Softwarica College of IT & E-Commerce
STW7072CEM Machine Learning
Assignment Brief [TBA]

in collaboration with Softwarica College of IT & E-commerce | Coventry University

Module Title: STW7072CEM Machine Learning	**Ind/Group:** Individual	**Cohort:** TBA	**Module Code:** STW7072CEM
Coursework Title: CW			**Handout Date:** TBA
Lecturer: Er. Shrawan Thakur			**Due Date:** TBA
Word Limit: up to 3000	**Coursework Type:** Project Development		**% Of Module Mark** 100%
Submission arrangement online via Softwarica Portal: The link to your GitHub classroom repository containing app source code must be submitted on Softwarica Portal. File types and method of recording: Mark and Feedback date: 3 weeks after submission Mark and Feedback method: written feedback using Softwarica Portal			

Module Learning Outcomes Assessed:

Module Learning Outcomes:

Please read **all** instructions and information carefully.

1. This assignment contributes 100% to your final module mark and will assess the following learning outcomes:
2. Knowledge
3. Knowing a variety of Machine Learning approaches in use in industry or research can help them select and use the best Machine Learning strategies for the challenges that occur.
4. Awareness of recent and upcoming developments in machine learning and its applications.
5. Skills
6. Assess real-world problems and determine which Machine Learning approaches are suitable for their solutions.
7. Apply various Machine Learning models and techniques in the solutions of a range of problems, and characterize the expected performance of a model, and compare with other techniques.

Task:

This assignment will involve the development of practical work, software planning for an intelligent application prototype and the development of that prototype.

You have been issued with exercises associated with lab work throughout the module.

The planning and development of the intelligent application prototype will make use of the knowledge and materials that you have been collecting.

Intelligent application prototype in the python programming language using Jupyter Notebook or Google Colab.

Project Prototype: Machine learning to solve real world classification and clustering problems.

Context:

You were familiar with many machine learning methods, related ideas, and applications during this module. Several classification techniques were studied, including Generalized Logistic Regression, Linear Discriminant Analysis, Optimized K-nearest Neighbor, Bayesian and Statistical Methods, Support Vector Machines, and Decision Trees.
In addition, we examined feature selection and extraction techniques like Principal Component Analysis as well as clustering algorithms like K-means. You must choose an application that relates to a classification, clustering, or anomaly detection problem for this project and investigate the most effective way to use machine learning methods to solve it. You are open to choose any datasets available (**Sources:** https://www.kaggle.com or https://archive.ics.uci.edu/ml/datasets.php and need to apply 3 different classification or clustering techniques on it.

The purpose of this coursework is to

- Analyze the core ideas of machine learning, as well as how they are used.
- Ensure that a dataset is properly prepared, and then assess how well various learning algorithms work on it.
- Learn how to choose machine learning methods to address a real-world classification or clustering issue.

You will be required to:

- Write a report on your own, independently (using your own words to describe the deliverable and to highlight your personal contributions as well).
- If you are unclear about this, ask the professor.
- You are encouraged to regularly submit work in progress for formative feedback to help you improve your final submission.
- READ the five samples at the bottom of the "Module Essentials >> **Schoolworkshoppro:** **https://schoolworkspro.com** " page before you begin. Understanding the components and requirements of the CW is crucial for you.

The following sections should generally be included in the paper:

- Abstract
- The problem is introduced along with a brief literature review of related work in the introduction; if the literature review is longer, it is advised to be its own section, as this would be preferable.
- Problem and Data set(s) description (where you describe in detail the problem. you want to solve and its significance)
- Methods (where you shortly describe the machine learning methods and/or other methods employed to solve the problem)
- Experimental setup (including data pre-processing, feature selection and extraction, classification/clustering parameters)
- Results
- Discussion and Conclusions
- References

These are general section headings that you can modify to fit the application or problem under investigation. Sections discussing fresh and unique developments or modifications of algorithms relevant to your work may be included.

