

# LDS7003M Assignment Brief (Component 1)

## Contents

Module Details .....	1
Assignment Description.....	2
Learning Outcomes .....	3
Advice and Guidance .....	4
How is this assessment marked? .....	5
Marking Criteria .....	6

Module Details			
<b>Module code:</b>	LDS7003M- Artificial Intelligence and Machine Learning	<b>Level of Study:</b>	7
<b>Module Leader(s):</b>	Swathi Ganesan	<b>Credits:</b>	20
<b>Assessment format:</b>	Creative Artefact - An artificial Intelligence practical project.	<b>Method of submission:</b>	Turnitin within Moodle
<b>Deadline or Assessment Period:</b>	<b>28<sup>th</sup> Jan 2025, 12Noon</b>	<b>Feedback date and place:</b>	19 <sup>th</sup> Feb 2025 Turnitin within Moodle
<b>Assessment limits:</b> length, load, word count, etc.	N/A	<b>Component number:</b>	1 of 2
<b>Is this exempt from anonymous marking under the policy?</b>	No	<b>Component weighting:</b>	60%

## Assignment Description

### Part 1: Theoretical Analysis (20%)

**Examine** the significance of ontology in structuring knowledge within AI systems. Critically analyse how the structure and components of an ontology can influence reasoning and decision-making in AI applications. Discuss the potential of using ontologies in emerging fields (e.g., healthcare, finance). Analyse a case where ontology has been successfully implemented and discuss what could have been done differently to enhance its effectiveness.

### Part 2: Practical Implementation (80%)

This task is designed to assess your ability to solve a real-world problem using both classification and clustering techniques within a specific domain. You will need to demonstrate your understanding of these machine learning methods, along with critical thinking and problem-solving skills in applying them to relevant data.

**Scenario:** You are required to develop an intelligent system that applies both classification and clustering techniques to solve a problem in a particular domain. The domain could be anything from healthcare, finance, education, or e-commerce. Your challenge is to identify a problem where classification can be used to predict outcomes or categorize data, and where clustering can help discover hidden patterns or segment data into meaningful groups.

#### Your Task:

##### 1. Problem Identification

- Identify a problem within a specific domain where both classification and clustering techniques can be applied to enhance decision-making or provide insights.
- Clearly define the problem, explaining why it is important and how AI techniques can address it.

##### 2. Dataset Selection

- Select or find a relevant dataset that contains features suitable for both classification and clustering. The dataset should have sufficient complexity and size to allow for meaningful analysis.
- Justify your choice of the dataset, detailing its relevance to the problem and how it will help in solving it.

##### 3. Model Development

- **Classification Task:** Develop a machine learning classification model to predict an outcome or categorize data based on the features in your dataset. This could involve binary classification (e.g., yes/no decisions) or multi-class classification (e.g., categorizing into multiple groups).
- **Clustering Task:** Apply a clustering algorithm to the same dataset to uncover hidden patterns, group similar data points, or segment the data into distinct clusters.
- Document your entire process, including data preprocessing, feature selection, model selection, and optimization. Provide code snippets, outputs, and visualizations to illustrate your approach.

##### 4. Critical Evaluation

- **For Classification:** Evaluate the performance of your classification model using appropriate metrics such as accuracy, precision, recall, and F1-score.
- **For Clustering:** Assess the quality of your clustering results using metrics such as silhouette score or elbow method to determine the optimal number of clusters or any other relevant clustering evaluation methods.

## Assignment Description

- Discuss the insights gained from both techniques, the strengths and limitations of your models, and suggest potential improvements or alternative approaches.

### 5. Reflection

- Reflect on the challenges encountered during the project, including any difficulties in applying both classification and clustering to the same dataset.
- Discuss the implications of your findings and how these techniques could be applied in a real-world scenario within your chosen domain.

## Learning Outcomes

You must successfully achieve the following Learning Outcomes to pass this assessment:

PLOs **7.1-7.5, 7.8, 7.9**

7.1 Critically apply skills, techniques, and knowledge from a range of data analysis methods and algorithms for enhancing and solving problems in various domains.

7.2 Develop abstract thinking and design ability to analytically demonstrate concepts relating to data science.

7.3 Use research-based knowledge for the design of experiments, analysis, and interpretation of data to provide valid results.

7.4 Critically evaluate and analyse advanced data science topics, and concepts, and implement them in workplace.

7.5 Identify and implement appropriate programming and software tools to critically analyse big data applications in workplace.

