**Project Design Phase**

**Problem – Solution Fit Template**

|  |  |
| --- | --- |
| Date | 15 February 2026 |
| Team ID | LTVIP2026TMIDS79711 |
| Project Name | **Rising Waters: A Machine Learning Approach to Flood Prediction** |
| Maximum Marks | 2 Marks |

# Problem Solution – Selected Algorithm

## Project Title:

**Rising Waters: A Machine Learning Approach to Flood Prediction**

## 📌 Selected Algorithm: XGBoost Classifier

After training and evaluating multiple machine learning models (Decision Tree, Random Forest, KNN, and XGBoost), **XGBoost (Extreme Gradient Boosting)** was selected as the final algorithm for flood prediction.

## 🎯 Why XGBoost Was Chosen

### 1️⃣ High Accuracy

During model comparison, XGBoost achieved **96.55% accuracy**, which was equal to or better than other models. It provided consistent and reliable performance on the test dataset.

### 2️⃣ Better Generalization

Unlike a single Decision Tree, XGBoost uses a **boosting technique**, where multiple weak learners are combined to create a strong predictive model.

This reduces:

* Overfitting
* Variance
* Prediction errors

### 3️⃣ Handles Complex Patterns

Flood prediction depends on:

* Seasonal rainfall distribution
* Cloud cover patterns
* Annual rainfall trends

XGBoost can effectively capture **non-linear relationships** between these features and the target variable.

### 4️⃣ Robust Performance on Imbalanced Data

Flood datasets may have fewer severe flood cases compared to non-flood cases.

XGBoost handles class imbalance effectively and provides better classification stability.

### 5️⃣ Feature Importance Capability

XGBoost allows analysis of feature importance, helping understand which rainfall periods contribute most to flood occurrence.

### 6️⃣ Faster and Efficient

* Optimized gradient boosting implementation
* Parallel processing support
* Efficient memory usage

This makes it suitable for real-time prediction in a web application.

## 📊 Performance Metrics of Final Model

* Accuracy → 96.55%
* Precision → 1.0
* Recall → 0.75

Interpretation:

* High precision indicates no false flood alarms.
* Good recall shows most flood cases were correctly detected.