Institution of Engineering and Technology) there are some exemptions from the standard regulations and these are detailed in your Programme Handbook

Cheating and Plagiarism

Both cheating and plagiarism are totally unacceptable and the University maintains a strict policy against them. It is YOUR responsibility to be aware of this policy and to act accordingly. Please refer to the Academic Registry Guidance at https://icity.bcu.ac.uk/Academic-Registry/Information-for-Students/Assessment/Avoiding-Allegations-of-Cheating

The basic principles are:
- Don't pass off anyone else's work as your own, including work from "essay banks". This is plagiarism and is viewed extremely seriously by the University.
- Don't submit a piece of work in whole or in part that has already been submitted for assessment elsewhere. This is called duplication and, like plagiarism, is viewed extremely seriously by the University.
- Always acknowledge all of the sources that you have used in your coursework assignment or project.
- If you are using the exact words of another person, always put them in quotation marks.
- Check that you know whether the coursework is to be produced individually or whether you can work with others.
- If you are doing group work, be sure about what you are supposed to do on your own.
- Never make up or falsify data to prove your point.
- Never allow others to copy your work.
- Never lend disks, memory sticks or copies of your coursework to any other student in the University; this may lead you being accused of collusion.

By submitting coursework, either physically or electronically, you are confirming that it is your own work (or, in the case of a group submission, that it is the result of joint work undertaken by members of the group that you represent) and that you have read and understand the University's guidance on plagiarism and cheating*.*

You should be aware that coursework may be submitted to an electronic detection system in order to help ascertain if any plagiarised material is present. You may check your own work prior to submission using Turnitin at the Formative Moodle Site. If you have queries about what constitutes plagiarism, please speak to your module tutor or the Centre for Academic Success.

Electronic Submission of Work

It is your responsibility to ensure that work submitted in electronic format can be opened on a faculty computer and to check that any electronic submissions have been successfully uploaded. If it cannot be opened it will not be marked. Any required file formats will be specified in the assignment brief and failure to comply with these submission requirements will result in work not being marked. You must retain a copy of all electronic work you have submitted and re-submit if requested.

Learning Outcomes to be Assessed:
1. Critically evaluate machine learning techniques for suitability to a given industry problem.
2. Implement machine learning techniques to formulate and solve real-life data-based problems.
3. Professionally report findings from a data analysis project to a technical audience.
4. Evaluate learning algorithms and utilise model selection effectively.

 Assessment Details:

Title: *Individual Machine Learning Report*

Type: *Coursework*

Style: Report

Rationale:
The assessment is an individual task that requires each student to use modern machine learning techniques and tools to identify, analyse and solve a domain-specific problem.
It will help the students to have the required technical expertise that are required by data scientists to construct and evaluate the effectiveness of AI systems.

Description:
The assessment is in the form of an individual report. The students have to follow the below steps in order to complete the assessment successfully:

*   Each Student will select a domain of interest (e.g. Healthcare, Education, Sport, Finance, Meteorology, etc.)
    *   In consultation with the course tutor, each student has to find a data set that is related to the chosen domain. The chosen data set should be publicly available (eg, Kaggle, UCI repository, etc.) with 1,000 rows/instances and 5 columns/attributes (at least).
    *   In consultation with the course tutor. each student has to identify three different machine learning problems (where at least one will be supervised & at least one will be unsupervised) that are related to the dataset.
    *   Each student will critically evaluate the chosen machine learning problems and propose solutions for them using supervised and unsupervised techniques. The solution should include data prepressing, data exploration, model building and model evaluation. Finally, the student should discuss the results using appropriate methods as discussed in the module.
    *   Each student has to submit a report of max 4,000 words (excluding code, tables, references and figures) in DOCX/PDF format.
    *   Ideally, the reports should include the following sections:
        o Cover page (report title, student ID and name).
        o Domain description.
        o Problems definition.
        o Data set description.

- o Data set exploration (preprocessing and wrangling).
- o Experiments (three machine learning techniques) and evaluations.
- o Analysis and Results.
- o Conclusions.
- o References.
- o Appendixes: (codes, charts, etc.)

Additional information:
- References (as per Harvard Referencing Style)
- Students must use Harvard Referencing Style https://icity.bcu.ac.uk/Library-and-Learning-Resources/Referencing/harvard-referencing).
- Report format for the submissions is **DOCX**.
- A video course on **Impactful Writing** is available at: https://www.youtube.com/playlist?list=PLU98PJIZ0JfIj_msDROdMU85wdV72ihks

For advice on writing style, referencing and academic skills, please make use of the Centre for Academic Success: https://icity.bcu.ac.uk/celt/centre-for-academic-success

For advice on writing style, referencing and academic skills, please make use of the Centre for Academic Success: https://icity.bcu.ac.uk/celt/centre-for-academic-success

Workload:

A typical student will spend up to 90 hours to complete this work. The word count limit is the equivalent of 4000 words.