7.8 Critically analyse the data and apply predictive modelling technique in the field of Machine Learning and Artificial Intelligence.

7.9 Critique legal, social, and ethical issues within the field of data science and applicable ancillary sectors, as applied to contemporary research and industrial practice.

## Advice and Guidance

### Assessment Guidance

#### Deliverables:

1. **Report:** Write a report (3000 words) covering both part 1 and part 2. Convert the file into single PDF format for final submission.
2. **Dataset:** Submit the dataset you used or provide a link to it. Explain why it was chosen and how it relates to the problem in the report clearly.

## Advice and Guidance

3. **Code:** Provide well-documented Python code that includes all steps from data preprocessing to model evaluation for both classification and clustering tasks. The code file in .pynb format should also be submitted.

### General Guidance

#### General considerations

The work you present should be your own work, and not just copied from others. You can quote from others, but you must say who the author is and use quotation marks or paraphrase. If you do not do so, we will investigate your work for academic misconduct. This is particularly likely if your Turnitin similarity score is above 25% and/or individual matches are above 6%.

If you require support with your study skills, please visit <https://www.yorks.ac.uk/students/study-skills/>

#### Assessment Regulations

Please refer to the York St John University Code of Practice for Assessment and Academic Related Matters 2023-24.

We ask that you pay particular attention to the academic misconduct policy. Penalties will be applied where a student is found guilty of academic and/or ethical misconduct, including termination of programme ([Policy Link](#)).

You are required to keep to the word limit set for an assessment and to note that you may be subject to penalty if you exceed that limit. You are required to provide an accurate word count on the cover sheet for each piece of work you submit ([Policy Link](#)).

For late or non-submission of work by the published deadline or an approved extended deadline, a mark of 0NS will be recorded. Where a re-assessment opportunity exists, a student will normally be permitted only one attempt to be re-assessed for a capped mark ([Policy Link](#)).

An extension to the published deadline may be granted to an individual student if they meet the eligibility criteria of the ([Policy Link](#)).

## How is this assessment marked?

Your work will be marked according to the assessment instructions provided within this document and the selected Learning Outcomes' (LOs) (see above).

Furthermore, this assessment is marked using the assessment marking criteria or a similar rubric that aligns with the University's Generic Assessment Descriptors (see below).<sup>1</sup> This is to ensure all assessment decisions are comparable regardless of the discipline or mode of assessment.

Please note that you **must** meet the required baseline standards (50 – 59%) which will include the LOs and minimum expectations of the assessment. Further still, you must ensure you meet the requirements of each grade boundary to progress to the next, i.e., you should demonstrate your

<sup>1</sup> A rubric is a type of scoring guide that markers use to set out specific components and expectations for an assignment for their students. It is then used to guide the marking they undertake.

### **How is this assessment marked?**

learning through the standards of the Pass, Merit and Distinction to reach a Distinction (70 – 84%). These standards are designed to scaffold and build your learning to achieve your fullest potential in each criterion being assessed.

## Marking Criteria

**Pass Grade Bands (100 – 50)** (Learning Outcomes must be met)

**Fail Grade Bands (49 – 0)** (Learning Outcomes are not met)

Assessment Criteria	Description	Mark (100%)
<b>Theoretical Analysis</b>	Examine the significance of ontology in AI systems. Critically analyse its structure and components, discussing their impact on reasoning and decision-making. Explore potential applications in fields such as healthcare or finance. Analyse a case of ontology use and suggest improvements.	<b>20%</b>
<b>Problem Identification</b>	Clearly identify a domain-specific problem that can be solved using classification and clustering techniques. Provide a well-defined explanation of the problem's importance and relevance to AI techniques.	<b>10%</b>
<b>Dataset Selection</b>	Select a relevant dataset for classification and clustering, ensuring sufficient complexity. Justify its selection based on the problem and provide a rationale for how the dataset supports solving the identified problem.	<b>10%</b>
<b>Model Development</b>	Develop classification and clustering models. Justify the choice of algorithms and demonstrate an understanding of their workings. Provide a detailed explanation of data preprocessing, feature selection, model optimization, and present outputs with visualizations.	<b>25%</b>
<b>Evaluation</b>	Evaluate the performance of classification models using metrics such as accuracy, precision, recall, F1-score. Assess clustering models using silhouette score or elbow method. Provide critical insights into model strengths and weaknesses, suggesting improvements.	<b>15%</b>
<b>Reflection</b>	Reflect on the challenges encountered in applying classification and clustering techniques. Discuss the real-world applicability of the models and ethical considerations related to data handling, privacy, and potential biases.	<b>10%</b>
<b>Report Writing and Referencing</b>	Evaluate the clarity, structure, and professionalism of the report. Assess proper citation of sources, adherence to academic standards, and appropriate use of references to support arguments.	<b>10%</b>

