

The document "Report for Flood Prediction Model" provides detailed steps for building a predictive model for flood occurrence using weather data. Here are the libraries used in the build process, categorized by the steps in which they are employed:

1. Data Loading and Preprocessing

- `pandas`: Used for data manipulation and analysis, especially for loading the dataset and handling dates.
- `numpy`: Used for numerical operations.
- `sklearn.preprocessing.LabelEncoder`: Used for encoding categorical features.

2. Feature Scaling

- `sklearn.preprocessing.StandardScaler`: Used for feature scaling to ensure numerical features are on a similar scale.

3. Model Building

- `sklearn.model_selection.train_test_split`: Used for splitting the dataset into training and testing sets.
- `sklearn.ensemble.RandomForestRegressor`: Used for building the Random Forest model.
- `sklearn.ensemble.GradientBoostingRegressor`: Used for building the Gradient Boosting model.
- `sklearn.ensemble.VotingRegressor`: Used for combining the ensemble models.
- `sklearn.model_selection.GridSearchCV`: Used for hyperparameter tuning using Grid Search.

4. Model Evaluation

- `sklearn.metrics.mean_absolute_error`: Used for calculating Mean Absolute Error (MAE).
- `sklearn.metrics.mean_squared_error`: Used for calculating Root Mean Squared Error (RMSE).
- `sklearn.metrics.r2_score`: Used for calculating R-squared (R^2).
- `sklearn.metrics.mean_absolute_percentage_error`: Used for calculating Mean Absolute Percentage Error (MAPE).

5. Visualizations

- `matplotlib.pyplot`: Used for creating plots to visualize actual vs. predicted values, residuals, and prediction error distribution.
- `seaborn`: Used for enhanced visualizations.

These libraries form the core tools used throughout the process, facilitating various stages from data preprocessing to model evaluation and visualization.