

Assessment Brief

Module Leader: Pratikshya Sharma		Level: 5
Module Name: Artificial Intelligence & Machine Learning 1		Module Code: 55-508228
Assignment Title: Connect 4 Game Playing AI Agent		
Individual Assessment	Weighting: 100 %	Magnitude: 30 hrs workload
Submission date/time: 1st May, 2025 , 15:00 (UK Time)	Blackboard submission: Yes Turnitin submission: No	Format: <ul style="list-style-type: none"> • <i>Source code</i> • <i>Code walkthrough video</i>
Planned feedback date: 23 rd May 2025	Mode of feedback: <ul style="list-style-type: none"> • <i>IT Sessions- Formative feedback</i> • <i>Blackboard "My Grades"-Summative Feedback</i> 	
In this assessment are students asked to consider:	Inclusivity and accessibility	Not applicable
	Sustainability	Not applicable
<u>Module Learning Outcomes</u>		
<ul style="list-style-type: none"> • LO1: Gain a foundational knowledge of artificial intelligence, its core concepts, and fundamental principles. 		
<ul style="list-style-type: none"> • LO2: Develop the ability to apply range of AI to solve complex problems and make informed decisions. 		
<ul style="list-style-type: none"> • LO3: Acquire practical skills in AI programming languages and tools, enabling implementation of AI algorithms and & ML models for building systems with decision making capabilities. 		

1.0 Assessment Brief

This is an individual assessment with 100% weighting and aims to assess your practical understanding of two core concepts of this module:

(a) AI algorithms for building an intelligent system.

(b) Machine learning models for decision making.

In this assessment there are tasks that you are required to work on, focusing on the above mentioned core concepts. To assess your understanding for the AI&ML algorithm concepts you are required to build a **Connect4 Game Playing AI agent** with a Think-Act Cycle. This means the agent must perceive the game environment, analyse it intelligently, and respond with calculated moves. The game must be able to correctly function for the scenarios described under the various section below.

1.1 Introduction to Connect 4 Game

Connect 4 is a two-player game played on a grid consisting of seven columns and six rows (i.e. 6x7) as demonstrated in Figure 1, with total 42 slots. Each player takes turns to drop a disc in one of the seven columns. The disc then drops to the lowest empty slot in that column. The disc can be dropped only from one of the columns. For example, given the state of the game in Figure 2, if the player chooses to drop green disc in column 3, the state of the game changes to that depicted in Figure 3. Each player is associated with a unique colour to distinguish the disc dropped. To win the game the player must align four discs (of the colour assigned to that player) consecutively either horizontally, vertically or diagonally. First player to connect four of their discs in a line (i.e., horizontally or vertically or diagonally), as mentioned earlier, wins the game. Instances of win states for the player associated with yellow disc is demonstrated in Figure 4. You can click the link provided here to explore and play the game.

<http://faculty.otterbein.edu/wittman1/games/connect4/>

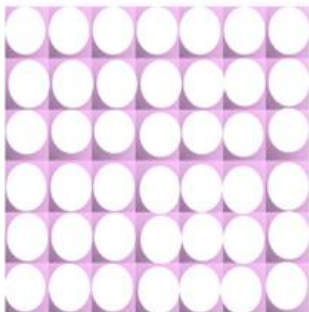


Figure 1: Connect 4 Game Board

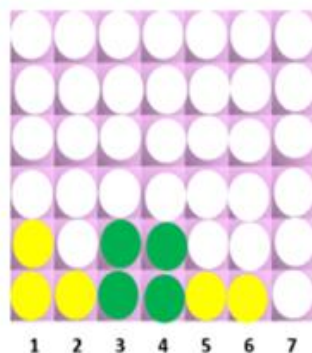


Figure 2: Before green disc drop.

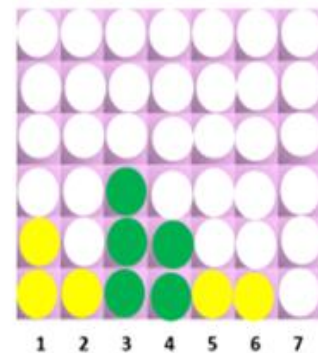
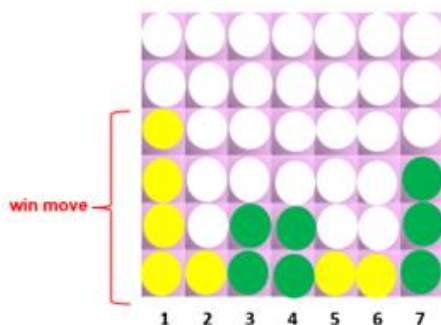
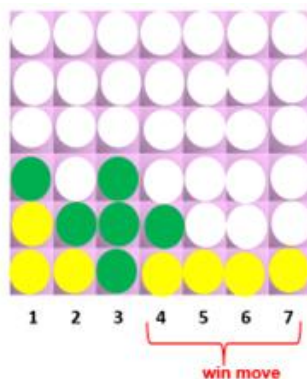