## Level 7 GAD Descriptor for Assessment Matrix

Assessment Criteria		Pass (50 – 59)	Merit (60 – 69)	Distinction (70 – 84)	Distinction (85 – 100)	Borderline Fail (45 – 49)	Fail (30 – 44)	Fail (0 – 29)
<b>Theoretical Analysis (20%)</b>	<b>Thinking Skills &amp; Research Skills</b>	Basic understanding of ontology, with limited critique of its structure and impact.	Clear understanding of ontology's significance, with thoughtful critique.	In-depth understanding with strong analysis of ontology's role in AI.	Exceptional grasp with insightful analysis and novel suggestions for improvement.	Limited or surface-level analysis of ontology, with minimal critical thinking.	Inadequate understanding and poor analysis of ontology's role.	No or very poor analysis of ontology, lacking any critical thinking.
<b>Problem Identification (10%)</b>	<b>Thinking Skills &amp; Research Skills</b>	Basic identification of a problem that can be solved using classification and clustering.	Clearly defined problem, relevant to AI, with appropriate explanation.	Thorough identification and explanation of the problem's relevance to AI.	Insightful problem identification, well explained, showing deep understanding.	Vaguely identified problem, lacking relevance to AI techniques.	Poor problem identification with little explanation or relevance.	No proper identification of a problem relevant to classification and clustering.
<b>Dataset Selection (10%)</b>	<b>Thinking Skills &amp; Research Skills</b>	Basic selection of a dataset with limited justification of its relevance.	Well-justified dataset selection, relevant to the problem and techniques.	Strong justification for dataset complexity and relevance to the problem.	Excellent dataset selection with deep insights into its suitability for AI techniques.	Minimal explanation of the dataset's relevance or selection criteria.	Inadequate dataset selection, lacking relevance or justification.	No or very poor dataset selection with no justification.
<b>Model Development (25%)</b>	<b>Practical Skills &amp; Professional Learning Skills</b>	Basic development of classification and clustering models with limited explanation.	Clear implementation of models, with appropriate justification for algorithms.	Strong model development with advanced use of algorithms and optimisation.	Expert-level implementation and explanation of models, with exceptional use of optimisation techniques.	Limited or poorly implemented models, lacking clear explanation.	Poor implementation of models with little understanding of AI techniques.	No implementation of models or incorrect application of techniques.
<b>Evaluation (15%)</b>	<b>Practical Skills &amp; Professional Learning Skills</b>	Basic evaluation using limited metrics, with minimal insight into model performance.	Good evaluation of models using appropriate metrics, with clear insights into performance.	Strong evaluation using advanced metrics, providing in-depth analysis of strengths and weaknesses.	Exemplary evaluation with insightful analysis and suggestions for improvement.	Limited evaluation of models with basic use of metrics.	Inadequate evaluation with poor use of metrics, lacking critical insight.	No meaningful evaluation of models or use of performance metrics.
<b>Reflection (10%)</b>	<b>Thinking Skills &amp; Research Skills</b>	Basic reflection on challenges and potential real-world applications.	Thoughtful reflection on the challenges and real-world implications, with some consideration of ethical issues.	Strong reflection with clear discussion of real-world applications and ethical considerations.	Exceptional reflection, offering deep insights into real-world applications, challenges, and ethical concerns.	Limited reflection on challenges and practical applications.	Inadequate reflection with little consideration of real-world implications or ethics.	No reflection or discussion of challenges, real-world applications, or ethics.

## London School

<b>Report Writing and Referencing (10%)</b>	<b>Communication and Creativity</b>	Basic report structure and referencing with minimal citation adherence.	Well-structured report with good referencing and appropriate citations.	High-quality report with accurate citations and appropriate referencing, adhering to academic standards.	Exemplary report, structured clearly, with correct and extensive use of academic references and citations.	Basic report with minimal structure or clarity, and inconsistent citations.	Inadequate structure and referencing, with improper citation and poor clarity.	No structure, improper or no referencing, with no clear adherence to academic standards.
---	-------------------------------------	---	---	--	--	---	--	--