



# Artificial Intelligence

## ASSIGNMENT SPECIFICATIONS

<b>Aims</b>	<ol style="list-style-type: none"> <li>1. Enable students to analyse and employ appropriate Artificial Intelligence (AI) techniques to design intelligent systems and solve problems.</li> <li>2. Enable students to use any relevant tools and technology, such as Python programming, to develop intelligent computer programs.</li> </ol>
<b>Learning Outcomes Assessed</b>	<p>CLO 2: Propose AI techniques and strategies to solve a given problem (A3, PLO9).</p> <p>CLO 3: Produce AI application using programming language or other relevant technology (P4, PLO3).</p>
<b>Outline of Problem</b>	<p>This is a <b>group assignment</b>. All members from each group are required to critically evaluate the current technologies, and then propose a project for the selected AI topic, implement an AI solution to solve the problem in the proposed project using either Python or other relevant tools.</p>
<b>Details</b>	<p>This assignment consists of <b>TWO (2)</b> parts that are related to each other. Each group is required to develop a program using <b>Python</b> or <b>any other relevant tools</b>. Kindly refer to <b>Project Details Part 1</b> and <b>Project Details Part 2</b> for more information.</p> <p>Form a group of <b>TWO (2)</b> to <b>THREE (3)</b> members. Each group has to complete and submit a <b>Documentation (Part 1)</b> and <b>Program/Prototype with source code (Part 2)</b>. Team leader has to compile and submit the deliverables before the due date.</p> <p><b>SIX (6)</b> options of research areas are listed in the following section named <b>TITLE</b>. Each group should select <b>ONE (1)</b> option and is expected to produce ideas that originated from the respective group members, but not to take the work or an idea of someone else (including from the Web) and pass it off as your own. Besides, <b>NO GROUP is allowed to share the same idea</b>. In other words, each group must propose a unique title or solution.</p> <p>The basic requirements for each research area are available in the <b>Titles</b> section (see the following page). It is important to note that, fulfilling those requirements might only be helping you to get an <b>Average</b> or <b>Good grade</b>. In order to achieve an <b>Excellent grade</b>, extra efforts are required such as learning new skills, introducing new ideas, implementing complex AI algorithms, demonstrating the ability to process big data, and/or producing excellent reports with a working prototype.</p>
<b>Titles</b>	<p><b>1. Machine Learning (Supervised)</b></p> <ol style="list-style-type: none"> <li>a) Identify a classification problem to be solved (e.g., predicting disease presence, credit approval and etc).</li> <li>b) Perform a background study on the selected problem and the supervised learning methods to be used (e.g., ANN, SVM, KNN, etc.).</li> <li>c) Search for the respective dataset. You may download a public dataset online or perform your own data gathering if required.</li> <li>d) Perform data pre-processing and data representation to prepare the dataset for model training (e.g., handling missing values, normalization, and feature engineering).</li> </ol>

- e) Each group member must provide a solution using a preferred classification method (e.g., ANN, SVM, KNN).
- f) Compare the results of different classification methods in terms of evaluation metrics such as accuracy, precision, recall, and F1 score.
- g) **Hint: students may refer to the web page for idea and dataset:**
  - <https://archive.ics.uci.edu/datasets>
  - <https://www.kaggle.com/datasets>
  - <https://www.openml.org/search?type=data&sort=runs&status=active>

## 2. Machine Learning (Unsupervised)

- a) Identify a clustering problem to be solved (e.g., customer segmentation, image segmentation, or anomaly detection).
- b) Perform a background study on the selected problem and the unsupervised learning methods to be used (e.g., K-means, MeanShift, DBSCAN, etc.).
- c) Search for the respective dataset. You may download a public dataset online or perform your own data gathering if required.
- d) Perform data pre-processing and data representation to prepare the dataset for clustering analysis (e.g., scaling, encoding, and dimensionality reduction).
- e) Each group member must provide a solution using a preferred clustering method (e.g., K-means, MeanShift, DBSCAN).
- f) Compare the results of different clustering methods based on evaluation metrics such as inertia, silhouette score, or other appropriate metrics.
- h) **Hint: students may refer to the web page for idea and dataset:**
  - <https://archive.ics.uci.edu/datasets>
  - <https://www.kaggle.com/datasets>
  - <https://www.openml.org/search?type=data&sort=runs&status=active>

## 3. Recommender System

- a. Identify a real-life scenario where a recommender system can be applied to suggest products, services, or content. Examples include:
  - i. Recommending products or brands for an e-commerce platform.
  - ii. Suggesting movies or TV shows for a streaming service.
  - iii. Proposing books or academic articles for a library system.
  - iv. Offering tailored services for procurement activities in a business.
- b. Perform a background study on:
  - i. The selected problem/scenario.
  - ii. The type of recommender system to be implemented
  - iii. The expected functionalities and benefits of the system for the chosen scenario.
- c. Each group member must provide a solution using a preferred recommender system (e.g., collaborative filtering, content-based, or hybrid).
- d. Test and evaluate your recommender system to assess its efficiency and accuracy. Use appropriate evaluation metrics, such as:
  - i. Precision, recall, or F1 score.
  - ii. Mean Squared Error (MSE) or Root Mean Squared Error (RMSE) for predicted ratings.
  - iii. User satisfaction through questionnaire.
- e. **Hint: Students may refer to the source below for idea:**
  - <https://github.com/IBM/product-recommendation-with-watson-ml>
  - <https://stackabuse.com/creating-a-simple-recommender-system-in-python-using-pandas/>

	<ul style="list-style-type: none"> <li>• <a href="https://towardsdatascience.com/how-did-we-build-book-recommender-systems-in-an-hour-the-fundamentals-dfee054f978e">https://towardsdatascience.com/how-did-we-build-book-recommender-systems-in-an-hour-the-fundamentals-dfee054f978e</a></li> </ul> <p><b>4. Natural Language Processing</b></p> <ol style="list-style-type: none"> <li>Identify a problem or task related to Natural Language Processing (NLP). Examples include:             <ol style="list-style-type: none"> <li>Sentiment Analysis: Determining whether a piece of text (e.g., tweets, reviews) is positive, negative, or neutral.</li> <li>Text Classification: Categorizing text into predefined categories (e.g., spam detection).</li> </ol> </li> <li>Perform background study on the problem and method to be used.             <ol style="list-style-type: none"> <li>The chosen NLP problem or task.</li> <li>The significance and applications of solving this problem in real-world scenarios.</li> <li>Common methods and techniques used for the task (e.g., Bag-of-Words, TF-IDF, word embeddings, transformers).</li> </ol> </li> <li>Create a web crawler to crawl sample data from a forum/social media or to use the dataset from any reliable website.</li> <li>Preprocess the data to make it suitable for analysis. This may include:             <ol style="list-style-type: none"> <li>Text cleaning (removing stop words, punctuation, or special characters).</li> <li>Tokenization, stemming, and lemmatization.</li> <li>Feature extraction (e.g., TF-IDF, word embeddings using Word2Vec or BERT).</li> </ol> </li> <li>Each member is required to develop and implement a different solution for the task using a preferred NLP method. Examples include:             <ol style="list-style-type: none"> <li>Sentiment Analysis: Use methods such as Naïve Bayes, Support Vector Machine (SVM), or transformer-based models like BERT or GPT.</li> <li>Text Classification: Implement algorithms like Logistic Regression, Decision Trees, or deep learning models.</li> </ol> </li> <li>Compare and evaluate the performance of different NLP models using appropriate metrics, such as: Accuracy, Precision, Recall, and F1 Score for classification tasks.</li> <li><b>Hint: Students may refer to the web page for idea and dataset:</b> <ul style="list-style-type: none"> <li>• <a href="https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html">https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html</a></li> <li>• <a href="http://adilmoujahid.com/posts/2014/07/twitter-analytics/">http://adilmoujahid.com/posts/2014/07/twitter-analytics/</a></li> <li>• <a href="https://www.yelp.com/dataset">https://www.yelp.com/dataset</a></li> <li>• <a href="https://ai.stanford.edu/~amaas/data/sentiment/">https://ai.stanford.edu/~amaas/data/sentiment/</a></li> </ul> </li> </ol>
	<p><b>5. Chatbot Development</b></p> <ol style="list-style-type: none"> <li>Identify a real-life scenario where a chatbot can be utilized. Examples include:             <ol style="list-style-type: none"> <li>Customer support chatbot for an e-commerce platform.</li> <li>FAQ chatbot for a university or organization.</li> <li>Personal assistant chatbot for scheduling and reminders.</li> </ol> </li> <li>Perform a Background Study             <ol style="list-style-type: none"> <li>Research existing chatbot technologies and applications.</li> <li>Explore the capabilities and limitations of chatbots built using machine learning or platforms like Pandorabots, Google Dialogflow, Rasa, or Botpress.</li> </ol> </li> <li>Choose a Development Approach             <p>Option 1: Build the chatbot using machine learning techniques. This may involve:</p> <ol style="list-style-type: none"> <li>Training a Natural Language Processing (NLP) model for intent recognition and response generation.</li> </ol> <p>Option 2: Develop the chatbot using a platform such as:</p> <ol style="list-style-type: none"> <li>Google Dialogflow.</li> <li>Pandorabots.</li> <li>Rasa or similar tools.</li> </ol> </li> </ol>

	<p>d. Define the Chatbot's Functionalities. Specify the features of the chatbot, such as:</p> <ol style="list-style-type: none"> <li>Answering FAQs.</li> <li>Conducting natural conversations.</li> <li>Providing recommendations, booking services, or troubleshooting issues.</li> </ol> <p>e. Each member is required to develop and implement different Chatbot</p> <p>For ML-based development:</p> <ol style="list-style-type: none"> <li>Collect a dataset for training and testing (e.g., conversation data, FAQs).</li> <li>Preprocess the data (cleaning, tokenization, etc.).</li> <li>Train a machine learning or deep learning model for intent classification and response generation.</li> </ol> <p>For platform-based development:</p> <ol style="list-style-type: none"> <li>Use the platform's GUI or APIs to configure intents, entities, and responses.</li> <li>Integrate additional functionalities (e.g., connecting to external APIs).</li> </ol> <p>f. Test the Chatbot. Conduct various tests to evaluate and compare different chatbot's performance:</p> <ol style="list-style-type: none"> <li>Accuracy of intent recognition.</li> <li>Response relevancy and quality.</li> <li>User satisfaction through feedback collection.</li> </ol> <p>g. Evaluate the Chatbot. Use evaluation metrics to measure performance, such as:</p> <ol style="list-style-type: none"> <li>F1 Score, Precision, and Recall for intent classification.</li> <li>BLEU or ROUGE scores for response generation.</li> <li>Usability and user satisfaction ratings.</li> </ol> <p>h. <b>Hint: Students may refer to the webpage for idea and dataset:</b></p> <ul style="list-style-type: none"> <li><a href="https://www.pandorabots.com/">https://www.pandorabots.com/</a></li> <li><a href="https://www.ibm.com/watson/how-to-build-a-chatbot/">https://www.ibm.com/watson/how-to-build-a-chatbot/</a></li> <li><a href="https://www.kaggle.com/code/melkmanszoon/building-a-chatbot">https://www.kaggle.com/code/melkmanszoon/building-a-chatbot</a></li> </ul>
	<p><b>6. Image Processing and Computer Vision</b></p> <p>a. Identify an image processing or computer vision problem to be solved. For example:</p> <ol style="list-style-type: none"> <li>Text detection and Optical Character Recognition (OCR)</li> <li>Object Detection and Classification</li> <li>Pedestrian Detection and Human Action Recognition</li> <li>Face Recognition</li> <li>Image Segmentation</li> </ol> <p>b. Perform background study on the problem and method to be used. Explore the state-of-the-art methods and technologies used to address the problem (e.g., Convolutional Neural Networks (CNNs), YOLO, Mask R-CNN, etc.).</p> <p>c. Search for the respective dataset (you may download the public online dataset or perform data gathering by yourself).</p> <p>d. Data pre-processing and representation such as</p> <ol style="list-style-type: none"> <li>Image resizing, normalization, and augmentation.</li> <li>Label encoding for classification tasks.</li> <li>Annotation for object detection or segmentation tasks.</li> </ol> <p>e. Each member is required to develop using different algorithms such as CNN, Machine Learning, YOLO and etc. Compare the results of different methods in terms of accuracy, precision, recalls, etc. You can also demo the development system in real-life scenario (optional).</p> <p>f. Based on the results, explain the performance of the chosen methods (e.g. compare the advantage, disadvantage, and characteristic of the methods).</p> <p>g. <b>Hint: students may refer to the webpage for idea and dataset:</b></p> <ul style="list-style-type: none"> <li><a href="http://yann.lecun.com/exdb/mnist/">http://yann.lecun.com/exdb/mnist/</a> (Handwritten Digits recognition)</li> </ul>