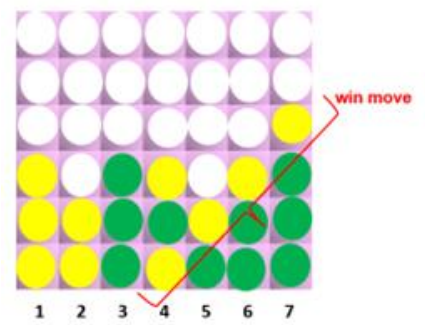
Figure 3: After green disc drop.



(a) Vertical



(b) Horizontal



(c) Diagonal

Figure 4: Instances of win moves for player with yellow discs.

1.2 Initial Game Board & General Game Settings

Create a 6x7 board that is empty initially to implement a basic game grid. The setup must have two players, player1(e.g. human) and player2 (e.g. AI agent). Player1 is assigned with a disc symbol '●' as and the player2 is assigned with a disc symbol '○'. An instance of the game board with players chosen symbol is in Figure5. Player1 uses '●' and the Player2 uses '○' to mark its place on an empty cell in the board. The coding must guarantee that both the players refrain from replacing a slot that has already been occupied.

After every player's move, the status of the board must be displayed. A player emerges victorious when their designated disc aligns consecutively four times in a row, column, or diagonal on the game board. The game comes to a draw when the game board is filled, and there are no four consecutive sequences of '●' (for player1) or '○'(for player2) in any row, column, or diagonal. This back-and-forth play continues until either one player wins or a tie occurs.

Upon reaching this point, the program displays the outcome of the game and exits accordingly. There are three possible outcomes of this game and always end with one of these outcomes as mentioned below.

- a.) Player1 wins.
- b.) Player2 wins.
- c.) Game ends in a draw.

In order to showcase your understanding and proficiency in Artificial Intelligence (AI), you are required to develop and implement four distinct AI agents that can play the Connect 4 game. These agents are Random Agent, Smart Agent, Mini-max Agent and Machine Learning (ML) Agent. These agents will play against various opponents, as outlined from section 1.3 onwards, to evaluate their performance in different contexts. After creating these agents, you will conduct a performance analysis as detailed in section 1.7. This analysis will involve evaluating the effectiveness of each agent, comparing their decision-making abilities, and determining how well each agent performs under different game scenarios. Through this process, you will not only develop multiple AI agents with varying complexities but also gain insight into their strategic capabilities. The performance analysis will provide a deeper understanding of how each approach—ranging from simple random decisions to sophisticated AI algorithms like Minimax and machine learning—affects the outcome of the Connect4 game. This will allow you to demonstrate your skills in building intelligent systems, analysing performance, and understanding the strengths and weaknesses of different AI approaches.

1.3 Random Agent

Build a random agent that plays the Connect4 game against an opponent (e.g. human). This agent simply makes a list of columns where a move is possible and then randomly picks one from the list. The agent places its disc randomly on any empty grid and follows no other specific rules while making its move.

1.4 Smart Agent

Extended the code to build a smart agent by incorporating a rule-based algorithm. The Smart Agent always follow simple rules outlined below during its turn. The Smart Agent is setup to play against its opponent e.g. any other AI

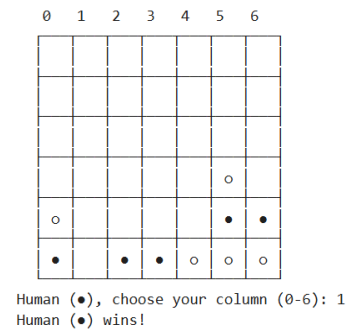


Figure 5: Game Boards.