Transferable skills:

- Problem solving
- Programming skills
- Analytical skills
- Time management
- Project management
- Verbal and written communication skills

Marking Criteria:
Marking criteria should be used as a general guide to the expectations at different grades. Ensure it is clear how the criteria will be used to derive the final grade for the component. Ensure it is clear how the component grade goes towards the final mod

Table of Assessment Criteria and Associated Grading Criteria

Assessment Criteria	1. Critically evaluate machine learning techniques for suitability to a given industry problem.	2. Implement machine learning techniques to formulate and solve real-life data-based problems	3. Professionally report findings from a data analysis project to a technical audience	4. Evaluate learning algorithms and utilise model selection effectively.
Weighting:	**25%**	**25%**	**25%**	**25%**
Grading Criteria **0 – 29%** **F**	Report shows little or no understanding of machine learning techniques. Missing description of data set used. Missing of machine learning problems definitions. inaccurate and incomplete comparison of machine learning methods.	No data-pre-processing. Inaccurate and incomplete application for the machine learning techniques. No analysis of the application of the machine learning methods is provided. Less than three machine learning techniques have been mentioned with no or limited description.	A superficial interpretation of the results. No data exploration. No clear solution provided for the underlying machine learning problems. Analysis and Results sections are not Provided or not accurate.	No evaluation methods. Incomplete numerical validation of machine learning techniques. No interpretation of the evaluation results.
30 – 39% **E**	Report shows poor understanding of machine learning techniques. Vague or no description of the data set used. Few techniques have been	Inadequate data-pre- processing. Vague application of a few machine learning techniques. basic analysis of the application of the machine learning methods is provided.	Poor description of the results. data is poorly explored with a limited number of charts. Inaccurate or vague solution provided for the underlying machine learning problem.	Inaccurate numerical validation of machine learning techniques. Inaccurate interpretation of the evaluation results.

	mentioned with non-detailed description. Unclear comparison of machine learning methods.	Less than three machine learning techniques have been mentioned with poor description.	Poor conclusions and poor references provided. Poor analysis, Inaccurate results.	
40 – 49% D	Report shows basic understanding of machine learning techniques. Limited description of the data set used. Few techniques have been mentioned with basic description. Incomplete comparison of machine learning methods. There may be some inaccuracies.	Incomplete data-pre- processing. Limited application of machine learning techniques. There may be some inaccuracies. Limited analysis of the application of the machine learning methods is provided. Less than three machine learning techniques have been mentioned with basic description.	Basic description of the results with some errors. Basic data exploration with a limited number of charts. The charts interpretation is not complete or not accurate. A non-efficient solution provided for the underlying machine learning problem. Vague conclusions and poor references provided. Basic analysis, Inaccurate results, may contain some errors.	Basic numerical validation of machine learning techniques, with some accurate results. Basic interpretation of the evaluation results, with some accurate interpretations.
50 – 59% C	Report shows general understanding of machine learning techniques. Reasonable description of the data set used.	Valid data-pre-processing but it may be inaccurate. A number of appropriate machine learning techniques have been applied. There may be few inaccuracies.	Results described but it may be inaccurate. data exploration is completed with different types of charts. The charts interpretation is basic and may be inaccurate.	More detailed numerical validation of machine learning techniques, with some accurate results. More detailed interpretation of the evaluation

	Relevant techniques have been described with justification of why it has been used. A comparison of machine learning methods has been provided but improvable.	Analysis of the information derived from the application of machine learning techniques is limited. Three machine learning techniques have been mentioned with basic description.	A non-efficient but a valid solution provided for the underlying machine learning problem. Simple conclusions are drawn with Incomplete references list. Basic analysis, results are provided.	results, with some accurate interpretations.
60 – 69% **B**	Report shows good understanding of machine learning techniques. Fair description data set used. Machine learning problems addressed to a good standard and impact level. A comparison of machine learning methods has been provided.	Good data-pre-processing but improvable. A good range of machine learning techniques has been correctly applied. A reasonable analysis of the information derived from the application of machine learning techniques. three machine learning techniques have been mentioned with good description and interpretations.	Sufficient interpretation of the results. data exploration is completed with different types of charts (statistical/graphical). The charts interpretation is good. Valid solution is provided for the underlying machine learning problems. Conclusions are briefly expressed, References are provided but improvable. fair analysis, results are provided.	More detailed numerical validation of machine learning techniques but it needs to be more strongly aligned with the analysis. More detailed interpretation of the evaluation results, but it's not strongly linked to the numerical evaluation.
70 – 79% **A**	Report shows good understanding of machine learning techniques.	Good and Complete data-pre- processing. An accurate application of a wide range of	Good interpretation of the experiment and results. Data exploration is	Good numerical validation of machine learning techniques and linked directly to the analysis.

