IN-COURSE ASSESSMENT (ICA) SPECIFICATION

Module Title: **Applied Machine Learning**	Module Leader: **Dr Ala Alkafri**
	Module Code: **CIS3036-N**
Assignment Title: **Part One (50%) Machine learning application** **Part Two (50%) Conference-style paper (approximately 2,000 words).**	Deadline Date: **7th Jan 2025**
	Deadline Time: 4:00PM
	Submission Method: Online (Blackboard) ☑ Middlesbrough Tower ☐

Online Submission Notes:

- Please carefully follow the instructions given in this Assignment Specification.
- When Extenuating Circumstances (e.g. extension) has been granted, a fully completed and signed Extenuating Circumstances form must be emailed to scedt-assessments@tees.ac.uk or submitted to the School Reception.

Central Assignments Office (Middlesbrough Tower M2.08) Notes:

- All work (including media) needs to be secured in a plastic envelope or a folder and clearly marked with the student name, number and module title.
- An Assignment Front Sheet should be fully completed before the work is submitted.
- When Extenuating Circumstances (e.g. extension) has been granted, a fully completed and signed Extenuating Circumstances form must be emailed to scedt-assessments@tees.ac.uk or submitted to the School Reception.

FULL DETAILS OF THE ASSIGNMENT ARE ATTACHED
INCLUDING MARKING & GRADING CRITERIA

Full details of the assignment

Assignment Instructions Specification

This document describes the assessment for the module. It is an in-course assignment (ICA) composed of two parts. **(strictly working without help from AI)**

Assessment for **Applied Machine Learning** requires you to develop machine learning applications and make predictions about unseen data. The summative assessment for this module is via **in-course assessment (100%)** which will evaluate all learning outcomes (see below).

The assessment will emulate the "shared task" framework appearing in several machine learning venues, in which participants are supplied with a task description and annotated data and must develop a machine learning solution that makes predictions about an unannotated data set.

1. **Component 1 (Group Work)** Machine learning application and its predictions on the labelled/unlabelled dataset (50%). Students will work on selection, implementation and experimental validation of machine learning techniques on a challenging benchmark dataset solving a real-world problem. Project demonstration of your prediction model - a recorded demo/walkthrough video (around 2-3 minutes). **[50 points].**

2. **Component 2 (Individual Work)** (50%) consists of a conference-style paper of approximately 2,000 words that reports on the development of their system (50%) **[50 points].**

Task Description

Problems in machine learning vary from one domain to another. In this coursework, you will select a dataset related to a real-world problem that best suits your area of interest. There are abundant of websites that provide publicly available datasets. A categorised list of datasets from GitHub can be found at https://github.com/caesar0301/awesome-public-datasets.The UCI Machine Learning Repository at https://archive.ics.uci.edu/ml/index.php is another long-standing source of benchmark datasets for data mining and machine learning research. Kaggle (https://www.kaggle.com/datasets) has interesting real-world problems and datasets.

You can select a dataset from the above sources, or another one that is available online. The dataset should be publicly available. The chosen dataset should have a minimum of

1,000 instances (rows) and a minimum of 5 attributes (columns). You have to complete the following stages in this assignment:

1. Define the problem for the selected data set and identify the machine learning algorithms that are applicable to this problem.
2. Data exploration and preparation: The nature of the dataset may dictate some data exploration and preparation that can help inform the solutions. For example, higher dimensional datasets (those with too many attributes/columns) may require applying a data reduction method like Principal Component Analysis (PCA).
3. Propose solutions: In this step, each student in the group should propose one machine learning algorithm that applicable to the selected dataset/problem.
4. Design, implementation, modelling and evaluation: design, model and implement the proposed solutions and critically evaluate the solutions. Use appropriate visualisation for the results.
5. Reflect on professional, ethical, and legal issues in relation to the problem and the data set.

Please note that for the component 1 (group work), one dataset should be selected and used by the whole group members, and then everyone in the group will use the same dataset for the implementation and evaluation of their selected algorithm (which is part of component 2). As part of the component 1, the dataset should be prepared (pre-processing, etc.) before it can be used in your model.

Every student in the group should select one algorithm that best fits the selected dataset, and this should be done after a group discussion and coordination between all the members of the group. The implementation and experimental evaluation of the algorithm is an individual work (part of the component 2), but the algorithms selection is part of the group work (component 1).

For more details about the two components and what to hand-in, please see the following section.

Tools to be used: ***You can use any tools that can help in delivering your project such as R/R Studio, Excel, Ms Word, and PDF.***

Deliverable

Component 1 (Group Work) Deliverable – Contribute 50% of the Module Mark

What to Hand-In

This component you will work as a group and your work will be evaluated based on your contribution (see individual mark calculation section below for more details) to the project.

Submission method is online on Blackboard. You are required to submit the following files in one zipped file appropriately labelled (student_ID_surname_firstname):

- A **file** in a **pdf format** via Blackboard that **includes all source code** and **screenshots** from your experiments appropriately labelled and commented (student_ID_surname_firstname).

