WIA1006/WID3006 Machine Learning: Course Assignment

Overall mark allocation: 20%

Contents

1 Preliminary Notes: Depression Indicator

The theme of the Machine Learning Hackathon is **Depression Indicator**.

Choose, *define*, and *design* a specific depression indicator that can be implemented using Machine Learning.

Examples:

- Can you compute the probability of a person having depression symptoms?
    - Compute depression probability.
- Can you classify whether a person is having mild or severe depression?
    - Depression classifier (mild or severe) based on symptoms/behaviors.
- Can you recommend a few activities to overcome depression?
    - Depression treatment recommendation, i.e., consult psychiatrist, society engagement.
- Etc.

Marks shall be awarded based on a specific rubric (*Section 2*), aimed at evaluating the overall level of understanding of the students behind the machine learning techniques used to solve the problems.

Machine learning techniques are not limited to the techniques presented in class (i.e. can use other methods such as XGBoost or Random Forest).

IMPORTANT:

Note that, there are **limited** datasets for the different problems on depression. Thus, you may need to develop or collect your own datasets for the problems that you have chosen. There is even a possibility of developing a **pseudo-dataset** to demonstrate your proposed problem & and solution.

2 Rubric

Choice of assignment:

Criteria/ Marks Allocations	1-8 marks	9-15 marks	16-25 marks
Validity of the problem from a machine learning aspect, and the choice of data to support the problem.	Minimally valid with limited data (5)	Adequately valid and convincing data(15)	Highly valid and convincing choice/design of data (25)
Initial flowchart of the proposed solution, gantt chart and tasks distribution among group members for the project (25 marks)	Showcases bare minimal understanding of the problem.	Showcases clear understanding of the problem	Showcases very clear understanding of the problem and well-designed flowchart
Report Evaluation 1: (Structure and Clarity, 25 marks) • Introduction to Problem • Hypothesis made for the problem • Project Objectives • Methodology • Elaboration on Data & Features Used • Results and discussions • Suggestion for future works	Report meets the bare minimal standard of structure and clarity. Contains minimal data exploration Methodology lacks clarity and requires refinement. Hypothesis is validated but not in entirety via the results and discussions.	Report is well written in terms of format and presentation. Contains enough data exploration. (e.g. plot of histogram, correlation matrix etc.) Methodology is adequately clear Hypothesis is adequately validated via the results and discussions.	Report is very well written and presented in a clear and concise manner. Contains a well-presented amount of data exploration and explanation. (e.g. plot of histogram, correlation matrix etc.) Methodology is presented clearly and convincing Hypothesis are validated and correlates well with the results, and highly convincing.

Report Evaluation 2: (Technical Content, 25 marks) ● Discusses all processes involved to solve the problem. ● Presents the background theory clearly for the reader for the different techniques used in solving the problem. ● Experimental protocol is well defined and presented (usage of confusion matrix or other metrics to determine the performance) ● Commented source code	Technical content meets the bare minimal standard required for a technical report. Results are presented at a bare minimum. Source code of the demo is commented at a minimal level.	Technical content presents well on all the methodologies, processes and contains enough background knowledge to help users understand the overall solution to the problem. Results are well presented and accompanied with performance metrics (e.g. confusion matrix, area under the ROC curve etc.) Source code is well structured and commented.	Technical content is of highest quality and presents excellently on all the methodologies, processes and contains enough background knowledge to help users understand the overall solution to the problem. Results are well presented with well accompanied performance metrics and limitations of the results are well discussed (e.g. confusion matrix, area under the ROC curve etc.) Source code is of high quality and well commented as well as structured.

Deliverables:

- Report & PPTX & 5 mins video
- Deliverables detailing the solution to the problem that your group had chosen.

3 Hackathon Timeline

- Kick-off Week: Week 8
    - Register group members.
    - Initial design of the problem containing the problem statement and proposed hypothesis.
    - Gantt Chart and tasks distribution between members
- Submission of Assignments: **FRIDAY, 17/6/2022 (W13)**
- Voting of projects: **18-19/6/2022**
- Pitching Day: **MONDAY, 20/6/2022 (W14)**

Cash Prizes

1. First Place : **RM350**
2. Second Place: **RM100**
3. Third Place : **RM 50**

Note:

- Evaluation of all assignments for all groups, shall be done based on Rubric 2 using the assignments submitted in Week 13.
- Winners of hackathon shall earn bragging rights.

4 Soft skills Evaluation

This assignment shall evaluate soft skills elements.

COMMUNICATION SKILLS (CS1, CS2, CS3) *

- REPORT (Report Writing, Review on related cognitive theory(ies)).
- PRESENTATION (Flow of discussion, Presentation (language/fluency/idea coherency), Teamwork, Effort and Q&A skills).

*Using the standard CS rubric. (note: applicable to Online presentation via Zoom / Meet)

Critical thinking and Problem solving (CT1, CT2, CT3)

- Creative and critical thinking (Innovation, connecting, synthesizing knowledge and transforming ideas into new forms/solutions) for Individual Project.

Moral and Professional Ethics (EM1, EM2)

- Moral & Professional Ethics – ethics in presentation of results in both reports and live presentations.