In the dataset, 'x' is the first player; 'o' the second and the outcome class is the game theoretical value for the first player. You must use the data in "connect-4.data" to train a ML model that can predict the next move for a given board state. The move it chooses must lead to a win, loss, or draw based on previous game outcomes of this dataset. This means that for any given board state, the model will predict the optimal move that has historically led to one of these outcomes. Once the model is trained, it should be integrated to build a machine learning-based agent that can play the game. This ML agent should be able to engage in two types of games:

1. **Human vs. ML Agent:** The trained agent will play against a human player, where the human provides the input for their move, and the ML agent predicts and makes its next move based on the trained model.
2. **Mini-Max Agent vs. ML Agent:** The ML agent will also play against a Mini-Max agent as its opponent. Here, the ML agent will predict the next best move based on the board state, while the Mini-Max agent makes decisions according to its own strategy. The objective is for the ML agent to apply its learned strategy and adapt to the outcomes of these games.

In both scenarios, the model must make a move based on the dataset and the outcomes it has learned, and the agent should play the game according to the predictions made by the trained model. This setup allows you to test the predictive capability of your model against both a human and an algorithmic opponent, providing valuable insights into its performance and accuracy in different game contexts. Remember to cite the following when you use the dataset in your implementation to acknowledge the use of the dataset.

J. Tromp. "Connect-4," UCI Machine Learning Repository, 1995. [Online]. Available: <https://doi.org/10.24432/C59P43>.

1.7 Performance Evaluation

Your code must have a section for performance evaluation and analysis. A performance evaluation is important for realising the efficiency of the agent. You must use appropriate and meaningful metrics, showcasing deep understanding of the AI agent's working mechanisms. You can use table/graph/bar charts/plots or other graphical visualisations to demonstrate the analysis. A list of some metrics that can help measure agent's performance is given below but not exhaustive, you can use any other relevant metrics as well.

- (a) Accuracy Metrics (win rate, draw rate & loss rate).
- (b) Search Performance Metrics (Search Depth, Nodes Expanded & branching factor)
- (c) Efficiency Metrics (Execution time, Pruning Effectiveness – if using alpha-beta pruning)
- (d) Heuristic Evaluation Quality & Counterplay Ability
- (e) Game-Level Metrics (Game Length, Winning Patterns: frequency of horizontal, vertical, diagonal winning strategies).
- (f) Robustness Metrics (i.e., performance against different opponents).
- (g) Resource Utilization Metrics (i.e., memory usage).

To achieve the required performance evaluation of these agents you can run 500(approximately) games each for the following:

- Random agent vs Smart (Rule-based) agent.
- Smart (Rule-based) agent vs mini-max agent.
- Mini-max agent vs ML Agent.

1.8 Innovation and Beyond

While the assessment task encourages the utilization of the listed AI functionalities, you are also welcome to introduce additional AI capabilities, within or beyond those covered during AI module. This presents an opportunity to push the boundaries of your system's capabilities and explore creative avenues beyond the prescribed techniques. The game

should not only reflect technical proficiency but also exhibit a deep understanding of AI principles. The assessment task serves as an avenue for you to demonstrate your AI knowledge, creativity, and practical implementation skills.

2.0 Deliverables

2.1 Code

- Use Python to implement your solution. Submit your complete implementation with all dependencies. The code must be error/bug free with a readme file outlining execution step or any essential information.
- You may have multiple versions of your code (especially when you have many different experiment settings). You are expected to submit all the used and unused codes for the assessment.
- Ensure that your code is well-structured with proper in-code comments. The code must have suitable indentation for full readability and understanding.
- You will use many sources in programming the given problem. Include references to these sources in the code and a readme file.

2.2 Video

You are required to create a 12-minute video (recorded and playable at normal speed) that provides a comprehensive walkthrough of your code for the task, demonstrating the key tasks outlined previously. In this video, you will need to record your screen as you execute each of the tasks, ensuring that the corresponding output is clearly shown. During this video, you should clearly explain and demonstrate the working of the Mini-max agent & ML agent, highlighting any key aspects. You should also cover your implementation steps, walking through the code structure and key algorithms used. Lastly, it is important to explain your experiment setup, detailing how you tested your code and validated the results. By the end of the video, there should be a clear understanding of how your code works, how you approached the problem, and the logic behind your solutions along with result analysis. You must also clearly demonstrate & highlight any additional innovative AI functionalities (if used).

3.0 Submission Guidelines

You must submit, to Blackboard, by **3:00 PM (UK Time) on Thursday 1st May 2025**, a single ZIP file that contains all your materials (code & video). If you create a large file (more than the file limit of the Blackboard site), you should add a Google Drive link to the zip file. Ensure the timestamps in the shared drive are earlier than the deadline. The zip file submission must be named as **StudentId_FirstName.zip** and should contain the code and video. Any content in video beyond the specified limit will be ignored and left unmarked.