- **Copy of the source code** and **screenshots** from your experiments appropriately labelled and commented (student_ID_surname_firstname) via Blackboard.

- Group demonstration of your prediction model - a **recorded demo/walkthrough video** (around 2-3 minutes). Please make sure you save your RStudio working environment for the demonstrations.

- **Individual mark calculation form** (see below for more details about the form) submitted by each single member of the group.

The code and experiments will be assessed on:

- Appropriateness of machine learning algorithm selected for the given task.

- Quality of software architecture and implementation.

- Quantitative performance of application.

For Component 1, all the files/documents previously listed must be uploaded to the relevant sections on Blackboard by each student. Each member of a group must submit the same documents. Please do not alter your own copy. Only one copy per team will be marked and it will be assumed that all copies are identical.

Component 2 (Individual Work) – 50% of the Module Mark

What to Hand-In

(Individual Work)

Everyone in the group should do this component individually using their selected algorithm in component 1.

- A case study **report** maximum of **2,000 words** that documents the process of the entire case study, including data set, problem, data preparation and exploration, selected algorithm, critical evaluation and justification of the algorithm and findings.

Submission method is online on Blackboard. You are required to submit a file in a pdf format via Turnitin on Blackboard

The hand-in is electronically via Blackboard, all deliverables shall be labelled with project name, your student name and university number.

The report will be assessed on:

- understanding of machine learning task

- review of relevant literature
- development methodology
- justification of design decisions
- consideration of professional, ethical, and legal issues

The report could broadly include the following sections:

- Abstract
- Introduction (introduce the problem and its significance, write short literature review of related work)
- Data exploration and features selection
- Experiments
- Results
- Discussion, Conclusions and Future Work
- References

These are generic section titles, which you may adapt appropriately to the application/problem that is investigated. You may include sections describing modifications of algorithms or developments that are novel and specific to your work.

Learning Outcomes

Personal and Transferable Skills Development

1. Decide suitable machine learning algorithms for a given task and elaborate it in a final work report.

Research, Knowledge and Cognitive Skills

2. Demonstrate a deep understanding of machine learning concepts.

3. Analyse and discuss the empirical results of the selected machine learning algorithms and justify the performance.

4. Work in a team to design appropriate data input and apply machine learning techniques to real-life problems.

Professional Skills, Values and Behaviours

5. Analyse, design, and construct a solution to a project by exploring suitable machine learning techniques and evaluating performance at managing and executing a development project.

6. Critically appraise the ethical, social, legal and security implications related to machine learning applications.

Component 1 (50%) assesses learning outcomes: 1, 2, 4, 5

Component 2 (50%) assesses learning outcomes: 3, 5, 6

Assessment Criteria

For the component 1 the students will work in a group and assessment will be group-based. Formative feedback will form a significant part of the learning and assessment strategy for this module with regular formative feedback meetings operating through the assessment period.

Component 1 – (Group Work) Machine learning application (50%):

Students will be assessed on:

- Experimentation/analysis of different machine learning algorithms and the appropriateness of the algorithm selected for the given task
- Performance evaluation of the selected algorithms
- Evidence of reflecting upon feedback and incorporating it into the final work
- Evaluation of the effective management and execution of the final solution.

Component 2 – (Individual Work) Conference-style report (50%):

Students will be assessed on:

- Methodological understanding of the machine learning task
- Development of the methodology and interpretation of results
- Evaluation of results from their practical work
- Justification of the selected algorithms

Your submission will be assessed according to the following criteria:

1. Machine Learning application **[50 points]**.

2. Report (conference-style paper) **[50 points].**

Below is a provisional indication of the criteria applied to determine points for each component.

Please note: Exceptionally, whilst points are allocated to specific parts, outstanding work in one area may be used to trade-off points against poorer work in another area.

Task 1 Assessment Criteria [50 points] *Source Code Documentation and Demo*	**100%**
Clear evidence of running the experiments with code that is excellently organised and commented. Machine learning algorithms selected are appropriate for the given task. Excellent quality of software architecture and implementation. Excellent quantitative performance of the application.	70 -100

Deep understanding shown. Excellent collaboration with others within the group has been successful, professional and effective.	
Very good evidence of running the experiments with code that is well-organised and commented. Machine learning algorithms selected are appropriate for the given task. Very good quality of software architecture and implementation. Very good quantitative performance of the application. Very good understanding. Very good Collaboration with others within the group.	60 - 69
Good evidence of running the experiments with code that is well-organised and commented. Machine learning algorithms selected are appropriate for the given task. Good quality of software architecture and implementation. Good quantitative performance of the application. Good understanding. Good Collaboration with others within the group.	50 - 59
Satisfactory evidence of running the experiments with code that is organised and commented. Machine learning algorithms selected are appropriate for the given task. Satisfactory quality of software architecture and implementation. Satisfactory quantitative performance of the application. Satisfactory understanding. Satisfactory collaboration with others within the group.	40 - 49
Little evidence of running the experiments with code that is not well-organised and commented. The machine learning algorithms selected are not appropriate for the given task. Poor quality of software architecture and implementation. Poor quantitative performance of the application. Poor understanding. The group did not collaborate well which has had an impact on the completion of the project.	0-39
N/A	**NS** NON-SUBMISSION

