System Architecture

A. Overview

The proposed system is an AI-driven decision-support framework designed for the classification and treatment recommendation of Household Hazardous Waste (HHW). The architecture integrates synthetic data generation, machine learning–based decision modeling, and explainable artificial intelligence (XAI) techniques to ensure transparency and trustworthiness. The system operates as a command-line application, enabling interactive scenario-based analysis while maintaining a modular and extensible design.

B. Architectural Layers

The system follows a layered architecture to separate concerns and improve interpretability, scalability, and maintainability. Each layer performs a distinct role while contributing to the overall decision-making pipeline.

1) User Interaction Layer

This layer provides the interface between the user and the system. A menu-driven command-line interface allows users to select predefined hazardous waste scenarios or enter custom waste parameters. The interface ensures controlled input acquisition and supports exploratory analysis through repeated scenario evaluation.

Functions: - Scenario selection and custom input handling - Input validation and normalization - Triggering prediction and explanation workflows

---

2) Data Simulation Layer

Due to the limited availability and safety constraints of real-world HHW datasets, a synthetic data generation module is employed. This layer generates domain-aware data using rule-based logic derived from hazardous waste management principles.

Generated Attributes: - Waste Type (categorical) - pH Level (continuous) - Flammability Index (continuous) - Toxicity Score (continuous) - Container Integrity (binary)

The target variable, *Recommended Treatment*, is assigned using conditional safety rules, ensuring that the dataset reflects realistic disposal practices.

---

3) Feature Engineering and Encoding Layer

This layer transforms raw and categorical data into a numerical format suitable for machine learning algorithms. Waste categories are encoded using label encoding, while chemical and safety indicators are retained as normalized numerical values.

Feature Vector Representation:

$$X = [Waste\_Type\_Encoded, pH\_Level, Flammability, Toxicity\_Score, Container\_Intact]$$

This transformation ensures consistency, reproducibility, and compatibility with ensemble learning models.

---

4) Machine Learning Layer

The core decision engine is implemented using a Random Forest Classifier. This ensemble-based model is selected for its robustness, ability to model non-linear decision boundaries, and suitability for multiclass classification tasks.

Responsibilities: - Learning complex relationships between waste properties and treatment strategies - Generating treatment recommendations - Producing class probability distributions to quantify prediction confidence

The trained model serves as the primary inference component of the system.

---

5) Local Explainability Layer

To provide transparency at the individual decision level, SHAP (SHapley Additive exPlanations) is employed for local interpretability. For each user-selected scenario, feature-wise contribution values are computed relative to the predicted treatment class.

Outputs: - Ranked feature importance for a single prediction - Directional influence (positive or negative impact) - Human-readable textual explanations

This layer answers the question: *Why was a specific treatment recommended for a given waste instance?*

---

6) Global Explainability Layer

Global model behavior is analyzed using SHAP summary (beeswarm) plots generated over the test dataset. Unlike local explanations, this layer provides insight into overall feature importance and interaction trends across all waste samples.

Purpose: - Understanding dominant decision drivers - Validating model behavior - Detecting potential bias or over-reliance on specific features

This layer supports system-level transparency and model validation.

---

7) Visualization Layer

The visualization layer translates explainability outputs into interpretable graphical representations. Color-coded SHAP plots illustrate both the magnitude and direction of feature influence, enabling intuitive analysis by technical and non-technical stakeholders.

C. End-to-End Workflow

1. Synthetic HHW data generation
2. Feature encoding and dataset preparation
3. Model training and validation
4. SHAP explainer initialization
5. Interactive user input and prediction
6. Local explanation generation
7. Global feature importance visualization

D. Architectural Significance

The proposed architecture uniquely combines interactive decision support with explainable machine learning for hazardous waste management. By integrating both local and global explainability mechanisms, the system moves beyond black-box prediction and enables informed, transparent, and trustworthy decision-making.

E. Summary

This layered architecture ensures modularity, interpretability, and extensibility, making it suitable for academic research and practical decision-support applications. The design aligns with IEEE and Springer expectations by emphasizing clarity, reproducibility, and explainable intelligence in safety-critical domains.