Reassessment Brief

M505 Intro to AI and Machine Learning

Winter Semester, 2023

Part I: General Information

Module	M505 Intro to AI and Machine Learning
Term	Semester 1, Quarter 1
Assignment Title	Individual Final Project
Weighting	70% Primary Assessment Task 15% Online Assessments 15% Class Participation
Distributed on:	W/C November 27, 2023
To be submitted on:	January 12th, 2024. 18:00 CET
Submission Method	This assignment must be submitted as a Jupyter Notebook (converted to a *.html file) in the corresponding submission folder to be found on Canvas. You must submit your work with an Assessed Submission Form, which must be completed in full. Please take a screenshot of the signed Assessed Submission Form and add it to the end of your Jupyter Notebook as an image. The assignment will not be accepted by the Registry unless the form is completed correctly.
Length	The submission should contain all the required textual and code sections in a correct, complete, and concise manner. To avoid verbosity, please keep the total number of characters in your submission below 20,000.

Part II: Assessment Details

Primary Assessment Task Topic	You are the newly appointed data scientist of the company. For a given dataset, you are required to build an end-to-end machine learning pipeline in a Jupyter Notebook. Your designed and implemented pipeline will be submitted to the team lead data scientist of the company. The notebook should contain the following information: • A problem statement that elaborates the task. For example, what is the underlying business problem and why is it important? How solving this problem will benefit the company? How would you collect relevant data? How would you formulate this problem as a machine learning task? • A data exploration discussion on the characteristics of the given dataset. For example, how do distributions of the features and the target class look like? Does this dataset

	require any sampling or balancing techniques? Which evaluation metrics are fitting this dataset?
	• Data preprocessing and feature engineering steps to prepare the dataset for learning.
	• A model training step to select the best machine learning algorithm and tune its hyperparameters.
	• A model assessment step to evaluate the final performance of the best trained model.
	• A final discussion on the overall pipeline. For example, what are the overall strengths and limitations of the proposed solution? What are the implications of the results for the business problem? What are your data-driven recommendations for solving the initial business problem? What are the most informative features of your model? Is your model explainable?
Assessment Guidelines	The submitted Jupyter Notebook should contain both textual sections and runnable codes in a rational structure. The texts and codes should be written in a clear and easy-to-follow manner.

All the design decisions should be made in a principled manner. In fact, all the choices should be justified in the notebook, either by explaining the intuition or by conducting empirical experiments.

Be creative. Get inspired from any public documents (e.g., blogs, documentation, open-source projects) but design and implement your notebook yourself. Reproducing another source will lead to plagiarism issues.

Please choose a new dataset that was not used in the exercises. Please clearly mention the URL of your dataset. It allows us later to rerun your notebook, if necessary, as the dataset is accessible via its URL.

The use of generative AI technologies (such as ChatGPT) in your final assignments is not allowed unless the assessment guidelines explicitly clarify, under which terms, you are allowed to use these technologies. Any violation of this rule will result in an investigation of academic misconduct. |
| Purpose | Designing and implementing such end-to-end machine learning pipelines is one of the key responsibilities of data scientists in practice. This assignment is designed to assess your ability to build such pipelines. We are especially interested to see that you can apply various techniques that you have learned in the module in a systematic and principled way. |

Links to Module Intended Learning Outcomes	The assignment relates to the following intended learning outcomes for the module: • Critically evaluate the business contexts that can benefit from machine learning. • Critically analyse and evaluate core machine learning concepts and algorithms, including supervised learning and unsupervised learning. • Critically design and implement machine learning systems for various problems using Scikit-Learn.
Special Instructions	
Additional Assessment Components	GISMA University rewards in-class participation, and engagement with asynchronous content, at a rate of 30% per module. Students participating ≥ 80% (factoring on possible extenuating circumstances) of their synchronous classes as per their due mode of delivery will gain 15% towards their final module mark. Students successfully engaging with asynchronous material on the gamification/microlearning path and completing all summative assessments in the asynchronous environment will equally gain 15% towards their final module mark. Designated asynchronous tasks should be completed by the deadlines specified by the tutors. Do note that all tasks must be completed by the deadline applicable for the principal assessment task. The above also entails that students falling below 80% of participation, although they will still be allowed to submit, will have their final mark capped at 85/100. Equally, if they fail to engage with the asynchronous material and complete the short summative assessments included in specific checkpoints during each term (usually 4), their module mark, irrespective of their engagement and participation in synchronous delivery, will drop by a maximum rate of 15%.

Part III: Marking Criteria / Assessment Criteria

Mark Weight 100%	Fail (0 - 49%) 5,0	Sufficient (50 – 59%) 4,0 - 3,7	Satisfactory (60 – 74%) 2,7-3,3	Good (75-89%) 1,7-2,3	Very Good (90-100%) 1,0-1,3
Marking Criteria	Does not fulfil the requirements	Demonstrates acceptable knowledge and	Demonstrates substantial knowledge	Demonstrates a comprehensive knowledge and	Demonstrates a comprehensive knowledge and

	of	the	understanding of the subject-matter and achievement of learning outcomes at low to average level of performance.	and understanding of the subject-matter and achievement of learning outcomes at average to above average performance levels.	understanding of the subject-matter and achievement of learning outcomes at well above average levels of performance.	understanding of the subject-matter and achievement of learning outcomes at high (highest) levels of performance.

Assessment Criteria	Your primary assessment task will be assessed based on the following criteria: • The correctness, completeness, and conciseness of runnable codes. (35%) • The structure of the report, quality of writing, and critical evaluation of codes and results in the text. (35%)
Notes about Marking	

Part IV: Tips for Successfully Engaging with this Assessment

Answer the Question	It may seem obvious, but make sure you are answering the question you have been set, not the question you would prefer to answer. If the brief has a number of tasks or parts, answer all of them. Parts that involve evaluation or analysis are usually longer and worth more marks than parts that ask for description or explanation. Keep the brief in front of you and check it regularly.
How to Use Assessment Criteria	The assessment criteria document is not usually a guide to the structure of your assignment. Each section of the criteria is not a separate paragraph in your assignment, but qualities that you need to demonstrate throughout. Treat the assessment criteria as a checklist at the end not as a plan at the beginning. Also, the criteria document often tells you what to demonstrate (e.g., critical analysis) but not necessarily how to do it. For how to do it, look back at the skills and activities you have covered in the rest of the module. Above all, remember this is not a test of how much you know or how much you have read about the topic. It is a test of how well you can use your knowledge to answer the specific question set.
Planning and Preparation	Make sure you attend the lectures, especially the first and the last one, where we will be 'unpacking' this assignment in greater detail.
Referencing	GISMA Business School requires that students use Harvard Referencing.
Plagiarism and Cheating	Your attention is drawn to the University's stated position on plagiarism. THE WORK OF OTHERS THAT IS INCLUDED IN THE

	ASSIGNMENT MUST BE ATTRIBUTED TO ITS SOURCE (a list of references and bibliography must be submitted).
	Please note that this is intended to be an individual piece of work. Ensure that you read through your work prior to submission. Action will be taken where a student is suspected of having cheated or engaged in any dishonest practice. Students are referred to the University regulations on plagiarism and other forms of academic misconduct. Students must not copy or collude with one another or present any information that they themselves have not generated.

For more information on Plagiarism, please see the relevant section in your Programme Handbook. |