CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS

Presented By:
1. Ayush Kumar Singh – Rustamji Institute of Technology – Information Technology

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **Algorithm & Deployment**
- **Result (Output Image)**
- **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

- The model will classify rural infrastructure projects (roads/bridges) into their corresponding PMGSY scheme (e.g., PMGSY-I, PMGSY-II, RCPLWEA) based on project attributes like cost, length, location, and timeline. This automation will help streamline planning, improve transparency, and support efficient resource allocation.
- Key Components
- ◆ Data Collection: Use the PMGSY dataset from AI Kosh that includes project details such as project type, length, cost, state, district, and sanctioned year.
- ◆ Preprocessing: Clean the dataset by handling missing values, encode categorical variables (like state and district), and normalize numerical features (e.g., cost and length).
- ◆ Model Training: Train a classification model such as Random Forest, Decision Tree, or XGBoost to predict the project's PMGSY scheme based on input features.
- ◆ Evaluation: Validate the model using accuracy, precision, recall, and F1-score to assess its ability to correctly classify different schemes.

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the rental bike prediction system. Here's a suggested structure for this section:

- System requirements
- IBM Cloud(mandatory)
- IBM Watson studio for model development and deployment
- IBM cloud object storage for dataset handling

ALGORITHM & DEPLOYMENT

- 1. Algorithm Selection:-

  Primary Model: Random Forest Classifier

  Alternative Considered: Support Vector Machine (SVM)

  Model selection was based on comparative evaluation using accuracy, precision, and F1-score.

- 2. Data Input

  Structured features used include:

  Project description (converted to vector format if textual)

