**Risk Prediction and Service Mapping for Vulnerable Populations in Afghanistan**

**Group 8**  
Rafiullah Samim

Halima

Laila

**1. Project Idea:**

**Risk Prediction and Service Mapping for Vulnerable Populations in Afghanistan**

* **Problem:** Vulnerable populations in Afghanistan face heightened risks of violence, displacement, and lack of access to essential services. These risks are poorly documented, leading to inefficient resource allocation by humanitarian organizations.
* **Goal:** To develop a machine learning-based risk prediction model using protection monitoring surveys and create a service mapping app that provides real-time information on available services, such as shelters, healthcare, and legal aid, in high-risk areas.

**2. Relevance to Sustainable Development Goals (SDGs):**

This project aligns with the following SDGs:

* **SDG 5: Gender Equality** – Promotes the safety of women and girls by identifying high-risk areas and improving access to support services.
* **SDG 16: Peace, Justice, and Strong Institutions** – Contributes to reducing violence, enhancing access to justice, and strengthening institutions that provide aid.  
  By enabling data-driven decision-making and service accessibility, the project addresses critical barriers to safety and support for vulnerable groups.

**3. Literature Examples:**

1. **"Predicting Spatial Distribution of Crime Using Machine Learning"**
   * This study used machine learning to analyze historical crime data and predict crime-prone areas. It employed classification algorithms like Random Forest and Logistic Regression.
   * **Relevance:** It demonstrates how predictive models can guide resource allocation to prevent incidents, similar to identifying high-risk areas in Afghanistan.
2. **"Humanitarian Data Exchange and Mapping for Refugees"**
   * This project mapped services for refugees in conflict zones using geospatial data combined with organizational service directories. It allowed users to locate resources efficiently.
   * **Relevance:** It highlights the importance of service mapping in crisis settings and inspires the integration of mapping into the proposed project.

**4. Describe Your Data:**

* **Source:**
  + Protection monitoring surveys from humanitarian organizations operating in Afghanistan (e.g., UNHCR, DRC).
  + UNHCR protection cluster service mapping data for mapping service with organizations and locations.
  + Reports on incidents and vulnerabilities collected from NGOs.
* **Format:**
  + CSV files for survey data.
* **Size:** Moderate; likely containing thousands of records depending on the availability of surveys data.
* **Preprocessing Steps:**
  + Cleaning and normalizing survey data.
  + Handling missing values.
  + Encoding categorical features (e.g., incident types, service categories).

**5. Approach (Machine Learning or Deep Learning):**

* **Chosen Approach:** Machine Learning
* **Justification:**
  + **Nature of the Data:** The data consists of tabular formats (survey responses and geospatial information), making ML models like Random Forest, Gradient Boosting, or Logistic Regression suitable for classification tasks.
  + **Task Complexity:** The goal is to classify regions by risk level and integrate the output into a service mapping app, which requires interpretable and efficient algorithms rather than complex models like deep neural networks.
  + **Scalability:** ML models are computationally efficient and easier to deploy for such tasks in low-resource settings.