UNIVERSITI MALAYA
UNIVERSITY OF MALAYA

PEPERIKSAAN IJAZAH SARJANA KEPINTARAN BUATAN
EXAMINATION FOR THE DEGREE OF MASTER OF ARTIFICIAL INTELLIGENCE

SESI AKADEMIK 2025/2026 : SEMESTER I
ACADEMIC SESSION 2025/2026 : *SEMESTER I*

WQF7009 : Kebolehjelasan Kepintaran Buatan
: *Explainable Artificial Intelligence (XAI)*

Jan/Jan 2026 Masa : 5 Minggu
Time : 5 Weeks

ARAHAN KEPADA CALON :
INSTRUCTIONS TO CANDIDATES :

Calon dihendaki menjawab **SEMUA** soalan (40 markah)
*Answer **ALL** questions (40 marks).*

Alternative Assessment (40%)

Students are required to identify and access a suitable dataset from an online data portal. The dataset could be related to **Environmental, Social, and Governance (**ESG) topics, such as social issues, healthcare, finance or environment.

After identifying a dataset, the students shall study the background of the dataset by understanding the meaning of all data features, and the target variable's nature (either classification or regression).

XAI Tool References:

GitHub - SquareResearchCenter-AI/BEExAI: Benchmark to Evaluate EXplainable AI

Welcome to XAIB! — xai-benchmark 0.4.0 documentation

Part 1:

Task 1: Propose an XAI-driven solution for the selected domain (20%)

Include:

- Why explainability matters in this domain
- Target stakeholders (domain specialists, regulators, end users)
- Recommended XAI strategy (which method → which stakeholder → why)
- Potential impacts:
    - Trust & transparency
    - Risk mitigation / safety
    - Bias detection / fairness
    - Regulatory compliance
    - Decision-making support

Part 2:

Select **One(1)** machine learning model from the list:

- Convolutional Neural Networks (CNN) classifier
- Random Forest
- XGBoost

- Transformer-based model
- Or other suitable models

Task 2 Build and evaluate a predictive model (black box preferred) (5 %)

- Dataset description
- Model details and performance metrics

Task 3 Select and Apply XAI models (10%)

For the selected model in Task 2:

1. Explain the **explainability challenges** of the model.
2. Select **One** Feature Importance explanation method & **One** Counterfactuals explanation method and categorise **selected methods** under:
    - Local vs Global
    - Post-hoc vs Intrinsic
    - Model-specific vs Model-agnostic
3. **Justify** why these XAI techniques best match the model and its data modality.
4. **Apply** the selected Feature Importance explanation method & Counterfactuals explanation method.
5. Analysis, evaluate and present your XAI explanation. Visual outputs are required (heatmaps, feature importance, textual explanations, etc.).

Task 4 Knowledge Assessment (5 marks)

1. **Differentiate** the usage of Concept-based Explanation from Counterfactual-based Explanation by giving examples based on Task 2.
2. **Explain** the advantages & disadvantages of Interpretable complex models with examples.

END