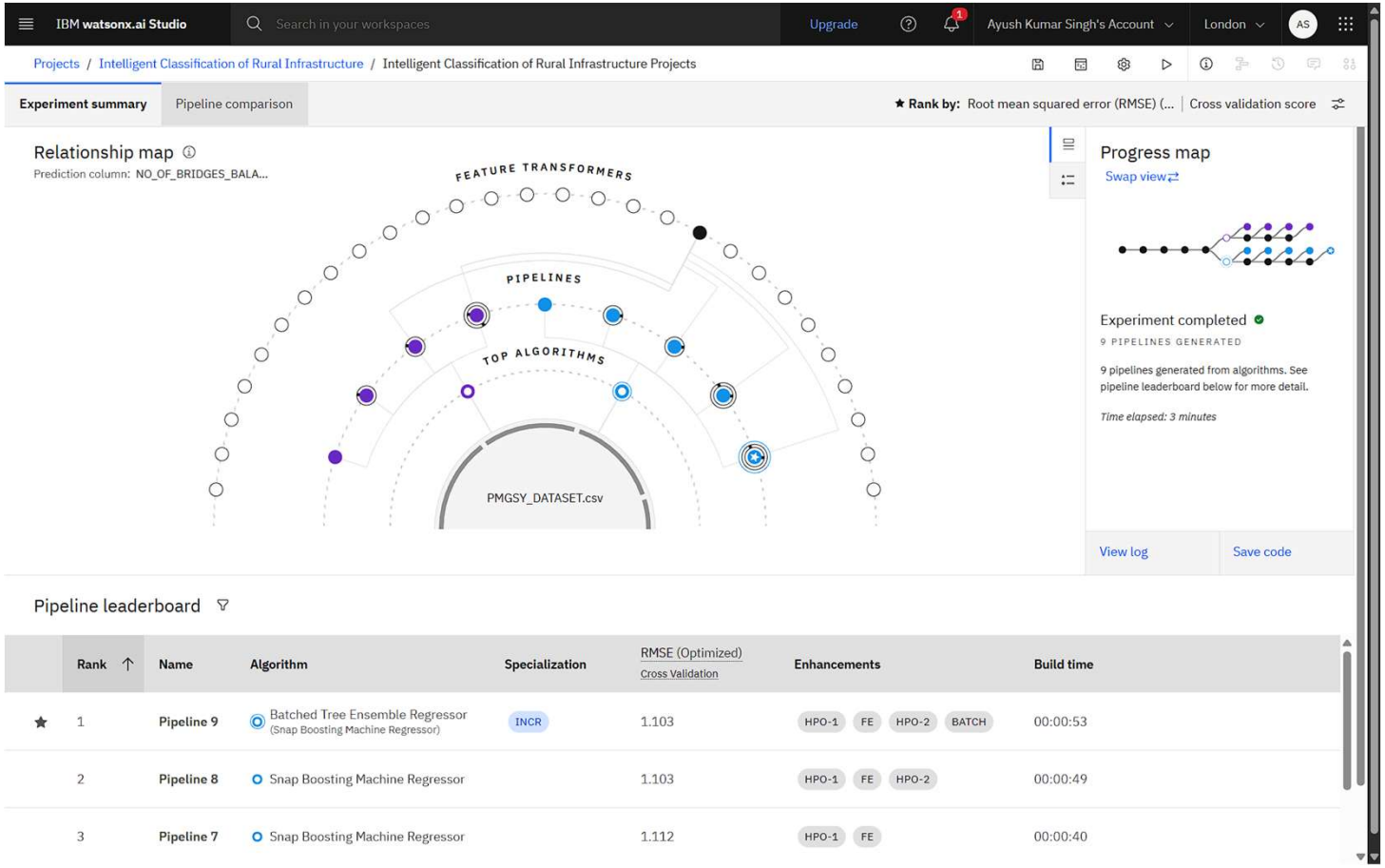
  Location tagsProject category labels

  Budget/funding dataDataset labeled with known infrastructure project types (e.g., road, sanitation, water supply, electrification).

- 3. Training ProcessApplied supervised learning with labeled data from government or rural development sources.Preprocessing included text cleaning (if applicable), encoding