جامعة الحسين التقنية
Al Hussein Technical University

ASSIGNMENT BRIEF

HTU Course No: 10204350	**HTU Course Name:** Machine Learning
BTEC Unit Code: J/615/1662	**BTEC UNIT Name:** Machine Learning

Version: 3

Student Name/ID Number/Section	
HTU Course Number and Title	10204350 Machine Learning
BTEC Unit Code and Title	J/615/1662 Machine Learning
Academic Year	2023-2024 Fall
Assignment Author	Rami Al-Ouran
Course Tutor	Rami Al-Ouran - Rami Ibrahim
Assignment Title	Assignment 2 ML Project
Assignment Ref No	2
Issue Date	21/12/2023
Formative Assessment dates	From 25/12/2023 to 11/01/2024
Submission Date	30/01/2024
IV Name & Date	Raneem Qaddoura 20/12/2023

Submission Format

Your submissions should be in the form of a soft copy via the eLearning school system (https://elearning.htu.edu.jo). The report should be:

- Written in professional style format.
- The report must be submitted in MS Word format (not pdf).
- Include a cover page: Student name, Class, Assignment Title, Date.
- Your work must be supported with references using the Harvard reference system.
- Any plagiarism shall result in failing the course.

Oral Exam: An oral discussion will be scheduled with your instructors to assess your understanding of the assignment.

Resubmission: If you lose 3 Ps or more, you will not be eligible for resubmission.

Submission checklist:

- No compressed files or folders (no.zip or .tar extensions)
- Technical report file named as “FirstName_LastName_ML_part 2.docx” and in word format.
- A jupyter notebook named as “FirstName_LastName_ML_Notebook_part_2.ipynb”.
- A python file of the notebook named as “FirstName_LastName_ML_Notebook_part_2.py”.
- Declaration form named as “FirstName_LastName_ML_Declaration_form_Fall 23-24.docx.

Unit Learning Outcomes

LO3 Develop a machine learning application using an appropriate programming language or machine learning tool for solving a real-world problem.

LO4 Evaluate the outcome or the result of the application to determine the effectiveness of the learning algorithm used in the application.

Assignment Brief and Guidance

You have been hired as a new intern at a Jordanian startup company which plans to establish a new data science department with emphasis on machine learning applications. You have been tasked with developing machine learning applications using an appropriate programming language or machine learning tool to solve different problems from the real world.

Your first task is to analyze a Room Occupancy Estimation problem which helps in optimizing energy usage. The link to the dataset is available at: https://archive.ics.uci.edu/dataset/864/room+occupancy+estimation and the Journal paper link: https://www.semanticscholar.org/paper/Machine-Learning-Based-Occupancy-Estimation-Using-Singh-Jain/e631ea26f0fd88541f42b4e049d63d6b52d6d3ac. You are requested to analyse the data using the following algorithms: random forest, SVM , gradient boosting, and XGBoost. For each model you are requested to find and report the best model parameters.

Data Description: The "data set for estimating the precise number of occupants in a room using multiple non-intrusive environmental sensors like temperature, light, sound, CO2 and passive infrared sensor (PIR)." There are 10129 instances and 18 features. The target variable is "Room_Occupancy_Count".

Prepare a technical report with the following sections.

1. Introduction

- Describe the problem you are addressing and why is it important?
- Describe the dataset's source, collection method, attributes, size, and domain.
- Describe the learning problem you are trying to solve.
- How did you prepare training and test data before implementing the machine learning models?

2. Methods

- Explain why the provided models are appropriate to solve this problem.
- Demonstrate how you will test the machine learning applications using a range of test data and explain each stage of this activity (Apply k-fold cross-validation).
- Explain in detail the machine learning algorithms you are using to address this problem.

3. Evaluation

Evaluate the effectiveness of the learning algorithms used by answering the following questions:

- What performance measures did you use to evaluate the effectiveness of your models?
- Why did you use these metrics?
- Evaluate how, based on the performance measures, you were able to enhance the model.

4. Results and Discussion

- Discuss the reliability of your results and whether they are balanced, overfitting, or underfitting.
- Analyse the result of the applications to determine the effectiveness of the algorithms.
- Draw conclusions regarding the strengths and weaknesses of the different algorithms.

- Identify further enhancements which can be done in the future? Discuss any limitations and future improvements of your project.

5. References

You need to write the references here using the Harvard style of referencing.

Learning Outcomes and Assessment Criteria			
Learning Outcome	**Pass**	**Merit**	**Distinction**
LO3 Develop a machine learning application using an appropriate programming language or machine learning tool for solving a real-world problem.	**P5** Choose an appropriate learning problem and prepare the training and test data sets in order to implement a machine learning solution **P6** Implement a machine learning solution with a suitable machine learning algorithm and demonstrate the outcome.	**M3** Test the machine learning application using a range of test data and explain each stages of this activity.	**D2** Critically evaluate the implemented learning solution and it's effectiveness in meeting end user requirements
LO4 Evaluate the outcome or the result of the application to determine the effectiveness of the learning algorithm used in the application.	**P7** Discuss whether the result is balanced, underfitting or over-fitting **P8** Analyse the result of the application to determine the effectiveness of the algorithm	**M4** Evaluate the effectiveness of the learning algorithm used in the application.	

STUDENT ASSESSMENT SUBMISSION AND DECLARATION

When submitting evidence for assessment, each student must sign a declaration confirming that the work is their own

Student name:	**Assessor name:**	
Issue date: 21/12/2023	**Submission date:** 30/01/2024	**Submitted on:**
Programme: Computing		
HTU Course Name: Machine Learning **HTU Course Code:** 10204350	**BTEC Course Title:** Machine Learning **BTEC Course Code:** J/615/1662	
Assignment number and title: 2, Assignment 2 ML Project		

Plagiarism:

Plagiarism is a particular form of cheating. Plagiarism must be avoided at all costs and students who break the rules, however innocently, may be penalised. It is your responsibility to ensure that you understand **correct referencing practices.** As a university level student, you are expected to use appropriate references throughout and keep carefully detailed notes of all your sources of materials for material you have used in your work, including any material downloaded from the Internet. Please consult the relevant unit lecturer or your course tutor if you need any further advice.

I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.

Student Name: **Student Signature:**

Date: