# Assessment Brief

|  |  |  |
| --- | --- | --- |
| Module Leader: Rizwana Rashad | | Level: 6 |
| Module Name: Machine Learning Algorithms and Heuristics | | Module Code: 55-602446-BF-20245 |
| Assignment Title: Machine Learning System | | |
| Individual Assignment | Weighting: 100% | Magnitude: Equivalent 2000 words |
| Submission date/time:  29th April 2025, 3PM  If you have an extension then  check your personal submission  deadline on My Student Record. | Blackboard submission: Y Turnitin submission: N | Format: 12 ±10% minute video presentation with supporting documentation |
| Planned feedback date:  21st May 2025  If your submission date is different to the standard submission date for the task, for example, due to an approved extension, we will endeavour to return your feedback within 3 weeks from the extended deadline. | Mode of feedback: Blackboard | In-module retrieval available: No |
|  | | |
| **Module Learning Outcomes**   * Understand and implement commonly used machine learning algorithms. * Apply machine learning techniques to develop applications which answer complex questions within a variety of contexts including business, industry and the social media domain. * Analyse and evaluate techniques for prediction, ranking and recommending using technical and contextual reasoning. | | |

**Assessment Brief**

**Introduction**

This assignment aims to assess your practical understanding of machine learning algorithms by implementing systems (using supervised learning, unsupervised learning and image processing), which can answer complex questions and make intelligent decisions, in the context of the dataset provided and a copyright free image dataset of your choice from Kaggle (Please ensure that you do not use a dataset used in lectures or tutorials).

This is an individual project. You are expected to demonstrate your end-to-end machine learning system design and development skills, including data exploration, feature analysis, machine learning, and performance evaluation.

You need to prepare the following content for the assessment:

1. A 12 ±10% minutes video presentation demonstrating your machine learning systems. Ensure that you consider both supervised and unsupervised learning.

* **Supervised Learning** – Use the complete dataset provided.
* **Unsupervised Learning** – Delete the label column, to ensure that the data is not labelled.
* **Image Processing** – Use a copyright free image dataset of your choice from Kaggle.

1. Supporting documentation for the video including presentation slides, code, and relevant screenshots.

**Individual Project**

Your machine learning project should include the following components:

For the Dataset Provided:

1. **Data Exploration**

You need to understand your data first. Your data may require cleaning and pre-processing steps. It will define which algorithms you can choose, and it will ultimately define the performance of your supervised and unsupervised learning models. During the project design and development, you should go through the following steps and try to get as many insights as possible about the data, and for supervised learning its relation to the target variable:

* Analyse the data distribution.
* For supervised learning analyse how the target variable is influenced by the features.
* Analyse the difficulty of your prediction task.

Please raise any ethical concerns with the tutor.

1. **Feature analysis**

Define a feature space based on the data samples. Make use of data visualisation tools to find correlated dimensions. You can also measure and analyse features and target dependencies for regression or classification tasks.

1. **Machine Learning**

Design and develop a suitable learning model using machine learning tools such as Scikit-Learn. Depending on the algorithms (select at least two for supervised learning and two for unsupervised learning), you may need to pre-process the data such as "scaling" and "de-noising". You can use cross-validation approaches to estimate the performance. You also need to keep the balance between under-fitting and over-fitting through adjusting the hyperparameters of the chosen machine learning algorithms.

**D. Evaluation**

The focus will be the accuracy performance of the developed system. You need to fine-tune your models to improve system accuracy performance. You are expected to compare at least two different algorithms, for each type of learning, in the project. During the system evaluation, you need to analyse and discuss the performance by using at least 2 different scientific evaluation approaches such as ROC curves and confusion matrices.

For the copyright image dataset of your choice from Kaggle (Please get image dataset approved from your tutor before working with it):

1. **Data Exploration**

Pre-process the dataset selected.

1. **Machine Learning**

Train a model for classification.

1. **Evaluate**

Measure accuracy and any other metrics of your choice.

Use the Python programming language and its machine learning packages, such as “Scikit-learn” for your project development.

**Your video and supporting documentation should be submitted electronically, through the module's Blackboard site as a single ZIP file, before the deadline (29th April 2025, at 15:00).**