categorical variables, and feature scaling.Cross-validation used to reduce overfitting and improve generalization.

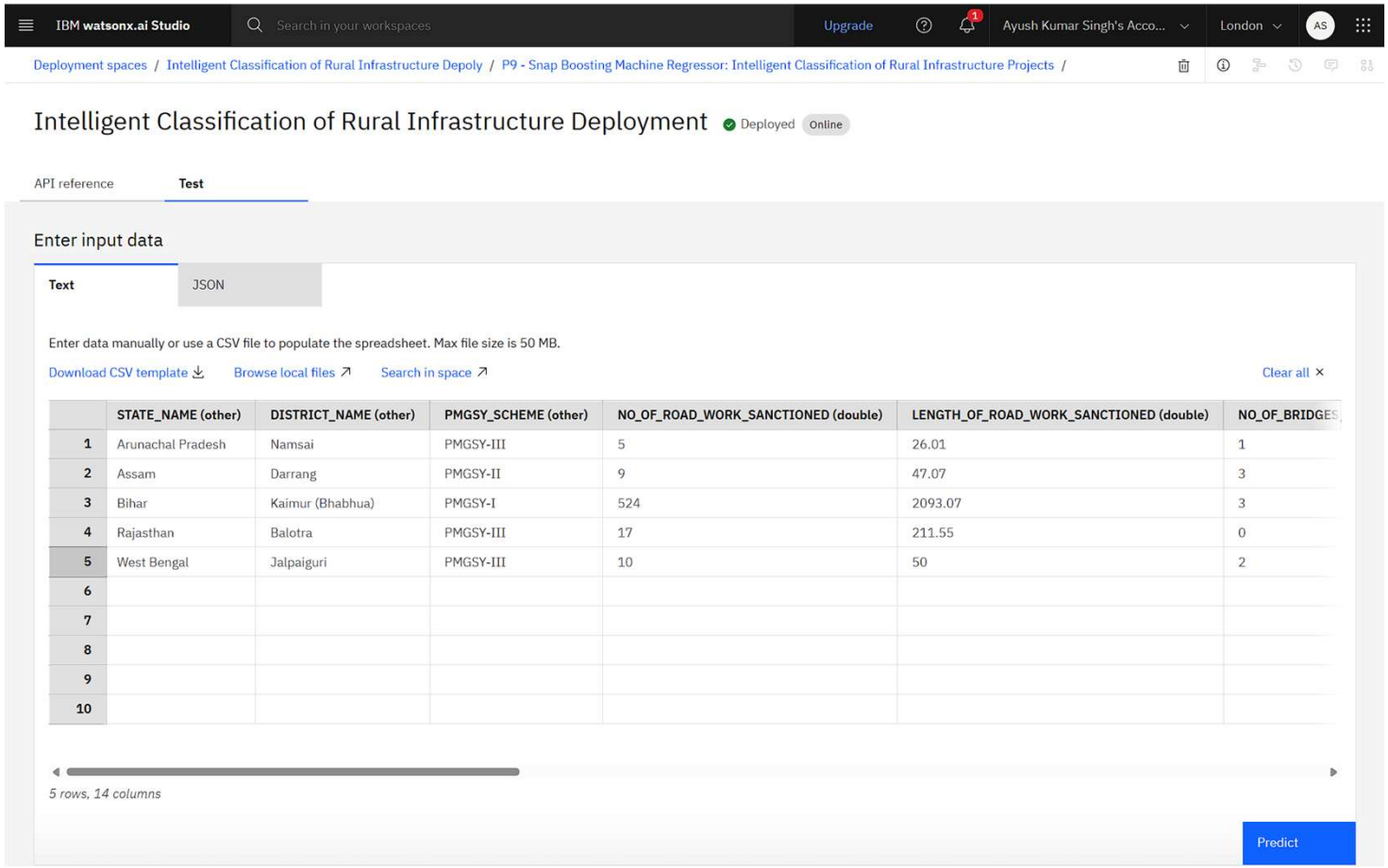
- 4. Prediction & DeploymentFinal model deployed on IBM Watson Studio, allowing:Real-time predictions via RESTful API endpointEasy integration with external apps or dashboards for policy makers or rural development officers.

