# CAPSTONE PROJECT

## INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS

Powered by:
Anu bharti

edunet foundation

# OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References

edunet foundation

# PROBLEM STATEMENT

The Pradhan Mantri Gram Sadak Yojana (PMGSY) is a flagship rural development program in India, initiated to provide all-weather road connectivity to eligible unconnected habitations. Over the years, the program has evolved through different phases or schemes (PMGSY-I, PMGSY-II, RCPLWEA, etc.), each with potentially distinct objectives, funding mechanisms, and project specifications. For government bodies, infrastructure planners, and policy analysts, efficiently categorizing thousands of ongoing and completed projects is crucial for effective monitoring, transparent budget allocation, and assessing the long-term impact of these schemes. Manual classification is time-consuming, prone to errors, and scales poorly. Your specific task is to design, build, and evaluate a machine learning model that can automatically classify a road or bridge construction project into its correct PMGSY_SCHEME based on its physical and financial characteristics.

# PROPOSED SOLUTION

**To build a scalable and intelligent system using IBM Cloud that classifies rural infrastructure projects based on multiple parameters like location, type, priority, cost, feasibility, and environmental impact.**

- **Data Collection:**
- **Downloaded from public government portals.**
- **"The dataset, containing real-world rural infrastructure project details, was uploaded securely to IBM Cloud Object Storage and served as the foundation for model training and testing."**
- **Data Preprocessing:**
  - **Thorough data processing ensured the model received clean and structured input, which improved training performance and boosted the model's classification accuracy."**
- **Machine Learning Algorithm:**
- **"The AutoAI pipeline trained several classification algorithms. After comparing their performance, the XGBoost Classifier was selected due to its high accuracy and confidence across multiple PMGSY categories."**
- **Deployment:**
  - **Model was deployed as a REST API endpoint using Watson Machine Learning.**
  - **"Once deployed, the model can automatically classify incoming rural infrastructure project data into PMGSY categories via API or web interface, supporting decision-makers in real time."**
- **Evaluation:**
  - **"Model evaluation shows promising accuracy and confidence for multiclass classification, with the potential to assist policy-makers in project categorization and prioritization."**
  - Result:

edunet foundation

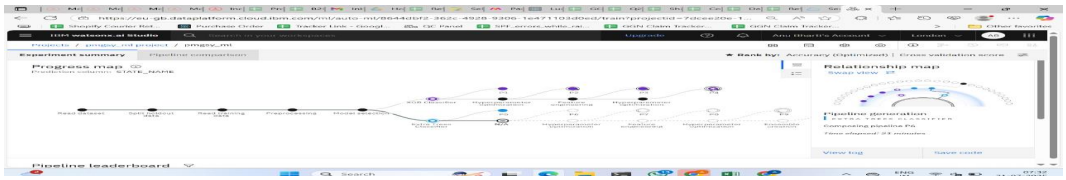
# SYSTEM APPROACH

Components used for making this machine learning project using watsonx studio AI:

- Data Storage : IBM Cloud Object Storage.
- Data Preprocessing : IBM Watson Studio.
- Model Building : IBM Watson Studio AI.
- Model Deployment : IBM Watson Machine Leaning .
- After deploying upload data to test it.
- Test the model by testing it in test tab .

# ALGORITHM & DEPLOYMENT

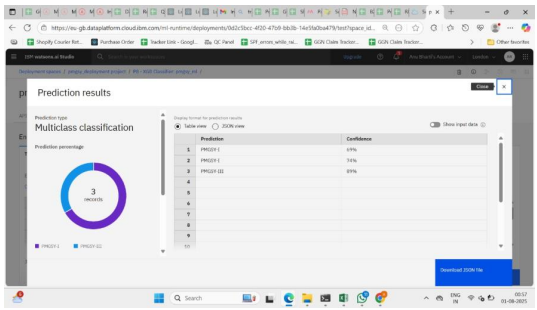
- **Multiclass Classification:**
- The objective was to classify rural infrastructure projects into categories like PMGSY-I, PMGSY-II, and PMGSY-III using supervised learning.
- **Algorithm selection**:
- IBM Watson AutoAI evaluated multiple models including Logistic Regression, Random Forest, and XGBoost. Based on accuracy and confidence, XGBoost Classifier was selected as the best-performing algorithm.
- **Data input:**
- Cleaned and preprocessed project data (CSV format) was uploaded to IBM Cloud Object Storage. Key features included project cost, length, duration, location, and scheme type.
- **Training process:**
- AutoAI automatically split the dataset, performed feature engineering, hyperparameter tuning, and trained multiple models. The selected XGBoost model achieved high accuracy across all classes.
- **Prediction process**:
- The final model was deployed using IBM Watson Machine Learning as a REST API. When new project data is input, the model returns the predicted PMGSY category along with a confidence score.

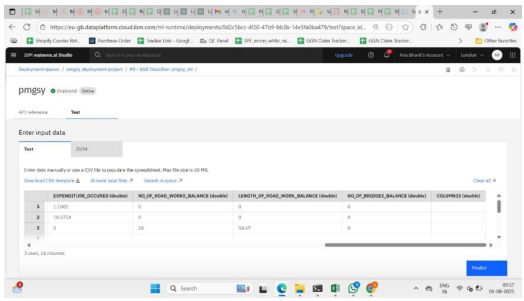


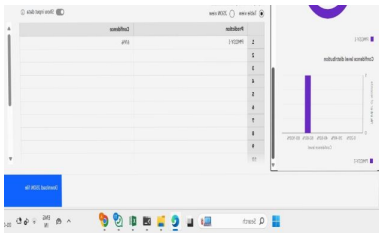

edunet foundation

# RESULT

- Model is trained to classify infrastructure projects into various PMGSY categories.
- “The AI model deployed on IBM Watson Machine Learning successfully classifies PMGSY projects into respective schemes (PMGSY-I, II, III), enabling automated, scalable categorization of rural infrastructure initiatives.”



edunet foundation

# CONCLUSION

This intelligent classification system, built using IBM Cloud and Watson services, can greatly enhance the way rural infrastructure projects are analyzed and prioritized, leading to improved governance, transparency, and developmental outcomes.

# FUTURE SCOPE

- Integrate geospatial data using IBM GeoSpatial Analytics.
- Expand classification to include risk assessment.
- Multilingual data support using IBM Watson NLP APIs.

edunet foundation

# REFERENCES

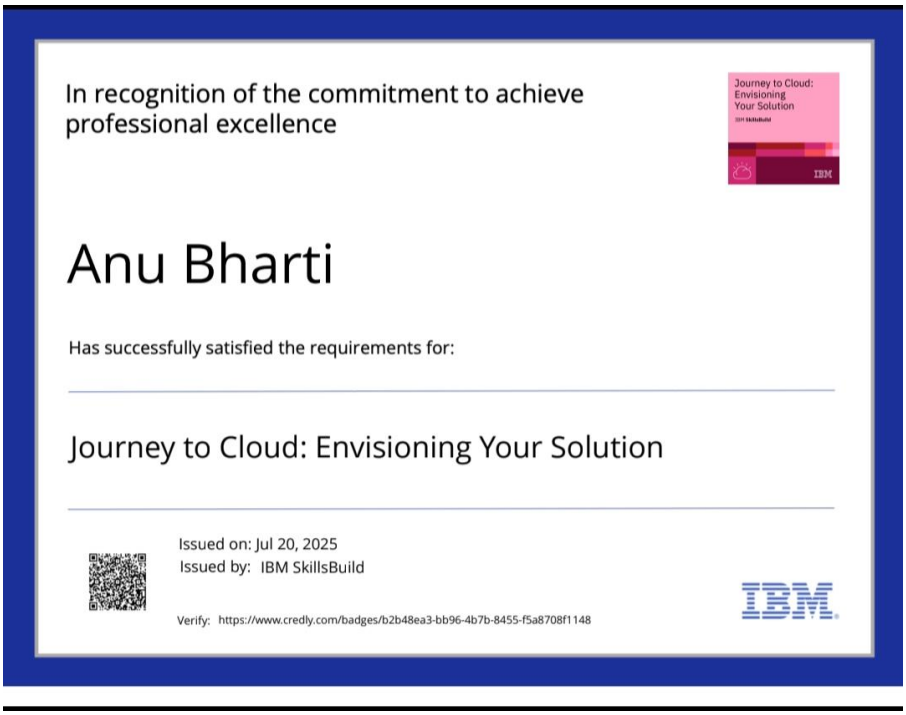
Data was collected from the official PMGSY portal and Data.gov.in. IBM Watson Studio, AutoAI, and Watson Machine Learning were used for model development and deployment. Python libraries like Pandas, NumPy, and Scikit-learn supported data processing. IBM Cloud documentation and ML resources guided the implementation.

edunet foundation

# IBM CERTIFICATIONS


edunet foundation

# IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM

Anu Bharti

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/b2b48ea3-bb96-4b7b-8455-f5a8708f1148

IBM

edunet foundation

# IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Anu Bharti

for the completion of

**Lab: Retrieval Augmented Generation with LangChain**

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

**Completion date:** 24 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

# THANK YOU

edunet
foundation