4.0 Who do I contact if I have a question?

You can speak to the module leader, (Pratikshya Sharma, email: pratikshya.sharma@shu.ac.uk) for assessment support and discussions.

5.0 Assessment Criteria

The submitted work will be assessed based on the rubrics and evaluation criteria outlined in this section. These rubrics provide a clear framework for grading and specify the key aspects that will be examined during the evaluation process. Each criterion is designed to measure specific elements of the task, such as the quality of the code, adherence to requirements, logical problem-solving, originality, presentation and more.

5.1 Marking Criteria & Weighting

Criteria	Weighting	Description
Criterion1	10 %	Prescribed AI Functionalities implementation & completeness: <ul style="list-style-type: none">• Initial Game Board Setting & General Game Settings• Random Agent• Smart Agent
Criterion 2	20 %	Prescribed AI Functionalities implementation & completeness: <ul style="list-style-type: none">• Minimax-Agent
Criterion 3	20 %	Prescribed ML Functionalities implementation & completeness: <ul style="list-style-type: none">• Machine Learning Agent
Criterion 4	20 %	Performance Evaluation: <ul style="list-style-type: none">• Random agent vs Smart (Rule-based) agent• Smart (Rule-based) agent vs mini-max agent• Mini-max agent vs ML Agent• Overall Analysis.
Criterion 5	10 %	Innovation & Beyond: <ul style="list-style-type: none">• Additional AI & ML Capabilities
Criterion 6	10 %	Code walkthrough: <ul style="list-style-type: none">• Understanding of the Code and AI Functionalities (Knowledge of Concepts)• Logic and Flow• Areas for improvement• Design Choices
Criterion 7	10 %	Code Quality: Adheres to best practices and standards (Clarity, Readability, Naming Conventions, Code Formatting, Comments, Scalability etc.).

