

TECHNOLOGIES USED:

1. IBM watsonx.ai Studio

A no-code AI and machine learning development environment within IBM Cloud Pak for Data. It allows users to upload datasets, use **AutoAI** to automate preprocessing, model selection, and hyperparameter tuning, and deploy models as APIs without writing code. Ideal for rapid prototyping and production deployment.

2. IBM AutoAI

An automated machine learning (AutoML) engine within watsonx.ai. AutoAI:

- Splits datasets into training and validation subsets
 - Handles missing values, categorical encoding, and scaling
 - Tests multiple algorithms (e.g., Random Forest, Gradient Boosted Trees, Logistic Regression)
 - Selects the best model based on metrics like accuracy and F1-score
 - Automatically generates a deployment-ready pipeline.
-

3. IBM watsonx.ai Runtime

The execution environment where deployed models run in the cloud. It:

- Hosts the model and handles API requests
- Processes input data in real time
- Returns prediction results with confidence scores.

4. IBM Cloud Pak for Data

An integrated platform for data collection, management, governance, and AI development. It provides:

- Secure cloud storage and access control
 - Tools for data preparation and cleaning
 - AI lifecycle management from data ingestion to model monitoring.
-

5. IBM Cloud Object Storage

A secure and scalable cloud-based storage service used to store:

- Training datasets (e.g., nsapallschemes.csv)
 - Model artifacts (trained model files, preprocessing pipelines)
 - Logs and monitoring data for retraining.
-

6. Random Forest Classifier

An ensemble-based supervised machine learning algorithm that:

- Trains multiple decision trees on random subsets of data and features
- Uses majority voting to make predictions
- Handles both categorical and numerical inputs
- Reduces overfitting and improves accuracy in complex datasets like NSAP eligibility.