جامعة التقنية
والعلوم التطبيقية
University of Technology
and Applied Sciences

Mini Project Guidelines

Index

Aim:

The aim of this mini-project is to apply machine learning techniques to solve a real-world problem by working through the data processing, model building, evaluation, and analysis phases. This will help you gain hands-on experience in the practical implementation of machine learning concepts and workflows.

Instructions:

1. **Project Topic:**
   - Choose a dataset relevant to a real-world application (e.g., healthcare, finance, retail, or social media).
   - Define a clear machine learning problem (e.g., classification, regression, clustering).
2. **Data Collection:**
   - You may use publicly available datasets from sources such as Kaggle, UCI Machine Learning Repository, or any other open-source platform.
   - Ensure the dataset is of a reasonable size to perform meaningful analysis (minimum 500 records and 5 features).
3. **Data Preprocessing:**
   - Handle missing values, outliers, and normalize/standardize the data as necessary.
   - Perform feature selection and engineering to improve model performance.
   - Visualize key features for better understanding (e.g., correlation heatmaps, pair plots).
4. **Model Building:**
   - Use two machine learning algorithms to model your data (use one dataset only) *(The algorithms should not have been implemented in lab activities. The students are free to explore algorithms beyond the course).*
   - Train and test your models using cross-validation or a train-test split.
   - Implement hyper parameter tuning for better model optimization.
5. **Model Evaluation:**
   - Evaluate models using appropriate metrics such as accuracy, precision, recall, F1-score, or RMSE (depending on the problem type).
   - Compare the performance of the different models and provide reasoning for why one model performs better than others.
6. **Analysis and Conclusion:**
   - Discuss the insights gained from the model outcomes.
   - Highlight the practical applications of your findings and their implications for the domain of choice.
7. **Report Submission:**
   - Submit a written report summarizing the problem, methodology, results, and conclusion.
   - Attach your Python code as a separate file or notebook for reproducibility.

Marks distribution

Sl. No.	Category	Marks
1.	Report cum code	13
2.	Presentation	3
Total		16

Mini Project Report Structure

Sl. No.	Content	Description
1.	Title Page	• Title of the project • Name of the author(s) • Date
2.	Abstract	• A brief summary of the project objectives, methodology, key findings, and conclusions. Typically, the abstract should not exceed 250 words.
3.	Table of Contents	• List of sections and subsections with corresponding page numbers.
4.	Introduction	• Background information on the problem domain. • Objectives of the project. • Importance of the problem being addressed. • Brief overview of the methodology and approach.
5.	Data Collection and Preprocessing.	• Description of the dataset used in the project. • Details about data collection methods. • Data preprocessing steps such as cleaning, normalization, handling missing values, and • feature engineering
6.	Methodology	• Overview of the machine learning algorithms or techniques employed. • Description of the model architecture (if applicable). • Explanation of parameter tuning and model selection processes. • Details about the evaluation metrics used to • assess model performance
7.	Model Evaluation	(Refer Report cum Code marking rubrics)
8.	Analysis and Conclusion	(Refer Report cum Code marking rubrics)

Report cum Code marking rubrics

Sl. No.	Criteria	Points	Hints for evaluation
1.	Problem Definition	1	• Clarity of the machine learning problem and alignment with real-world use cases.
2.	Data Preprocessing	1	• Handling missing values, outliers, feature engineering, and visualization.
3.	Model Building	4	• Suitability of selected algorithms, application of hyper parameter tuning, and cross-validation. *Attach your error free Python code as a separate file and notebook for reproducibility*.
4.	Model Evaluation	3	• Use of appropriate metrics, comparison of models, and thorough evaluation by showing the model results
5.	Analysis & Conclusion	4	• Interpretation of model results, practical insights, and application to the problem domain.
Total		13	

Presentation marking rubrics

Sl. No.	Criteria	Points	Hints for evaluation
1	Technical Understanding	1	• The presenter demonstrates a deep understanding of the technical aspects of the project, including data preprocessing, model development, and evaluation.
		1	• Code and complex concepts are explained clearly and concisely, demonstrating proficiency in the subject matter.
		1	• The presenter responds confidently to technical questions
Total		3	