Assessment Brief

Module Code	COM7039M	Module Lecturers	Dr Nalinda Somasiri (Director) Ms Swathi Ganesan (Lecturer)
Module Title	Machine Learning		
Level	7	**Credit Value**	15
Assessment Title	Creative Artefact- A practical project to design and develop ML prediction model with supporting documentation.		
Workload	N/A		
Assessment Number	1	**of** 1	**Weighting** 100%
Submission Type	Creative Artefact (Anonymous)		
Submission Method	Turnitin within Moodle		
Publication Date	Week commencing 24 October 2023 GMT		
Due Date			
Expected Feedback Date			
Format of Feedback	Written feedback within Turnitin/Moodle		

Programme Learning Outcomes (PLO)
PLOs **7.1-7.7**
7.1 Evaluate computer science concepts and principles and their application to the effective design, implementation, and usability of computer-based systems.
7.2 Apply the findings of advanced scholarship and/or contemporary research and practice to the solution of computer science problems
7.3 Critically evaluate computer science problems, including those at the forefront of field.
7.4 Demonstrate operation within applicable professional, legal, social and ethical frameworks.
7.5 Demonstrate originality and creativity in the solution of computer science problems.
7.6 Recommend, with detailed justification, the appropriate computer science principles and practices to apply to a significant domain-specific activity.
7.7 Apply standards, quality processes and engineering principles to the solution of computer science problems.

Assignment Description

This coursework is designed to demonstrate the broad understanding and knowledge of the module, assessing and evaluating the student's strength and level of analysis. The assessment consists of two tasks.

Task 1:40% You are given set of questions to critically evaluate the subject knowledge and understanding of the concepts behind Machine Learning.
Task 2:60% You are given dataset to work with. You are expected to predict the model and its accuracy using Machine Learning algorithms. Coding part must be accompanied by relevant and detailed explanation.

It is important that the content is underpinned with the inclusion of relevant academic theory, concepts, and models where appropriate, as well as contemporary industrial insights. These should be accurately cited and referenced according to **York St John Harvard Referencing** throughout.

TASK 1: (40 Marks)

1. What are regression and classification ML techniques? Explain in about 500 words how these techniques are helpful in different sectors? **(10 Marks)**
2. Explain overfitting and underfitting? How to combat Overfitting and Underfitting problems? **(10 Marks)**
3. Can you explain the difference between validation set and a test set? **(5 Marks)**
4. Explain SVM algorithm with an example? **(5 Marks)**
5. Explain the Decision Tree algorithm and how it works? **(5 Marks)**
6. What is forward and backward propagation and explain its working? **(5 Marks)**

TASK 2: (60 Marks)

Design and develop the ML prediction model for a given scenario and produce the report

A newly established bank "Bank ABC is eager to develop a **Loan Eligibility Prediction Model based on the applicants'** information. Currently, Bank ABC issuing a paper-based system to check the eligibility of applicants and it is a lengthy process to approve or reject the loan. The bank receives many applications every day and they need more time and staff to check the eligibility whether the applicant is eligible to take the loan or not. Therefore, to save time and cost, the bank is intending to develop an automation system by developing a computer-based loan eligibility prediction model. Assume that the bank has hired you to design and develop a model based on your expertise in Machine Learning. And you will need to design and develop ML models to automate their process.

Data:

The datasets are available to download from ML module in Moodle. You are given two set of data to develop ML prediction model

LoanDatasets.csv – to train the model
LoanDataset_test.csv – to test the model

Code:

We are looking of to show your coding skills using python programming language and Jupyter Notebook.

Challenge:

We would like you to work towards the above scenario to predict loan eligibility of the applicants for YB Bank.

Activity:

You could focus on the following

1. Data collection

2. Data preparation (cleaning, feature selection, feature engineering, etc.,) - Ensure you carry out the data cleaning process for the machine learning model once you explore and visualise the dataset. Provide detailed evidence of the process in the report.

3. Split the dataset for training and testing in your machine learning model. Split the dataset into different portions and observe the model output and see how it impacts the performance of the model

4. Data visualization – Visualise the trend and pattern in data using bar chart, histogram, line chart, scatter plot, etc.,

5. ML modelling (building and testing a predictive model) - Use suitable machine learning technique to model the dataset. Explain the rationale behind choosing your machine learning technique. Evaluate and compare the performance of the model using a different machine learning algorithm to the same dataset.

6. Model performance and evaluation - Evaluate the performance and effectiveness of the model – (with training, testing and new datasets and compare model's output and write findings with valid justification). Evaluate the trained model by predicting some input data.

7. Optimise the model by tuning the hyper parameters to get a better accuracy. (Use R-squared to find the accuracy of the model)

You should support your code implementation with written documentation for Task 2(must cover all the above points)

Note:

1. Both Task 1 and Task 2 (copy the code to the answer file) must be converted to **single pdf file** before submitting.
2. Code in jupyter notebook to be submitted separately in .pynb format.

Additional Information

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.yorksj.ac.uk/students/study-skills/

Assessment Regulations

Please refer to the York St John University Code of Practice for Assessment and Academic Related Matters 2022-23.

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**).

Please see the assessment criteria below.

York St John University Level 7 Assessment Descriptor

Note to Educator - ensure that you <u>contextualise</u> the four criteria within the descriptor below, to <u>each</u> assessment, aligning the content with YSJU's language

	PASS GRADES				FAIL GRADES		
	(100-85)	**(84 - 70)**	**(69 - 60)**	**(59 - 50)**	**(49 - 40)**	**(39 - 20)**	**(19 - 0)**
Overarching indicators:	All learning outcomes/assessment criteria have been achieved to an exceptionally high level, beyond that expected at Level 7, with features consistent with Level 8 (doctoral study).	All learning outcomes/assessment criteria have been achieved to a high standard, and many at an exceptionally high level.	All learning outcomes/assessment criteria have been met fully, at a good or very good standard.	All learning outcomes/assessment criteria have been met.	One or more of the learning outcomes/assessment criteria have not been met.	A significant proportion of the learning outcomes/assessment criteria have not been met.	Most of the learning outcomes/assessment criteria have not been met.
SUMMARY DESCRIPTOR: Learning accredited at Level 7 (Master's) will reflect the ability to display mastery of a complex and specialised area of knowledge and skills, employing advanced skills to conduct research or advanced technical/professional activity, accepting accountability for related decision-making, including use of supervision.							
Criteria	**Characteristics**						
Subject knowledge & understanding (30%)	Exceptional subject knowledge and conceptual understanding at the forefront of the discipline. Authoritative approach to complexity.	Comprehensive subject knowledge and conceptual understanding, informed by recent developments in the discipline, demonstrating reading/research at significant depth/breadth. Informed & confident approach to complexity.	Detailed subject knowledge and conceptual understanding demonstrating purposeful reading/research. Developing awareness of complexity.	Broad subject knowledge and conceptual understanding, demonstrating directed reading/research. Some awareness of complexity.	Reproduction of taught content and/or tendency to describe or report facts rather than demonstrate complex ideas. Any errors or misconceptions are outweighed by the overall degree of knowledge & understanding demonstrated.	Insufficient evidence of knowledge and understanding of the subject and its underlying concepts.	Little or no evidence of knowledge and understanding of the subject and its underlying concepts.
Higher cognitive skills & originality (30%)	Rigorous and sustained criticality, independent thinking and original insight; convincing conclusions and/or application to practice.	Strong and sustained criticality and independent thinking/original insight; persuasive conclusions and/or application to practice.	Detailed criticality and evidence of independent thinking/original insight; logical and sustained conclusions and/or application to practice.	General criticality and some evidence of independent thinking; logical conclusions and/or application to practice.	Limited criticality and independent thought, leading to conclusions and/or application to practice that is poorly supported.	Mainly descriptive and/or inadequately supported conclusions and/or application to practice.	Little or no evidence of criticality and independence of thought.
Advanced technical, professional and/or research expertise (30%)	Exceptional demonstration of advanced technical, professional and/or research expertise. Innovative. Work may challenge the existing	Purposeful, systematic, and sophisticated demonstration of advanced technical, professional and/or research expertise.	Purposeful, systematic, and skilled demonstration of advanced technical, professional and/or research expertise.	Skilled demonstration of advanced technical, professional and/or research expertise.	Developing expertise. Inconsistent demonstration of advanced technical, professional and/or research conduct.	Limited demonstration of advanced technical, professional and/or research conduct.	Little or no demonstration of advanced technical, professional and/or research conduct.

	PASS GRADES				FAIL GRADES		
	(100-85)	**(84 - 70)**	**(69 - 60)**	**(59 - 50)**	**(49 - 40)**	**(39 - 20)**	**(19 - 0)**
	boundaries of knowledge and/or practice.						
Professionalism (10%) [e.g., of information, results of research, ideas, concepts and arguments etc] **and adherence to academic conventions**	Professional, sophisticated/innovative communication, with exceptional clarity and/or audience-engagement, and exemplary academic conventions.	Professional and fluent communication, that holds the attention of its reader/audience throughout and which demonstrates academic conventions that are accurate and relevant to the level of study/beyond.	Fluent and coherent communication, which demonstrates consistent and accurate academic conventions.	Mostly fluent and coherent communication; demonstration of appropriate academic conventions, which may include some errors or inconsistencies.	Communication that is difficult to follow at times because of poor clarity/structure; inconsistent demonstration of academic conventions.	Limited clarity and/or structure in communication, and/or inadequate demonstration of academic conventions.	Highly limited clarity and/or structure in written and/or oral communication. Inadequate demonstration of academic conventions.