	Good description of the data set used. Machine learning problems addressed to a very good standard with significant impact made. A comparison of machine learning methods has been provided, including their strengths and limitations.	machine learning techniques. Good analysis of the information derived from the application of machine learning techniques but there may be scope for more details. Three machine learning techniques have been mentioned with good description and interpretations.	completed with different types of charts (statistical/graphical). The charts interpretation is good. Good solution is provided for the underlying machine learning problems. Appropriate conclusions are drawn and well expressed. References are provided with correct citations. Very good analysis, results are provided.	Good interpretation of the evaluation results, directly linked to the numerical evaluation
80 – 89% A+	Report shows Very good understanding of machine learning techniques. Very good and detailed description of the data set used. Relevant and effective techniques have been well described with justification of why it has been used. A comparison of machine learning methods has been provided,	Very good data-pre- processing. Very good and efficient application of a wider range of machine learning techniques. Excellent analysis of the information derived from the application of machine learning techniques. Three machine learning techniques have been mentioned with very good description and interpretations.	Very good interpretation of the experiment and results. Data exploration is completed with different types of charts (statistical/graphical). The charts interpretation is very good. Very good solution is provided for the underlying machine learning problems. Very good conclusions are drawn and well expressed. References are provided with	Very good numerical validation of machine learning techniques that is strongly aligned with the analysis. Very good interpretation of the evaluation results, directly linked to the numerical evaluation

	including their strengths and limitations.		excellent and complete citations.	
90 – 100% A*	Report shows excellent understanding of machine learning techniques. Excellent and detailed description of the data set used. Relevant and effective techniques have been well described with justification of why it has been used. A comparison of machine learning methods has been provided, including their strengths and limitations.	Excellent data-pre- processing. Excellent and efficient application of a wider range of machine learning techniques. Excellent analysis of the information derived from the application of machine learning techniques. Three machine learning techniques have been mentioned with excellent description and interpretations.	Excellent interpretation of the experiment and results. Data exploration is completed with different types of charts (statistical/graphical). The charts interpretation is excellent Excellent solution is provided for the underlying machine learning problems. Excellent conclusions are drawn and well expressed. References are provided with excellent and complete citations.	Excellent numerical validation of machine learning techniques that is strongly aligned with the analysis. Excellent interpretation of the evaluation results, directly linked to the numerical evaluation

Submission Details:

Format: *Upload MS Word file / PDF to Moodle*

Regulations:

If you submit an assessment late at the first attempt then you will be subject to one of the following penalties:

- if the submission is made **between 1 and 24 hours** after the published deadline the original mark awarded will be reduced by **5%**. For example, a mark of 60% will be reduced by 3% so that the mark that the student will receive is 57%. ;
- if the submission is made between **24 hours** and **one week (5 working days)** after the published deadline the original mark awarded will be reduced by 10%. For example, a mark of 60% will be reduced by 6% so that the mark the student will receive is 54%.
- **if the submission is made after 5 days following the deadline, your work will be deemed as a fail and returned to you unmarked.**

The reduction in the mark will not be applied in the following two cases:

- the mark is below the pass mark for the assessment. In this case the mark achieved by the student will stand

- where a deduction will reduce the mark from a pass to a fail. In this case the mark awarded will be the threshold (i.e. 50%)

Please note:
- **If you submit a re-assessment late then it will be deemed as a fail and returned to you unmarked.**

Feedback:

Feedback for the deliverable will be provided via Moodle. The students are also strongly encouraged to discuss their draft work with tutors in lecture sessions, whenever time permits.

Marks and Feedback on your work will normally be provided within 20 working days of its submission deadline.

Where to get help:

Students can get additional support from the library support for searching for information and finding academic sources. See their iCity page for more information: http://libanswers.bcu.ac.uk/

The Centre for Academic Success offers 1:1 advice and feedback on academic writing, referencing, study skills and maths/statistics/computing. See their iCity page for more information: https://icity.bcu.ac.uk/celt/centre-for-academic-success

Link to My Assignment Planner tool: http://library.bcu.ac.uk/MAP2/freecalc-mail/

Fit to Submit:

Are you ready to submit your assignment – review this assignment brief and consider whether you have met the criteria. Use any checklists provided to ensure that you have done everything needed.

Staff to include any checklists here, or alternatively a link to ones found on Moodle.