Your code needs to be developed in the python programming language using Jupyter Notebook or Google Colab and submitted as a zip file. You should include a **comprehensive set of instructions** for the successful execution of your code as well as a **short screencast/video of your code running** (or a link to the screencast/video if it is too large to include).

More information of how to write a paper is available at the following link: “Crafting Papers on Machine Learning”, by Pat Langley (which can be found here if the previous link does not work http://www.machinelearning.ru/wiki/images/0/07/Langley00crafting.pdf).

You will need to follow the formatting guidelines of the IEEE Manuscript Template for Conference Proceedings (https://www.ieee.org/conferences/publishing/templates.html).

As mentioned above, you must choose a real-world classification/clustering issue and one or more suitable dataset(s). Additionally, you can use the links below:

UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/;

ICML 2019 accepted papers: https://icml.cc/Conferences/2019/Schedule?type=Poster;

Kaggle competitions: http://www.kaggle.com/competitions;

Stanford machine learning projects: http://cs229.stanford.edu/projects2013.html, http://cs229.stanford.edu/projects2012.html,http://cs229.stanford.edu/projects2011.html, http://cs229.stanford.edu/projects2016.html .

- You will draft a proposal (no longer than one A4 page) outlining the project's title, the members of your group, the problem at hand, and your proposed course of action. You must present your tutor with this plan in order to receive constructive criticism.
- You will need to investigate and read related work in the next 2 days. You are welcomed to submit an individually written short literature review of your findings in order to get formative feedback.
- You must choose, implement, and apply the best machine learning algorithms to the chosen problem throughout the course of the next two to three weeks. If necessary, you must also execute data pre-processing.
- You are welcomed to report your progress to receive regular feedback from the module leader or tutor in the following 2 support weeks, via some scheduled meetings with the module leader.
- You have to write up your final paper, and submit it by the deadline specified on the first page.

Marking Criteria for the paper:

Criterion	Marks

Technical quality 1. Rigor and extent of the experiments. 2. Correct application of the selected algorithms and suitability of the methods. 3. Data preparation - technical quality. 4. Extent of evidence of running the experiments provided in appendices.	 10% 10% 10% 10%
Evaluation 5. Discussion and evaluation of the findings. Why are the outcomes significant? What applications would the findings have for other academics or professionals? 6. Is this a "serious" issue or just a minor "toy" issue? In what ways does the paper improve the field of study?	 15% 5%
7. **Social, ethical, legal and professional considerations** related to the problem in question.	5%
Clarity of the writing: 8. Is there sufficient information for the reader to reproduce the results? Is the language used in the paper good? 9. References and general presentation; Are results clearly presented, with appropriate visualizations?	 10% 5%
Originality: 10. Is there some original approach to the problem, original use of techniques? 11. Is there any (and how much) difference from previous contributions?	 10% 10%

General marking guidelines (for all courseworks)

Mark allocation guidelines to students (to be edited by staff per assessment)

0-39	40-49	50-59	60-69	70+	80+
Work mainly incomplete and /or weaknesses in most areas	Most elements completed; weaknesses outweigh strengths	Most elements are strong, minor weaknesses	Strengths in all elements	Most work exceeds the standard expected	All work substantially exceeds the standard expected

Submission Date: 2023 by 05:00 PM

Notes:

1. You are expected to use the Coventry University APA style for referencing. For support and advice on this, students can contact the Centre for Academic Writing (CAW).
2. Please notify your academic services team and module leader for disability support.
3. The college cannot take responsibility for any coursework lost or corrupted on disks, laptops, or personal computers. Students should therefore regularly back-up any work and are advised to save it on the cloud-based services.
4. If there are technical or performance issues that prevent students submitting coursework through the online coursework submission system on the day of a coursework deadline, an appropriate extension to the coursework submission deadline will be agreed. This extension will normally be 24 hours or the next working day if the deadline falls on a Friday or over the weekend period. This will be communicated via your Module Leader.
5. Collusion between students (where sections of your work are similar to the work submitted by other students in this or previous module cohorts) is taken extremely seriously and will be reported to the academic conduct panel. This applies to both coursework and exam answers.
6. A marked difference between your writing style, knowledge and skill level demonstrated in class discussion, any test conditions and that demonstrated in a coursework assignment may result in you having to undertake a Viva Voce in order to prove the coursework assignment is entirely your own work.
7. If you make use of the services of a proof reader in your work you must keep your original version and make it available as a demonstration of your written efforts.
8. You must not submit work for assessment that you have already submitted (partially or in full), either for your current course or for another qualification of this college, with the exception of resits, where for the coursework, you may be asked to rework and improve a previous attempt. This requirement will be specifically detailed in your assignment brief or specific course or module information. Where earlier work by you is citable, i.e., it has already been published/submitted, you must reference it clearly. Identical pieces of work submitted concurrently may also be considered to be self-plagiarism.

Mark allocation guidelines to students (to be edited by staff per assessment)

GRADE	ANSWER RELEVANCE	ARGUMENT & COHERENCE	EVIDENCE	SUMMARY
First **≥70**	Innovative response, answers the question fully, addressing the learning objectives of the assessment task. Evidence of critical analysis, synthesis and evaluation.	A clear, consistent in-depth critical and evaluative argument, displaying the ability to develop original ideas from a range of sources. Engagement with theoretical and conceptual analysis.	Wide range of appropriately supporting evidence provided, going beyond the recommended texts. Correctly referenced.	An outstanding, well-structured and appropriately referenced answer, demonstrating a high degree of understanding and critical analytic skills.
Upper Second **60-69**	A very good attempt to address the objectives of the assessment task with an emphasis on those elements requiring critical review.	A generally clear line of critical and evaluative argument is presented. Relationships between statements and sections are easy to follow, and there is a sound, coherent structure.	A very good range of relevant sources is used in a largely consistent way as supporting evidence. There is use of some sources beyond recommended texts. Correctly referenced in the main.	The answer demonstrates a very good understanding of theories, concepts and issues, with evidence of reading beyond the recommended minimum. Well organized and clearly written.
Lower Second **50-59**	Competently addresses objectives, but may contain errors or omissions and critical discussion of issues may be superficial or limited in places.	Some critical discussion, but the argument is not always convincing, and the work is descriptive in places, with over-reliance on the work of others.	A range of relevant sources is used, but the critical evaluation aspect is not fully presented. There is limited use of sources beyond the standard recommended materials. Referencing is not always correctly presented.	The answer demonstrates a good understanding of some relevant theories, concepts and issues, but there are some errors and irrelevant material included. The structure lacks clarity.
Third **40-49**	Addresses most objectives of the assessment task, with some notable omissions. The structure is unclear in parts, and there is limited analysis.	The work is descriptive with minimal critical discussion and limited theoretical engagement.	A limited range of relevant sources used without appropriate presentation as supporting or conflicting evidence coupled with very limited critical analysis. Referencing has some errors.	Some understanding is demonstrated but is incomplete, and there is evidence of limited research on the topic. Poor structure and presentation, with few and/or poorly presented references.
Fail **<40**	Some deviation from the objectives of the assessment task. May not consistently address the assignment brief. At the lower end fails to answer the question set or address the learning outcomes. There is minimal evidence of analysis or evaluation.	Descriptive with no evidence of theoretical engagement, critical discussion or theoretical engagement. At the lower end displays a minimal level of understanding.	Very limited use and application of relevant sources as supporting evidence. At the lower end demonstrates a lack of real understanding. Poor presentation of references.	Whilst some relevant material is present, the level of understanding is poor with limited evidence of wider reading. Poor structure and poor presentation, including referencing. At the lower end there is evidence of a lack of comprehension, resulting in an assignment that is well below the required standard.
Late submission	0	0	0	0