5.2 Marking Rubrics

	FAIL (insufficient)				THIRD (sufficient)			LOWER SECOND (good)			UPPER SECOND (very good)			FIRST (excellent)			
	No submission	Low Fail	Mid Fail	Borderline 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
Numeric Marks	0	10	25	35	40	45	48	52	55	58	62	65	68	74	81	89	96
Criterion 1 (10 %)	No submission	No evidence of use of prescribed algorithm & functionalities.	Highly insufficient & incomplete evidence of use of prescribed algorithm.	Prescribed algorithms & functionalities are all implemented incorrectly, with limited work clarity for the agents.	An attempt to implement some prescribed algorithms with some prescribed functionalities. All agents built have several obvious technical omissions and deficiencies in the core functionalities.			An attempt to implement all prescribed algorithms with all prescribed functionalities. Some agents built have some significant technical omissions.			All prescribed algorithms & functionalities are correctly implemented, but some minor deficiencies exist. All the agents are working properly with few acceptable omissions or oversights.			All prescribed algorithms & functionalities are correctly implemented according to the requirements. All agents are working correctly and meets the expected requirements with no omissions or oversights.		All prescribed algorithm & functionalities are correctly implemented. Agent's functionality surpasses the expected requirements with added AI capabilities.	
Criterion 2 (20 %)	No submission	No evidence of use of prescribed algorithm & functionalities.	Highly insufficient & incomplete evidence of use of prescribed algorithm.	Prescribed algorithms & functionalities are all implemented incorrectly, with limited work clarity for mini-max agent.	An attempt to implement the prescribed algorithms with some prescribed functionalities. The mini-max agent built has several obvious technical omissions and deficiencies in the core functionalities.			An attempt to implement the prescribed algorithm with all prescribed functionalities. Mini-max agent built has some significant technical omissions.			All prescribed algorithm & functionalities are correctly implemented, but some minor deficiencies exist. The mini-max agent is working with few acceptable omissions or oversights.			All prescribed algorithm & functionalities are correctly implemented according to the requirements. Mini-max agent is working correctly and meets the expected requirements with no omissions or oversights.		All prescribed algorithm & functionalities are correctly implemented. Mini-max agent's functionality surpasses the expected requirements with added AI capabilities.	
Criterion 3 (20 %)	No submission	No evidence of use of ML algorithm & functionalities.	Highly insufficient & incomplete evidence of use of ML algorithm.	ML algorithms & functionalities are all implemented incorrectly, with limited work clarity for the ML agent.	An attempt to implement ML algorithm with some prescribed functionalities. The ML agent built has several obvious technical omissions and deficiencies in the core functionalities.			An attempt to implement ML algorithms with all prescribed functionalities. ML agent built has some significant technical omissions.			ML algorithm & prescribed functionalities are correctly implemented, but some minor deficiencies exist. The ML agent is working with few acceptable omissions or oversights.			ML algorithm & all prescribed functionalities are correctly implemented according to the requirements. ML agent is working correctly and meets the expected requirements with no omissions or oversights.		ML algorithm & all prescribed functionalities are correctly implemented. ML agent's functionality surpasses the expected requirements with added AI capabilities.	
Criterion 4 (20 %)	No submission	No evidence of use of performance evaluation.	Highly insufficient & incomplete performance evaluation.	Incorrect usage of performance evaluation methods, fails to make meaningful analytical evidence.	Minimal ability for performance evaluation, is seen. Obvious deficiencies exist in the performance evaluation implementations.			Performance evaluation is present but lacks depth or clarity, with basic analytical evidence. Minor deficiencies exist in the performance evaluation. All agents are not included in the analysis.			Performance evaluation is presented correctly with some depth or clarity and acceptable analytical evidence. Majority of agents are included in the analysis.			Adequate performance evaluation with few appropriate metrics and satisfactory analysis. All agents are included in the analysis.		Performance evaluation with several meaningful metrics and detailed analysis, showcasing deep understanding is presented. All agents are included in the analysis.	
Criterion 5 (10 %)	No submission	No evidence of use of additional AI capabilities.	Incomplete attempt to add extra AI functionalities.	None of the additional AI functionality implemented is working correctly.	An attempt for one additional AI functionality is included in the implementation, but with major technical omissions.			Limited additional AI functionalities are present and are working correctly with minor and acceptable technical omissions.			Few additional AI functionalities are present and implemented correctly but lack creativity or significance. Majority of these functionalities are working correctly with no technical omissions.			Some additional AI functionalities are implemented correctly, adding value to the project. Has clear evidence of wider understanding of the applied methods.		Several additional AI functionalities are creatively implemented, enhancing the project's overall quality and functionality. The added functionality clearly surpasses the expectations.	
Criterion 6 (10 %)	No submission	No evidence of understanding of taught AI/ML concepts.	Significantly lacks understanding and proficiency of AI&ML concepts.	Explanation to exhibit understanding of the code and AI&ML functionalities is highly unsatisfactory, incorrect, with omission of important concepts and poor presentation.	Lacks a clear understanding of the code and AI&ML functionalities, with vague explanations, however a few AI&ML concepts are known with some incorrect explanations.			Indicates a basic understanding of the code and AI&ML functionalities, lacks depth but majority of explanations are correct.			Indicates a satisfactory understanding of the code and AI&ML functionalities with some depth and clarity.			Shows a good, wider understanding of the code and AI&ML functionalities with informative explanations.		Demonstrates a deep and thorough understanding of the code and AI&ML functionalities through insightful discussions.	
Criterion 7 (10%)	No submission	No evidence of best practices understanding of organisation in the code.	Incomplete code with significant errors.	The code is disorganized, lacks best practices, and in-code comments is absent.	The code is complete but lacks some organization and best practices, with major deficiencies and minimal acceptable in-code comments.			The code is somewhat organized and follows minimal basic practices, with minor faults and acceptable in-code comments.			The code is mostly organized and follows basic practices, with satisfactory and acceptable in-code comments.			The code is well-organized and follows good practices with clear in-code comments.		The code is well-structured, highly organized, and follows best practices with exemplary in-code comments.	

6.0 University Grade Descriptor (UGD)

Revised UGD3 – to be implemented for new delivery from September 2024

Level 5 - Generic grade descriptor: relationship of degree classification and equivalent percentage

Class	Category	Mark range	%	General Characteristics
1 st	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 beyond expectation. The ability to make decisions and carry out tasks/processes with autonomy; excellent leadership skills 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. Fully meets expectations set by the industry/employment context.
1 st	High 1st	85 - 92	89	Outstanding/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 autonomy; creative flair and the ability to (re)interpret predefined rules/conventions to select and justify individual working practice; highly developed 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. Broadly 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 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 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 analytical ; evidence of appropriate selection and evaluation of reading/research 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 coordinated performance; tasks are completed with a good level of independent thought; some autonomy is evident; 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	
3 rd	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; strong reliance on available support 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 basic level of independent thought; effective judgements have been made; basic 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	
Zero	Zero	0-5	0	Work not submitted, work of no merit, penalty in some misconduct cases.

7.0 Artificial Intelligence and Academic Integrity – AI&AI

It is important you do not use AI tools to generate an assignment/discussion post 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>