CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS

Presented By:
1. Rohit Shitalkumar Mudakude-PES University Bengaluru-CSE

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**

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 automate and streamline the classification of rural infrastructure projects, we propose a **machine learning-based classification system** that can :
  - Predict the appropriate PMGSY scheme (PMGSY-I, II, RCPLWEA, etc.).
  - Use project-level physical and financial features for training.
  - Deliver fast, scalable, and accurate predictions.
  - Be integrated into analytics platforms for policy and budget planning.
- **Data Collection:** Uses the AI Kosh dataset on NSAP scheme.
- **Data Preprocessing:** Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.
- **Machine Learning Algorithm:** Train classification model (e.g decision tree, random forest or SVM).
- **Evaluation:** Validate the model using accuracy, precision, recall, and F1 score.

SYSTEM APPROACH

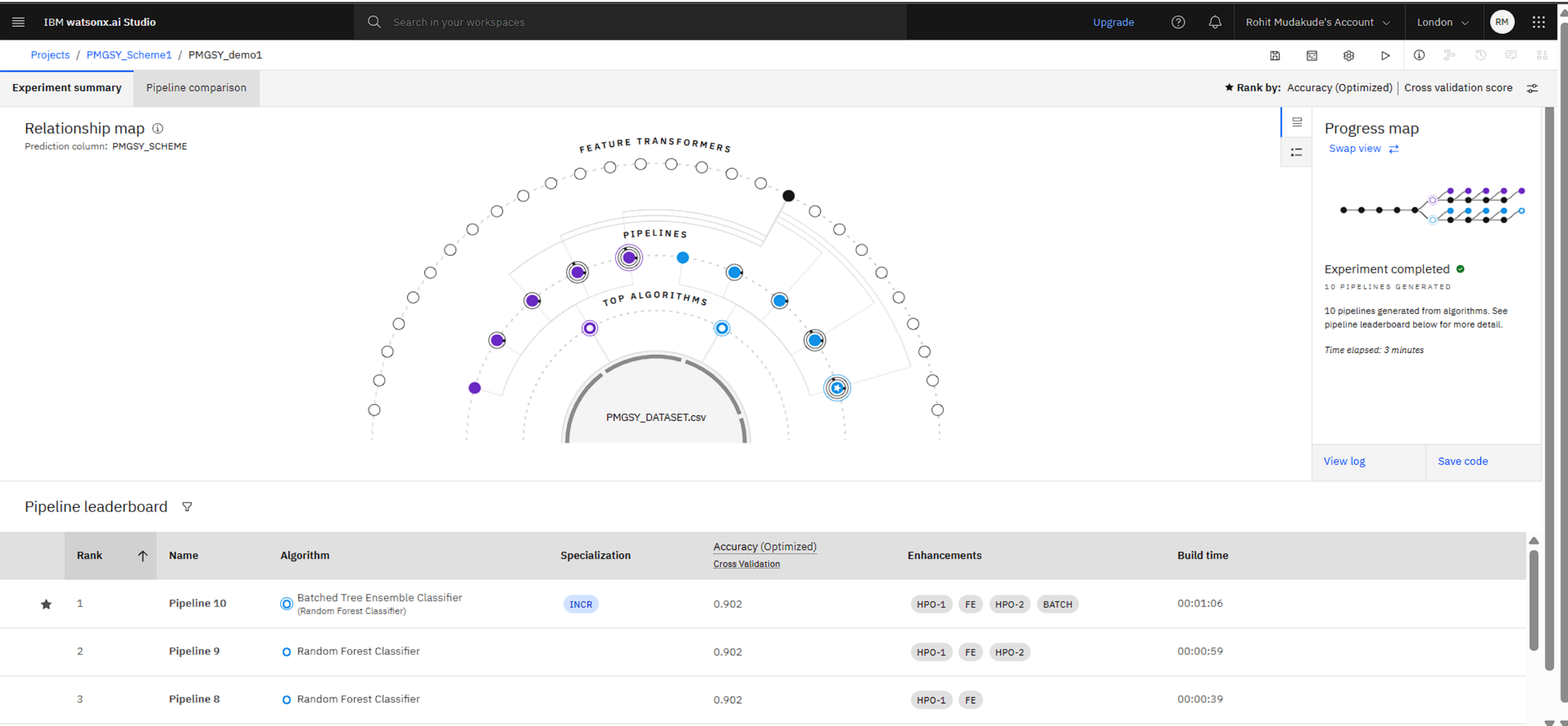
The "System Approach" section outlines the overall strategy and methodology for developing and implementing the intelligent classification of rural infrastructure projects. Here's a suggested structure for this section:

- System requirements:
  - **IBM Cloud Lite**: For cloud-based model development and deployment
  - **Watson Studio**: For creating notebooks, training ML models, and visualizing results
  - **AI Kosh Dataset**: Used as the source dataset for model training
  - **Python (Scikit-learn, Pandas, Matplotlib)**: For data preprocessing and model implementation
  - **IBM Cloud Object Storage**: To store datasets and model artifacts

ALGORITHM & DEPLOYMENT

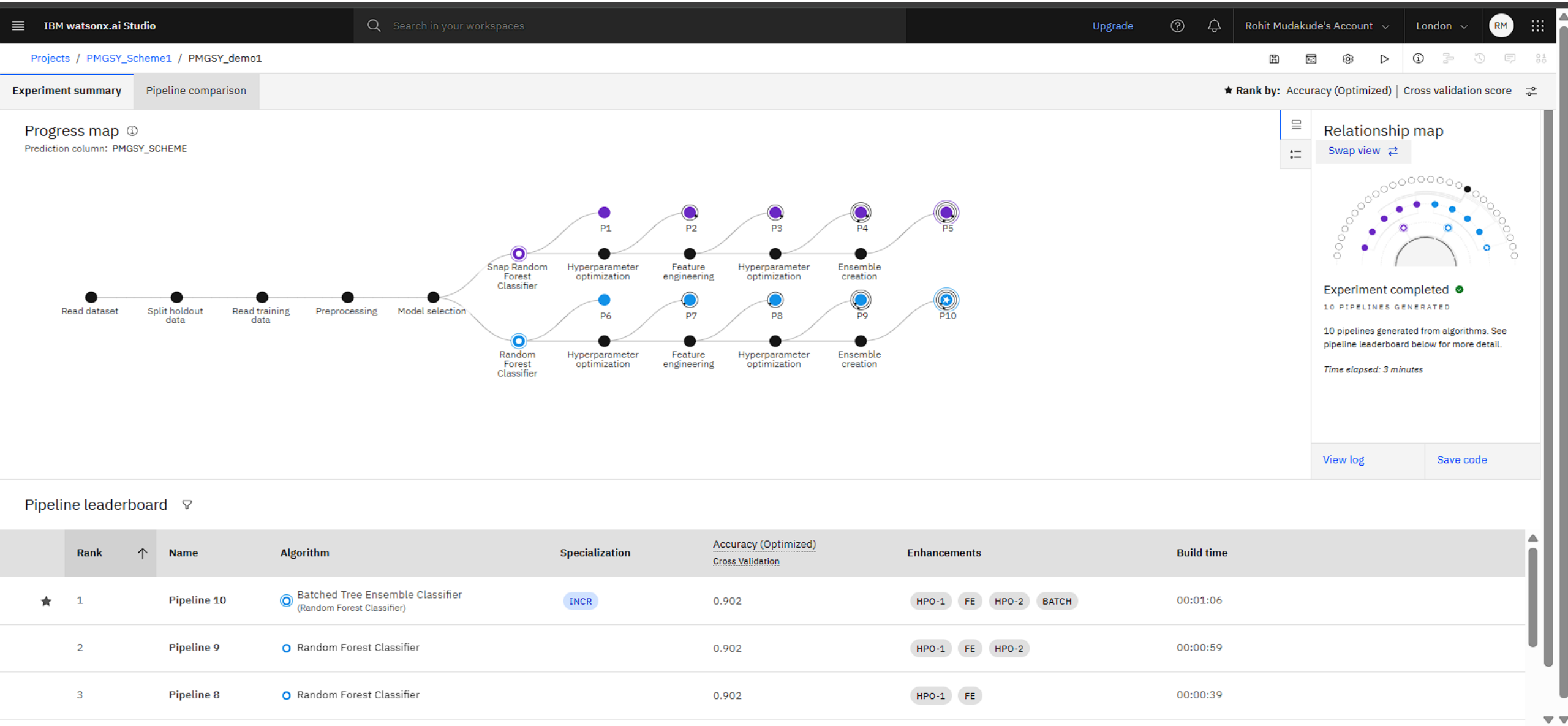
- In the Algorithm section, describe the machine learning algorithm chosen for predicting PMGSY Scheme. Here's an example structure for this section:
- Algorithm Selection:
  - Used Auto AI on IBM Watson to choose the best multi-class classifier (e.g., Random Forest), suited for predicting NSAP schemes from structured data.
- Data Input:
  - Inputs include state name, district, no. of road work sanctioned, length of road work sanctioned, no. of bridges sanctioned, and other socio-economic factors from the AI Kosh dataset.
- Training Process:
  - Performed data cleaning, encoding, and an 80-20 train-test split. Auto AI handled hyperparameter tuning and model evaluation using cross-validation.
- Prediction Process:
  - Deployed model on IBM Watson Machine Learning as an API. Takes applicant data and predicts the eligible NSAP scheme in real time.