# Assessment Criteria

|  | **FAIL** (insufficient) | | | | **THIRD** (sufficient) | | | **LOWER SECOND** (good) | | | **UPPER SECOND** (very good) | | | **FIRST**  (excellent) | | | **EXCEPTIONAL FIRST** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Zero | Low  Fail | Mid Fail | Marginal Fail | Low  3rd | Mid  3rd | High 3rd | Low  2.2 | Mid  2.2 | High 2.2 | Low  2.1 | Mid 2.1 | High 2.1 | Low  1st | Mid  1st | High  1st | Exceptional 1st |
| Criteria and weighting | **<19** | | **20-39** | | **40-49** | | | **50-59** | | | **60–69** | | | **70-92** | | | **93+** |
| **Video and Supporting Documentation:**  **Technical Difficulty**  40% | Highly insufficient or no evidence of understanding of machine learning system and its data-driven approaches. | | Used incorrect machine learning tools and algorithms. Insufficient knowledge and understanding of the area of study | | The project uses some machine learning techniques but may have obvious technical omissions. | | | The machine learning algorithms and data processing approaches are correct without major technical omissions. Alternative methods are used and discussed. | | | Plus, there are a wider understanding of the applied methods. Suitable evaluation approaches have also been used. | | | Plus, the implementation of effective machine learning techniques. Some more advanced techniques beyond what has been taught are implemented. | | | Plus, the technical difficulty and complexity are equal or beyond current research level and/or industrial standards. |
| ✓ |  | |  | |  | | |  | | |  | | |  | | |  |
| **Video and Supporting Documentation:**  **Quality**  20% | The project is not complete. The quality of the project is not appropriate to the level of the study. | | The project has poor quality, and the machine learning is missing major functions. | | The quality of the project is acceptable, but the quality may be inconsistent, and there may be obvious missing elements to the project. | | | The quality is consistently good, and there are no obvious missing elements to the project. | | | Plus, the quality is consistently high, and there is a good level of finish to the project. | | | Plus, the quality is consistently exceptional, and there is an excellent level of finish to the project. There is a certain amount of original contribution during the project development. | | | Work may achieve or be close to a publishable level. |
| ✓ |  | |  | |  | | |  | | |  | | |  | | |  |
| **Video and Supporting Documentation:**  **Achievement**  20% | The machine learning project is not able to fulfil the intended aim of objectives. | | The system has some functions for data science projects, the main aim and/or objectives are missing | | The project is theoretically capable of fulfilling the intended goals, but some of these have not been achieved. | | | All the project aims are achieved. | | | Plus, the system has comparable system performance against popular benchmarks (for example, the system has been tested on other datasets) | | | Plus, scientific approaches are used for system evaluation. The work shows evidence of in-depth understanding of how to develop a machine learning system. | | | Plus, the output from this project can be tested and evaluated in real-world scenarios. |
| ✓ |  | |  | |  | | |  | | |  | | |  | | |  |
| **Video: Demonstration**  20% | Fails to address the project design and development during the video demonstration | | A limited attempt has been made to present the developed work in the video. | | The video covers the essential contents of project design and development, but it may be missing many project components. | | | The video includes major system design and development components, with very little errors and omissions. | | | The video shows that the student has a deeper understanding of the implementation of machine learning algorithms from end to end. There is limited discussion about the system performance. Alternative approaches are mentioned. | | | Plus, a detailed discussion on the system performance. The video shows evidence of deep understanding of using scientific methods to compare and evaluate the system performances. | | | Plus, the student demonstrated the exceptional breadth and depth of the machine learning knowledge in the demonstration. |
| ✓ |  | |  | |  | | |  | | |  | | |  | | |  |
| **Overall mark:** |  | |  | |  | | |  | | |  | | |  | | |  |

