**Explainable AI for Breast Cancer Prediction**

BY: Yamani Swathi(MST030118)

Mentor:-Urooj Khan

Submitted to Meta Scifor Technologies

**I. Introduction**

**Problem Statement:** Many machine learning models, especially complex ones like Random Forests or Neural Networks, act as "black boxes," offering predictions without explanations. In sensitive areas such as healthcare, understanding how and why a prediction is made is critical. Explainable AI (XAI) provides transparency, enabling users to understand and trust model decisions.

**Objective:** To implement Explainable AI techniques on a real-world classification dataset (Breast Cancer) using SHAP values and visualize how each feature influences predictions.

**II. Methodology**

**Dataset Used:** Breast Cancer Wisconsin (Diagnostic) Dataset from sklearn.datasets. It contains 30 numerical features computed from digitized images of breast masses.

**Approach:**

1. Load and preprocess the dataset.
2. Train a machine learning model (Random Forest Classifier).
3. Apply SHAP to interpret model predictions.
4. Visualize feature contributions using SHAP plots.

**Technology Stack:**

* Python
* Libraries: pandas, scikit-learn, shap, matplotlib, seaborn
* Jupyter Notebook (for implementation)

**III. System Design and Architecture**

**Modules:**

* **Data Loading Module:** Loads the dataset and handles missing values.
* **Preprocessing Module:** Scales/encodes the data.
* **Modeling Module:** Trains the Random Forest classifier.
* **XAI Module:** Uses SHAP to explain predictions.
* **Visualization Module:** Generates summary and force plots.

**Architecture Flow:**

* Data Input → Preprocessing → Model Training → SHAP Explainer → SHAP Visuals

**IV. Use Cases**

* Assisting doctors in understanding model-based diagnoses.
* Identifying which cell features influence malignancy predictions.
* Increasing transparency in medical AI applications.

**V. Challenges & Limitations**

* SHAP computation can be slow for large datasets.
* Model interpretability vs accuracy trade-off.
* Not all explanation techniques are model-agnostic.

**VI. Future Enhancements**

* Incorporate LIME and compare with SHAP.
* Apply on larger and more diverse datasets.
* Build a Streamlit dashboard to allow interactive exploration.

**References**

* SHAP Documentation: https://shap.readthedocs.io
* Scikit-learn: https://scikit-learn.org
* Breast Cancer Dataset: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)

**Acknowledgment**

I would like to express my gratitude to mentor Urooj Khan for their guidance, and to Meta Scifor Technologies ,Bangalore for the opportunity to work on this project.