Faculty of Computing Engineering and Science Assessment Brief

Module Title: Computational Applications OF Artificial Intelligence

Module Code: CS4S773

Module Leader/Tutor: Dr Mabrouka Abuhmida

Assessment Type: Asynchronous Assessment

Assessment Title:

Weighting: 50%

Word count/duration/equivalent: NA

Submission Date: Wednesday 30-10-24

Return Date: in 20 working days from the submission date

## Assessment Description

**Objective**

You are a data scientist working with healthcare data containing patient records. Your goal is to analyse and generate a report on the dataset, perform data cleaning and transformation, and develop a machine learning model to predict test results (Normal, Abnormal, Inconclusive).

**Scope: This assignment will cover the following key tasks:**

**Data Cleaning**: Identifying and handling missing or ambiguous values, outliers, and irregularities in the dataset.

**Data Transformation:** Applying techniques to preprocess and transform the data for better model performance.

**Exploratory Data Analysis (EDA):** Visualising trends, distributions, and relationships between variables, focusing on healthcare trends across demographics, medical conditions, and other features.

**Feature Engineering**: Creating new features that could enhance the predictive power of your model.

**Data Visualization**: Visualize the distribution of key features like age, gender, and test results. Explore patterns in the dataset features.

**Feature Engineering:** Create new features by combining or transforming existing ones.

Investigate if specific combinations of medical conditions or medications lead to certain test results.

**Model Development & Training**: The Blackbox machine learning model predicts the "Test Results" category. Evaluate your model using appropriate performance metrics to answer the following questions.

* Given the often-imbalanced nature of the datasets, what strategies can you implement to balance it out? What are the trade-offs of each?
* Which features do you believe will be most relevant for a machine learning model? How would you validate these assumptions?
* How does the pre-processed data impact its performance?

## Guidance on Format of Assessment

Note: Students are reminded **not** to include this assignment brief with the assignment submission

Please name your files using your student number.

You will need to submit your notebook code with your report as a PDF file using the same submission link- use a ZIP file.

**Technical Report Guidelines:**

* Title: Make it relevant and engaging.
* Introduction: Contextualize the bank fraud detection problem.
* Dataset Synopsis: Summarize key dataset attributes and their potential significance.
* Preprocessing Journey: Narrate your journey through infographics or flowcharts.
* Visuals: Prioritize clarity and relevance in your visualisations.
* Insights & Revelations: Share notable patterns or findings from your analysis.
* Conclusions: Reflect on the implications of your findings and their potential real-world impacts.
* References: Properly attribute the dataset and any other pivotal resources.

## Learning Outcomes Assessed

* To demonstrate knowledge and understanding of the essential facts, concepts, and principles of programming within a mathematical context.
* To utilise essential facts, concepts, principles and theories in the analysis, specification, design, planning, documentation, implementation, and evaluation of solutions.

## Marking Criteria/Rubric

Note: All grades are provisional until they are ratified by the exam board

The provided weightings for the marking criteria are guidelines. Given the variability in student submissions, the lecturer have the flexibility to adjust these weightings to best evaluate each work's unique strengths and challenges. This way we ensure that assessments prioritise quality and originality over strict adherence to predefined weights**.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Criteria | Beginning | Developing | Proficient | Accomplished | Excellent | Exceptional |
| Weight | 5% | 30% | 50% | 60% | 80% | 100% |
| Feature Engineering (20%) | No clear evidence of feature creation. | Basic features extracted without much novelty. | Some relevant features were extracted, but lacked innovation. | Good set of features, with a mix of basic and innovative. | Strong feature set, demonstrating creativity and relevance. | Exceptional and innovative feature sets tailored perfectly to the problem. |
| Feedback | Needs significant improvement in understanding the relevance of features. | Shows a basic understanding but needs to explore more innovative approaches. | Decent effort but should delve deeper into the domain for feature engineering. | A good mix of features but can incorporate more domain-specific insights. | Very well done, just a step away from a perfect set. Explore more correlations. | Perfectly executed. Demonstrated excellent domain knowledge and creativity. |
| Critical Approach  (20%) | No evidence of critical questioning or exploration. | Asks surface-level questions without diving deep. | Some good questions were posed but lacked depth in seeking answers. | Consistent inquisitive approach but room for deeper exploration. | Almost consistently deep and meaningful questions. Just minor improvements are needed. | Outstanding depth in questioning with thorough exploration for answers. |
| Feedback | Needs to adopt a more inquisitive approach. | Dig deeper, don't stop at the first answer. | On the right track but push for deeper insights. | Commendable effort; just refine the depth of your inquiries. | Excellent! Just fine-tune some aspects. | A perfect balance of curiosity and depth. Exceptionally well-done. |
| Data Handling  (20%) | Struggles with basic data preprocessing. | Manages common issues but struggles with complex nuances. | Handles most preprocessing tasks but misses some nuances. | Proficient in handling common data issues, with few oversights. | Almost perfect data handling with minimal errors. | Mastery in preprocessing and handling all dataset nuances. |
| Feedback | Start with understanding the basics of data preprocessing. | Improve on handling outliers and missing values. | Focus more on nuances and edge cases. | Strong preprocessing skills; just refine and recheck your steps. | Almost perfect, just a few oversights. Double-check your work. | Exceptional skills showcased in handling the dataset. |
| Visualization  (10%) | No clear or coherent visualizations were presented. | Basic visualizations without much clarity or coherence. | Some good visuals but lacks a clear narrative. | Visuals are clear and coherent but could benefit from refinement. | Almost perfect visuals with a minor scope for improvement. | Outstanding visual storytelling, with clarity and coherence. |
| Feedback | Invest time in learning effective visualization techniques. | Focus on clarity; ensure visuals tell a story. | Good effort but align visuals with the narrative. | Well-executed; just refine it for more coherence. | Exceptional work; just fine-tuning to perfection. | Visuals are a perfect blend of clarity, coherence, and creativity. |
| Insights  (15%) | Lacks depth and applicability of findings. | Basic findings with limited real-world relevance. | Decent insights but limited in scope or depth. | Good depth in findings with some novel insights. | Deep, meaningful insights with just minor room for improvement. | Exceptional depth and breadth in findings with strong real-world applicability. |
| Feedback | Delve deeper into data for meaningful insights. | Good start, but focus on the applicability of findings. | Align your insights with real-world scenarios. | Commendable insights; push for even deeper revelations. | Near perfect, refine for broader applicability. | Demonstrated excellent analytical skills with perfect insights. |
| Code Quality  (15%) | Code is difficult with no comments or structure. | Some organizations lack efficiency and comments. | The code is decently structured but lacks efficiency or comments. | Good code quality with minor issues in efficiency or commentary. | Almost perfect code: minor tweaks are needed for cleanliness or comments. | The code is clean, efficient, and well-commented. Demonstrates mastery. |
| Feedback | Organize your code and always comment on your steps. | Improve code efficiency and add more comments. | Streamline your code and provide insightful comments. | Good code quality; focus on refining and commenting more. | Just a step away from perfection. Refine and recheck. | Exceptional coding skills demonstrated. Well-executed! |

## Submission Details: Wednesday 30-10-24

## What happens next?

Your marked assessment should be available 20 working days after submission. However, please be advised that this may be subject to change in the event of Bank Holidays, University Closure or staff sickness. If there is something about the feedback you have been given that you are unclear about, please see your module tutor.

**Mabrouka.abuhmida@southwales.ac.uk**

## Feedback Method

Individual feedback will be provided via the Blackboard rubric, offering specific written insights on your work. General feedback will be shared verbally in class, allowing for interactive discussions. Ensure you engage with both to get a thorough understanding of your performance.

## Late Submission

Submissions made within five working days after the deadline will be accepted but capped at a maximum mark of 40.

Any submissions beyond the five-day period will receive a mark of zero.

If you face extenuating circumstances preventing timely submission, please apply for an extension in advance to avoid any penalties. Documentation or proof may be required to grant such requests.

## Extenuating Circumstances

[https://advice.southwales.ac.uk/a2z/extenuating-circumstances](https://advice.southwales.ac.uk/a2z/extenuating-circumstances/)

## Referencing, Plagiarism and Good Academic Practice

[https://advice.southwales.ac.uk/a2z/referencing-plagiarism-and-good-academic-practice](https://advice.southwales.ac.uk/a2z/referencing-plagiarism-and-good-academic-practice/)

## Learning Support Resources

[https://studyskills.southwales.ac.uk](https://studyskills.southwales.ac.uk/)

## Your Assessment Queries

To further support your understanding of the assessment criteria and expectations, we will be hosting "Assessment Surgeries". These sessions will delve into the assessment's details, allowing you to ask questions, clarify doubts, and get insights into what is expected for a successful submission.