Assessment Tools

Introduction

As part of our Mozilla RCC Project, we have developed a structured assessment framework to help participants gain practical knowledge and hands-on experience in open-source development, security, and AI fairness. This document outlines the assessment techniques we use to track progress, encourage collaboration, and promote ethical innovation in our project.

1. Quizzes

- Tools: Google Forms, Mentimeter.
- Short quizzes on Mozilla RCC principles, XAI, and AI ethics to assess participants' understanding of key concepts and responsible computing principles.

2. Hands-on Lab Exercises

- Evaluate participants' ability to apply open-source development and responsible computing techniques.
- Tools: Google Colab, GitHub, Google Teachable Machine.
- Participants complete assignments where they implement case studies related to Explainable AI, ensuring hands-on exposure to responsible AI development practices.

3. Google Teachable Machine

- Objective: Introduce participants to machine learning model training using Google's no-code AI platform.
- Tools: Google Teachable Machine.
- Participants engaged in a hands-on session where they create, train, and deploy an ML model for image or speech recognition, fostering an intuitive understanding of AI fairness and transparency.

4. Case Study Analysis

- Objective: Assess the ability to apply responsible computing principles to real-world problems.
- Tools: Written reports, presentations, GitHub repositories.

5. Instructor Feedback & One-on-One Reviews

- Objective: Provide personalized guidance on Responsible computing practices.
- Tools: Office hours, feedback sessions.
- Reviewing participants' contributions.

6. Explainable AI Dashboard

- Enable participants to visualize and interpret AI model decisions using explainability techniques.
  Tools: SHAP, LIME, ExplainerDashboard
- Participants interact with an Explainable AI (XAI) dashboard that provides insights into model predictions. The dashboard includes:
  - **Feature Importance Visualization:** Highlights key features affecting model decisions.
  - **SHAP & LIME Explanations:** Offers local and global model interpretability.
  - **Model Fairness Analysis:** Identifies biases in AI predictions.
  - **Interactive Insights:** Allows participants to tweak input data and observe changes in model predictions.
- By incorporating this dashboard, participants develop a deeper understanding of AI fairness, interpretability, and ethical decision-making in machine learning models.