	<ul style="list-style-type: none"> <li>• <a href="http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html">http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html</a> (Object detection and recognition)</li> <li>• <a href="https://www.kaggle.com/c/cifar-10">https://www.kaggle.com/c/cifar-10</a> (Object recognition)</li> <li>• <a href="http://vis-www.cs.umass.edu/lfw/">http://vis-www.cs.umass.edu/lfw/</a> (Face Recognition)</li> </ul>
<b>Submission Deadlines</b>	<p>Submit <b>Documentation</b> and <b>Prototype Source Code</b> by the <b>07 SEPTEMBER 2025 (Week 11, Sunday, before 11.59pm)</b></p> <p>Please submit your work to Google Classroom. Late submission will be penalized. A demo session to present the prototype is required in week 12 to week 14.</p>
<b>Contribution</b>	<p>This assignment consists of the following <b>TWO (2)</b> components:</p> <ol style="list-style-type: none"> <li>1. <b>Documentation (40%)</b></li> <li>2. <b>Prototype development (60%).</b></li> </ol> <p>Please check <b>Appendix 1</b> and <b>Appendix 2</b> for the assessment criteria.</p>
<b>Academic Integrity and Plagiarism</b>	<p>There must be <b>ORIGINALITY</b> in your work. Thus, do not copy or refer to other group(s). You may only work with your team member(s) to produce the solution of this assignment. You must not share with nor refer to any part of the assignment (including the code) of anyone else except your team member(s) and your tutor.</p> <p>Before submitting your assignment, please make sure that you have complied with TARUMT Plagiarism Policy. Any cheating, attempt to cheat, plagiarism, collusion and any other attempts to gain an unfair advantage in assessment will cause the students concerned to be penalized.</p> <p><b>IMPORTANT:</b> Students found to be dishonest are liable to disciplinary action.</p>
<b>Late Submission &amp; Penalty</b>	<p>Late submission without valid reason will <b>NOT</b> be tolerated. For late submission, there will be a reduction of total marks:</p> <ul style="list-style-type: none"> <li>• Late 1 to 3 days after deadline of submission: <b>Deduction of 10 marks</b></li> <li>• Late 4 to 7 days after deadline of submission: <b>Deduction of 20 marks</b></li> <li>• Late more than 7 days after deadline of submission: <b>Deduction of 100 marks</b></li> </ul> <p>In certain circumstances, a student may be allowed to submit the assignment late with valid reason. S/he must contact the tutor at least one week before the assignment is due. The tutor will evaluate whether the circumstance warrants submitting the assignment late, but no guarantee that the students will not be penalized.</p> <p>Failing to submit the reports and code will lead to failure of the coursework.</p>

## Project Details Part 1: Documentation

<b>Introduction</b>	Your task for this part of the assignment is to identify a problem for the selected AI topic, perform a <b>literature review</b> and propose your respective AI solution(s) that helps in solving the problem in the proposed project.
<b>How to write a documentation?</b>	<p>The documentation for your selected AI topic should contain the following:</p> <p><b><u>Section 1: Introduction</u></b></p> <ul style="list-style-type: none"> <li>• Brief description or introduction with problem statement/background</li> <li>• Objectives/aims</li> <li>• Motivation</li> </ul> <p><b><u>Section 2: Research Background</u></b></p> <ul style="list-style-type: none"> <li>• Background of the applications</li> <li>• Analysis of selected tool with any other relevant tools</li> <li>• Justify why the selected tool is suitable</li> </ul> <p><b><u>Section 3: Methodology</u></b></p> <ul style="list-style-type: none"> <li>• System flowchart/activity diagram</li> <li>• Description of dataset</li> <li>• Description of algorithms</li> <li>• Proposed test plan/hypothesis</li> </ul> <p><b><u>Section 4: Result</u></b></p> <ul style="list-style-type: none"> <li>• Results</li> <li>• Discussion/interpretation</li> </ul> <p><b><u>Section 5: Discussion and Conclusion</u></b></p> <ul style="list-style-type: none"> <li>• Achievements</li> <li>• Limitation and future works</li> </ul> <p><b><u>References</u></b></p> <ul style="list-style-type: none"> <li>• Sources of the dataset and tool(s) used for the development</li> <li>• Articles or other references cited in the text</li> </ul> <p>The report shall be completed using the given Google Doc template in the Google Classroom.</p>

## Project Details: Part 2 – Prototype Development

<b>Introduction</b>	Your task for this assignment is to implement an AI solution using Python or any other relevant tools, perform testing and evaluation on the system and finally present the work to your tutor.
<b>What to hand in?</b>	Submit all of your source code to <b>Google Classroom</b> submission page. <b>EACH team member is required to present their own work, demonstrate the prototype and be ready for a Q&amp;A session in Week 12-14 based on the arrangement.</b>
<b>Format for Deliverable</b>	Compress the entire source code using zip format and submitted by leader.

## APPENDIX 1

## Documentation Assessment Rubrics (40%)

CLO	Item	Missing or Unacceptable (0-4)	Poor (5-9)	Accomplished (10-15)	Good (16-20)
2	Introduction	No or very little discussion on existing problem and the project. The proposed project already exists, or with very minor change.	Little discussion on existing problem and introduction of proposed project. Minor ideas are modified from existing system(s).	Good discussion and evaluation of existing problem and the proposed project. Ideas modified from existing system, with some creative ideas are added.	A very good discussion and evaluation of existing problem and the proposed project. Majority of the ideas are creative.
2	Research Background	Background study are retrieved directly from the literature without any paraphrasing. No discussion or very little of introduction given to the related system or technology.	Background study is lengthy, contents are retrieved directly from the literature without any critical evaluation. Introduction to the related system is given, but no evaluation provided.	Background study is concise and clear, which integrates critical and logical details from the peer-reviewed theoretical and research literature. Brief discussion and evaluation of the related system.	Background study is concise and clear, which integrates critical and logical details from the peer-reviewed theoretical and research literature. A very good discussion and evaluation of the related system.
2	Methodology	The description does not relate the case study. Brief design of proposed method is provided but lack of explanation or irrelevant.	Brief description of system design, with some explanations. Introduction to the related application of the methods is given but lack of examples, understanding or explanation.	System design is well-illustrated, and with clear explanation. Good discussion and evaluation of the methods applied.	System design is well-illustrated, with good explanation. Good discussion and evaluation of the relevant and practical methods applied to the project.
2	Results	Testing methods were missing or inappropriately aligned with data and research design. Results were confusing.	Testing methods were identified but the results were confusing, incomplete or lacked relevance to the research questions, data, or research design.	The testing methods were identified. Results were presented. All were related to the research question and design. Sufficient metric or measurement is applied.	Testing methods and results presentation were sufficient, specific, clear, structured and appropriate based on the research questions and research design. Extra metric or measurement is applied.
2	Discussion and Conclusion	Discussions or answers to the research objectives and results were omitted or confusing. No or very little discussion on limitation and future improvement.	Little discussions were presented. Answers to the research question and results were unclear or confusing. Only little discussion on limitation and future improvement.	Discussions of the results were presented. The research question and system performance were answered and identified. Some discussion on limitation and future improvement were given.	The significance of the results and achievements of objectives were answered and evaluated. Limitations and future improvements of the studies were identified.
Final score = sum of scores/100*40 (base 40%)					

## APPENDIX 2

## Prototype Assessment Rubrics (60%)

CLO	Item	Criteria		
		Poor (0-4)	Accomplished (5-7)	Good (8-10)
3	User interface / Output (10%)	Poor or confusing design of UI or output, which provides inadequate information/outputs. Most of the information/outputs generated are less accurate. Layout of information is not organized.	Adequate information/outputs needed are generated. The information/output generated are accurate but some with errors. Layout of information is organized.	All the necessary information/outputs are generated. All or most of the information/outputs generated are accurate. Minor errors can be ignored. Layout of information is well-organized.
3	Programming (20%)	The end product fails with many logic errors, many actions lacked exception handling. Solutions are over-simplified. Programming skill needs improvement. Minimal validations are provided. Business rules are not validated. (0-8)	Major parts are logical, but some steps to complete a specific job may be tedious or unnecessarily complicated. Program algorithm demonstrates acceptable level of complexity. The student is qualified to be a programmer. Important and necessary validations are provided. (9-15)	Correct and logical flow, exceptions are handled well. Demonstrates appropriate or high level of complex algorithms and programming skills. Thorough and thoughtful validations are provided. All important business rules are validated. (16-20)
3	Degree of completion (10%)	Too much still remain to be done. Basic requirements are not fulfilled. The end product produces enormous errors, faults or incorrect results.	All required features present in the interface within the required scope, but some are simplified. Or one or two features are missing. The system is able to run with minor errors.	All required features present in the interface within or beyond the required scope. No bugs during demonstration.
3	System implementation (10%)	The end product is produced with different system design or approach, which is not related to the initial proposal.	The end product conforms to most of the system design, but some are different from the specification.	The end product fully conforms to the proposed system design.
3	Presentation and on-the-spot coding (10%)	The student is unclear about the work produced, sometimes not even knowing where to find the source code.	The student knows the code whereabouts, but sometimes may not be clear why the work was done in such a way.	The student is clear about every piece of the work done.
Final score = sum of scores				