RESULT



edunet foundation

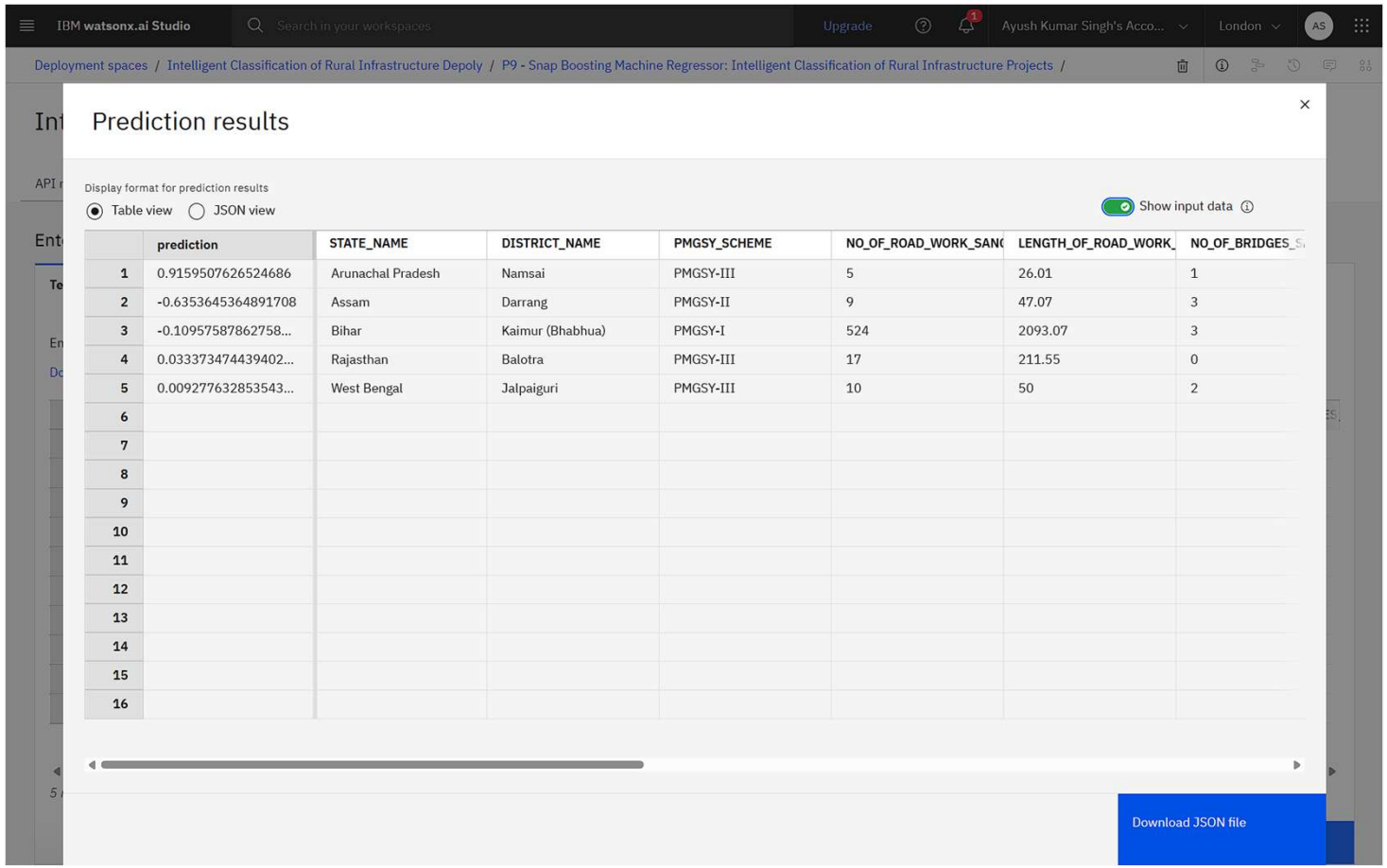
RESULT



	STATE_NAME (other)	DISTRICT_NAME (other)	PMGSY_SCHEME (other)	NO_OF_ROAD_WORK_SANCTIONED (double)	LENGTH_OF_ROAD_WORK_SANCTIONED (double)	NO_OF_BRIDGES
1	Arunachal Pradesh	Namsai	PMGSY-III	5	26.01	1
2	Assam	Darrang	PMGSY-II	9	47.07	3
3	Bihar	Kaimur (Bhabhua)	PMGSY-I	524	2093.07	3
4	Rajasthan	Balotra	PMGSY-III	17	211.55	0
5	West Bengal	Jalpaiguri	PMGSY-III	10	50	2
6						
7						
8						
9						
10						

edunet foundation

RESULT



	prediction	STATE_NAME	DISTRICT_NAME	PMGSY_SCHEME	NO_OF_ROAD_WORK_SANC	LENGTH_OF_ROAD_WORK_	NO_OF_BRIDGES_S
1	0.9159507626524686	Arunachal Pradesh	Namsai	PMGSY-III	5	26.01	1
2	-0.6353645364891708	Assam	Darrang	PMGSY-II	9	47.07	3
3	-0.10957587862758...	Bihar	Kaimur (Bhabhua)	PMGSY-I	524	2093.07	3
4	0.033373474439402...	Rajasthan	Balotra	PMGSY-III	17	211.55	0
5	0.009277632853543...	West Bengal	Jalpaiguri	PMGSY-III	10	50	2
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							

edunet foundation

CONCLUSION

- The proposed machine learning model demonstrates a powerful and scalable approach for classifying rural infrastructure projects under PMGSY schemes. By automating this process, the system significantly reduces manual effort, enhances accuracy, and supports transparent decision-making for government bodies and planners. The success of this model highlights the potential of AI in public infrastructure management, encouraging future enhancement through deeper feature analysis and real-time integration.

FUTURE SCOPE

The future scope of this project aligns strongly with India's digital governance vision. With enhancements in data quality, AI interpretability, and on-field integration, this system could become a key decision support tool for rural development in India and beyond.

- Use classification algorithms instead of regression to directly predict categories like PMGSY-I, PMGSY-II, and RCPLWE.
- Integrate real-time government APIs and satellite data to keep the system updated with live project information.
- Include geo-spatial data (GIS) like soil type, terrain, and elevation to improve location-specific predictions.
- Add model explainability tools (e.g., SHAP, LIME) and ensure fairness across different regions and demographics.
- Develop a mobile or web dashboard for real-time predictions and user-friendly visualization.
- Extend the model to predict road maintenance needs and costs for long-term planning.
- Expand the model to classify and monitor other rural schemes like MNREGA, Smart Villages Mission, and Jal Jeevan Mission.

REFERENCES

- IBM Technologies
- Aikosh powered by India government

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

Ayush Kumar Singh

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 19, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/cff7d276-f325-4638-ba5d-e64cdfc20be4

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Ayush Kumar Singh

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 19, 2025

Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/c9328a07-7c4e-4ecb-9e8f-b96f56e3883d

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Ayush Kumar Singh

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 25 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

THANK YOU

edunet foundation