School of Design and Informatics

Session 2022-2023

CMP310 Fundamentals of Machine Learning
Re-assessment Unit 1: Project (80%)

Module leader: Salma ElSayed

Submission date: 27th July 2023

Feedback return date: 17th August 2023

The student should refer to the original brief (below) and feedback received from the module tutors to make good any deficiencies in the original submissions. Submission is through the "Resit SUBMIT Unit 1" link on MyLearningSpace.

Learning outcomes assessed by this assignment:

- LO1 Understand the different learning paradigms and their applicability and challenges.
- LO2 Identify the use of various intelligent algorithms in real-world environments.
- LO3 Design and justify a suitable Machine Learning (ML) solution for a data-driven problem in a given context.
- LO4 Analyse and critically evaluate ML technologies and techniques for specific purposes.

Please note that you cannot pass the module without evidencing all learning outcomes for this unit.

Introduction

For CMP301, Unit 1 consists of a project and a report. You are required to build a machine learning solution informed by your understanding of the lectures and labs content. This is a **summative** assessment of 80% of your final grade. This document describes what is expected.

The Assessment

You are required to design and build a machine learning solution in a given context (e.g., security, computer vision, human-computer interactions, etc.). You are provided with some datasets to choose from, and it is up to you to decide what information the data can provide, what learning type is suitable (supervised, unsupervised), and what ML task is appropriate (regression, binary or multiclass classification, clustering). The datasets can be found on MyLearningSpace under the assessment folder.

Deliverables

1. Perform Exploratory Data Analysis (EDA) of the selected dataset.
2. Design any pre-processing steps that may be needed to prepare the data.
3. Investigate ML models suitable for the task. Select one and give rationale for the choice.
4. Define features to describe the inputs. Not all ML problems/models require a separate feature extraction step. You will have to decide and justify your decision.
5. Design and build an application that will implement the ML pipeline you designed.
6. Conduct a performance analysis of the system with appropriate measures and visualisations. This should include receiver operating characteristic (ROC) and precision recall curves.
7. Write a report detailing how you designed and developed the application, and how it was tested and assessed. The report should cover all stages of the data pipeline consisting of data ingestion/pre-processing, modelling, analysis, and communication of results.

Grading the Assessment

The Application (30% weighting)

The application should be efficiently coded and properly documented, following good programming practices. You will create a video, no more than 10 minutes long, demonstrating the application with an appropriate commentary detailing how you created and trained the model and demonstrating test cases with different inputs and corresponding outputs. A guide will be provided for you on what is expected in the video demo.

The report (70% weighting)

The report should fully describe the design process for the application, giving specific details of the steps you followed. You should give full rationale for the choice of features (if any) and techniques. For example, why you chose them and why they are particularly suited for your application. Your report should include details of testing, an explanation of the performance, and quantitative performance measures as appropriate (e.g., system accuracy) with relevant plots.

There should be a full set of test data and results as well as suitable conclusions drawn and discussion on the advantages and disadvantages of the adopted approach. Any references or images used MUST be properly cited. Refer to Abertay's referencing guide on how to cite different material. A template is provided for you to document the above deliverables.

Submission

Please submit your solution via MyLearningSpace as follows (separate files):

- Source files (.py or .ipynb)
- OneDrive link to video demo (adjust permission so staff can access the file)
- Report as .pdf

For the submission rules, refer to the academic regulations.

Guidance on submitting via MyLearningSpace is available at digital skills guide for students. Please contact the Support Enquiry Zone sez@abertay.ac.uk if you have any problems with the submission process.

Late Submission

Work submitted up to 5 days following the submission deadline will be accepted in the absence of valid cause but will be subject to a penalty of 1 grade (-0.5 GPA) for each day it is late. Work submitted after 5 days without valid cause will be treated as a non-submission and will be awarded a NS grade.

Grading Criteria

Your submission is used to demonstrate your achievements to the moderator and external examiner if any. You will receive feedback on your submission via MyLearningSpace. Module tutor will assess your submission against the grading criteria on the following page.

	Introduction (5%)	Data Specs (10%)	Methodology (15%)	Results (10%)	Discussion and Conclusions (20%)	References (5%)	Structure and Style (5%)	Application (30%)
A	Excellent introduction which gives an appropriate overview of the project.	Excellent description of data specs and rationale for datasets choice, including proper citation and sample pictures.	Excellent description of the methods used including a complete explanation and rationale for the chosen techniques.	Excellent results, clearly tabulated and relevant.	Excellent conclusions with a full analysis and summary of the project.	A good number of excellent references properly cited in Cite Them Right Harvard style.	Excellent formatting and structure, with clear flow of information between sections. Excellent use of language with excellent grammar and spelling.	Excellent application which clearly shows how the programme is operating, efficiently coded with excellent documentation.
B	Very good introduction which gives an almost perfect overview of the project.	Very good description of data specs with a mostly complete rationale for datasets choice, including proper citation and sample pictures.	Very good description of the methods used including a mostly complete explanation and rationale for the chosen techniques.	Very good results, very well tabulated and mostly relevant.	Very good conclusions with a very good analysis and summary of the project.	A number of very good references mostly properly cited in Cite Them Right Harvard style.	Very good formatting and structure, with good flow of information between sections. Very good use of language with very good grammar and spelling.	Very good application which clearly shows how the programme is operating, mostly efficiently coded with very good documentation.
C	A good introduction which gives a more than adequate overview of the project.	Good description of data specs including some rationale for datasets choice, citation, and sample pictures.	A good description of the methods used including an explanation and rationale for the chosen techniques.	Good results, reasonably well tabulated and relevant.	Good conclusions with a good analysis and summary of the project.	A reasonable number of good references mostly cited in Cite Them Right Harvard style.	Good formatting and structure with reasonable flow of information between sections.	Good application which mostly shows how the programme is operating, well coded with good documentation.
D	A brief introduction which gives an adequate overview of the project.	An adequate description of data specs with an incomplete rationale for datasets choice, citation, and sample pictures.	An adequate description of some of the methods used including an incomplete explanation and rationale for the techniques and features chosen.	Adequate results, with some tabulation and relevance.	Adequate conclusions with an adequate analysis and summary of the project.	A few adequate references properly cited in Cite Them Right Harvard style.	Satisfactory formatting and structure but lacking coherence of sections.	Adequate application which just shows how the programme is operating, adequately coded with some documentation.

MF	A very brief introduction which just fails to give an adequate overview of the project.	A not quite adequate description of data specs with incomplete rationale for datasets choice, no citations or sample pictures.	A not quite adequate description of the methods used with an incomplete explanation and rationale for the chosen techniques.	A few results, with little tabulation and relevance.	Inadequate conclusions with inadequate analysis and summary of the project.	Some references inadequately cited.	Some structure, readable in places with inadequate standards of spelling and grammar.	Inadequate application which does not show how the programme is operating clearly enough, inadequately coded with insufficient documentation.
F	Too brief and completely fails to give a sufficient overview of the project.	Totally inadequate description of data specs with very little explanation and no rationale for datasets choice, no citations or sample pictures.	Totally inadequate description of the methods used with very little explanation and rationale for the chosen techniques.	Inadequate results, not properly tabulated and barely relevant.	Scant conclusions with barely any analysis or summary of the project.	Barely any references cited, or poorly cited.	Poorly structured, not easy read with poor spelling and grammar. Unacceptable, messy, or incomplete report.	Very poor application which barely shows how the programme is operating, poorly coded with little documentation.
NS	There is no submission, or the submission contains no relevant material.							