Task 2 Assessment Criteria [50 points]	**100%**
Excellent structure and immaculate presentation of the materials, detailing a comprehensive understanding of the field. There is clear, concise and well-articulated evidence of reflection and appraisal of the implemented solution against the given brief. The project has been planned and executed flawlessly to satisfy a complex brief demonstrating the ability to work autonomously and effectively. The final presented paper has been successful, professional and effective.	70-100

Well-structured and presented materials, detailing a good understanding of the field. There is evidence of reflection and appraisal of the implemented solution against the given brief. The project has been planned and executed well to satisfy a complex brief demonstrating the ability to work partially autonomously. The final paper has been successfully presented.	60- 69
The structure and presentation of the materials is satisfactory, detailing some understanding of the field. Reflection and appraisal of the implemented solution has been attempted. The project has been completed satisfactorily but planning and execution could have been better. Collaboration with others has be completed with mixed results and there is room for improvement. The final paper has been completed with mixed results.	50-59
Limited structure and presentation of the materials demonstrating a limited level of understanding of the field. There has been very limited reflection and appraisal of the implemented solution. The project has been poorly completed and the planning and execution could been much improved. Collaboration with others has be completed with a weak results and there is room for improvement. The final paper has been completed with weak results.	40-49
Poor structure and presentation of the materials demonstrating a low level of understanding of the field. There has been insufficient reflection and appraisal of the implemented solution. The project requirements has not been followed which has had an impact on the completion of the project. The project lacked the required level of planning and execution. The final paper has not been fully completed.	0 - 39
NA	**NS** NON-SUBMISSION

KBS

Core Technical Knowledge

1. Contemporary techniques for design, developing, testing, correcting, deploying and documenting software systems from specifications, using agreed standards and tools.
2. How teams work effectively to produce technology solutions.

Core Behavioural Skills

1. Fluent in written communications and able to articulate complex issues.
2. Able to give and receive feedback constructively and incorporate it into his/her own development and life-long learning.
3. Applies analytical and critical thinking skills to Technology Solutions development and to systematically analyse and apply structured problem solving techniques to complex systems and situations.

4. Able to put forward, demonstrate value and gain commitment to a moderately complex technology-oriented solution, demonstrating understanding of business need, using open questions and summarising skills and basic negotiating skills.
5. Ability to perform under pressure
6. A thorough approach to work
7. Logical thinking and creative approach to problem solving

Specialism Outcomes - Software Engineer

Skills

1. Create effective and secure software solutions using contemporary software development languages to deliver the full range of functional and non-functional requirements using relevant development methodologies.
2. Undertake analysis and design to create artefacts, such as use cases to produce robust software designs.
3. Produce high quality code with sound syntax in at least one language following best practices and standards.
4. Perform code reviews, debugging and refactoring to improve code quality and efficiency.
5. Test code to ensure that the functional and non-functional requirements have been met.
6. Deliver software solutions using industry standard build processes, and tools for configuration management, version control and software build, release and deployment into enterprise environments.

Technical Knowledge (knows and understands)

1. How to operate at all stages of the software development lifecycle.
2. How teams work effectively to develop software solutions embracing agile and other development approaches.
3. How to apply software analysis and design approaches.
4. How to interpret and implement a design, compliant with functional, non-functional and security requirements.
5. How to perform functional and unit testing.
6. How to use and apply the range of software tools used in software engineering.

Feedback

Feedback will be given for your portfolio entries, outlining any possible improvements or clarifications where necessary.

Individual Mark Calculation

You will receive an individual mark for component 1 which will be calculated using an individual weighting factor. You must as individuals submit the following table to Blackboard:

Component 1

Member name	**Individual Weighting Factor**
Member 1	
Member 2	
Member 3	
Member 4	
Total	**13**

Divide 13 marks between the members of the group including yourself. **Please only use integers.** You do not need to do this in consultation with the rest of the group. Your individual mark will be calculated by determining your peer assessment factor and multiplying it by your group mark, where Peer Assessment Factor = (individual total weighting factor) / 13

E.g., If component 1 mark = 60% and Member 1 has an individual total of 12 then

Member 1 = 60 * (12/13) = 55%

The tutor will moderate the marks to ensure fairness and reserves the right to overrule the weightings in exceptional circumstances based on their observations within the class and during the presentations.

Your marks from both components will be combined as follows to give an overall module mark:

E.g., component 1 = 70% , component 2 = 78.5% gives an overall mark of 70*0.5+78.5*0.5 = 74.25%