**University Grade Descriptor (UGD)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class** | **Category** | **Mark range** | **%** | **General Characteristics** |
| 1st | Exceptional 1st | 93 - 100 | 96 | Exceptional breadth and depth of knowledge and understanding of the area of study, **significantly beyond what has been taught in all areas**; evidence of extensive and appropriate selection and critical evaluation/synthesis/analysis and of reading/research beyond the prescribed range, in both breadth and depth, to advance work/direct arguments; excellent communication; performance deemed to be beyond expectation. **Work at publishable or commercial standard**. The ability to make decisions and systematically carry out tasks/processes with autonomy in unpredictable situations; exercise of initiative in the completion of practical tasks; exceptional leadership skills and evidence of personal responsibility in group contexts; creative flair; extremely well-developed problem-solving skills; the ability to carry out sustained critical reflection on practical work within the wider context of industry/workplace. Exceeds expectations set by the industry/employment context. |
| 1st | High 1st | 85 - 92 | 89 | Excellent knowledge and understanding of the area of study as the student is **typically able to go beyond what has been taught (particularly for a mid/high 1st)**; evidence of extensive and appropriate selection and critical evaluation/synthesis/analysis of reading/research beyond the prescribed range, to advance work/direct arguments; excellent communication; performance deemed beyond expectation of the level. The ability to make decisions and carry out tasks/processes with a high level of autonomy; creative flair and the ability to (re)interpret predefined rules/conventions to select and justify individual working practice; excellent problem-solving skills; accuracy and fluency; excellent command of skills appropriate to the task; the ability to reflect critically on practical work within the wider context of industry/workplace. Meets expectations set by the industry/employment context. |
| Mid 1st | 78 - 84 | 81 |
| Low 1st | 70 - 77 | 74 |
| 2.1 | High 2.1 | 67 - 69 | 68 | Very good knowledge and understanding of the area of study as the student is **typically able to relate facts/concepts together with some ability to apply to known/taught contexts**; evidence of appropriate selection and critical evaluation of reading/research, some beyond the prescribed range, may rely on set sources to advance work/direct arguments; demonstrates autonomy in approach to learning; strong communication skills. Broadly autonomous completion of practical tasks/processes; ability to adapt in response to change or unexpected experiences; technical/artistic decision making is very highly developed; a clear command of the skills relevant to the task/process; ability to reflect on practical work and set future goals within the wider context of industry/workplace. Adherence to standards set by the industry/employment context. |
| Mid 2.1 | 64 -66 | 65 |
| Low 2.1 | 60 - 63 | 62 |
| 2.2 | High 2.2 | 57 - 59 | 58 | Good knowledge and understanding of the area of study **balanced towards the descriptive rather than critical or analytical;** evidence of appropriate selection and evaluation of reading/research, some may be beyond the prescribed range, but generally reliant on set sources to advance work/direct arguments; communication shows clarity, but structure may not always be coherent. A confident approach to practical tasks; a solid grasp of the related processes, tools, technology; creativity in the completion of the task; proficiency is demonstrated by an accurate and well-coordinated performance; tasks are completed with a good level of independent thought and autonomy; an ability to reflect on practical work and set future goals. General adherence to standards set by the industry/employment context. |
| Mid 2.2 | 54 - 56 | 55 |
| Low 2.2 | 50 - 53 | 52 |
| 3rd | High 3rd | 47 - 49 | 48 | **Knowledge and understanding sufficient to deal with terminology, basic facts and concepts** but fails to make meaningful synthesis; some ability to select and evaluate reading/research however work may be more generally descriptive; general reliance on set sources to advance work; arguments may be weak or poorly constructed; communication/presentation is generally competent but with some weaknesses. Competence in technical/artistic skills; tasks/processes are completed with a degree of proficiency and confidence; tasks are completed with a sufficient level of independent thought; effective judgements have been made; evaluation and analysis of performance in practical tasks is evident. Errors in workflow or completion of the task; general adherence to appropriate rules/conventions set by the industry/employment context. |
| Mid 3rd | 44 - 46 | 45 |
| Low 3rd | 40 - 43 | 40 |
| Fail | Borderline Fail | 30 - 39 | 35 | Insufficient knowledge and understanding of the subject and its underlying concepts; **some ability to evaluate given reading/research however work is more generally descriptive; naively follows or may ignore set material in development of work**; given brief may be only tangentially addressed or may ignore key aspects of the brief; communication shows limited clarity, poor presentation, structure may not be coherent. Practical tasks are attempted; skill displayed in some areas; there are a significant number of errors; a lack of proficiency in most areas; guidance may be needed to reproduce aspects of the task and/or apply learned skills. Tasks may be incomplete; failure to adhere to some of the rules/conventions set by the industry/employment context. |
| Mid Fail | 20 - 29 | 25 |
| Low Fail | 6-19 | 10 | No evidence of knowledge or understanding of the subject; **no understanding of taught concepts, with facts being reproduced in a disjointed or decontextualised manner**; ignores set material in development of work; fails to address the requirements of the brief; lacks basic communication skills. A general level of incompetency in practical tasks; an evident lack of practice; set tasks are not completed; few or no skills relating to tasks are evident. No adherence to rules/conventions set by the industry/employment context. |
| Zero | Zero | 0-5 | 0 | Work not submitted, work of no merit, penalty in some misconduct cases. |

**Artificial Intelligence and Academic Integrity – AI&AI**

It is important you do not use AI tools to generate an assignment and submit it as if it were your own work. Our regulations state:

Contract cheating/concerns over authorship: This form of misconduct involves another person (or artificial intelligence) creating the assignment which you then submit as your own. Examples of this sort of misconduct include:

* buying an assignment from an ‘essay mill’/professional writer
* submitting an assignment which you have downloaded from a file-sharing site
* acquiring an essay from another student or family member and submitting it as your own
* attempting to pass off work created by artificial intelligence as your own.

These activities show a clear intention to deceive the marker and are treated as misconduct.

Further guidance is available here: [https://blogs.shu.ac.uk/assessment4students/preparing-to-submit-work/#AI](https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fblogs.shu.ac.uk%2Fassessment4students%2Fpreparing-to-submit-work%2F%23AI&data=05%7C02%7Cm.jacobi%40shu.ac.uk%7Ce43edfef59d143757cc808dc9077c828%7C8968f6a1ac13472fb899f7316e439f43%7C0%7C0%7C638544091847747773%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=Mu35s2bWJFR7ESDUdLI4GvdJiCwl6CyBS42osK5Z468%3D&reserved=0)