RESULT



	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 10	Batched Tree Ensemble Classifier (Random Forest Classifier)	INCR	0.902	HPO-1 FE HPO-2 BATCH	00:01:06
	2	Pipeline 9	Random Forest Classifier		0.902	HPO-1 FE HPO-2	00:00:59
	3	Pipeline 8	Random Forest Classifier		0.902	HPO-1 FE	00:00:39

RESULT



Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 10	Batched Tree Ensemble Classifier (Random Forest Classifier)	INCR	0.902	HPO-1 FE HPO-2 BATCH	00:01:06
	2	Pipeline 9	Random Forest Classifier		0.902	HPO-1 FE HPO-2	00:00:59
	3	Pipeline 8	Random Forest Classifier		0.902	HPO-1 FE	00:00:39

RESULT

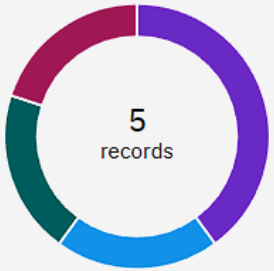
Prediction results

Close

Prediction type

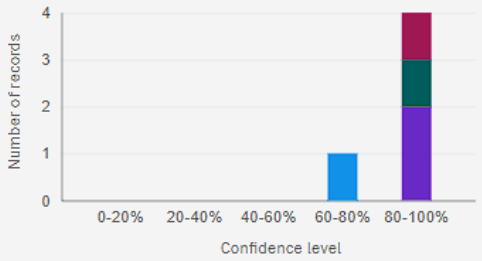
Multiclass classification

Prediction percentage



PMGSY-I RCPLWEA PMGSY-III PMGSY-II

Confidence level distribution



Display format for prediction results

(•) Table view ( ) JSON view

Show input data

	Prediction	Confidence
1	PMGSY-I	90%
2	RCPLWEA	60%
3	PMGSY-I	100%
4	PMGSY-III	95%
5	PMGSY-II	91%
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

Download JSON file

CONCLUSION

- The proposed machine learning model successfully automates the classification of PMGSY infrastructure projects into their respective schemes based on available financial and physical data. By eliminating the manual process, the system improves **speed, consistency, and transparency** in government infrastructure monitoring.
- It demonstrates how **AI/ML with cloud-based tools** like IBM Cloud can assist in large-scale public sector decision-making.

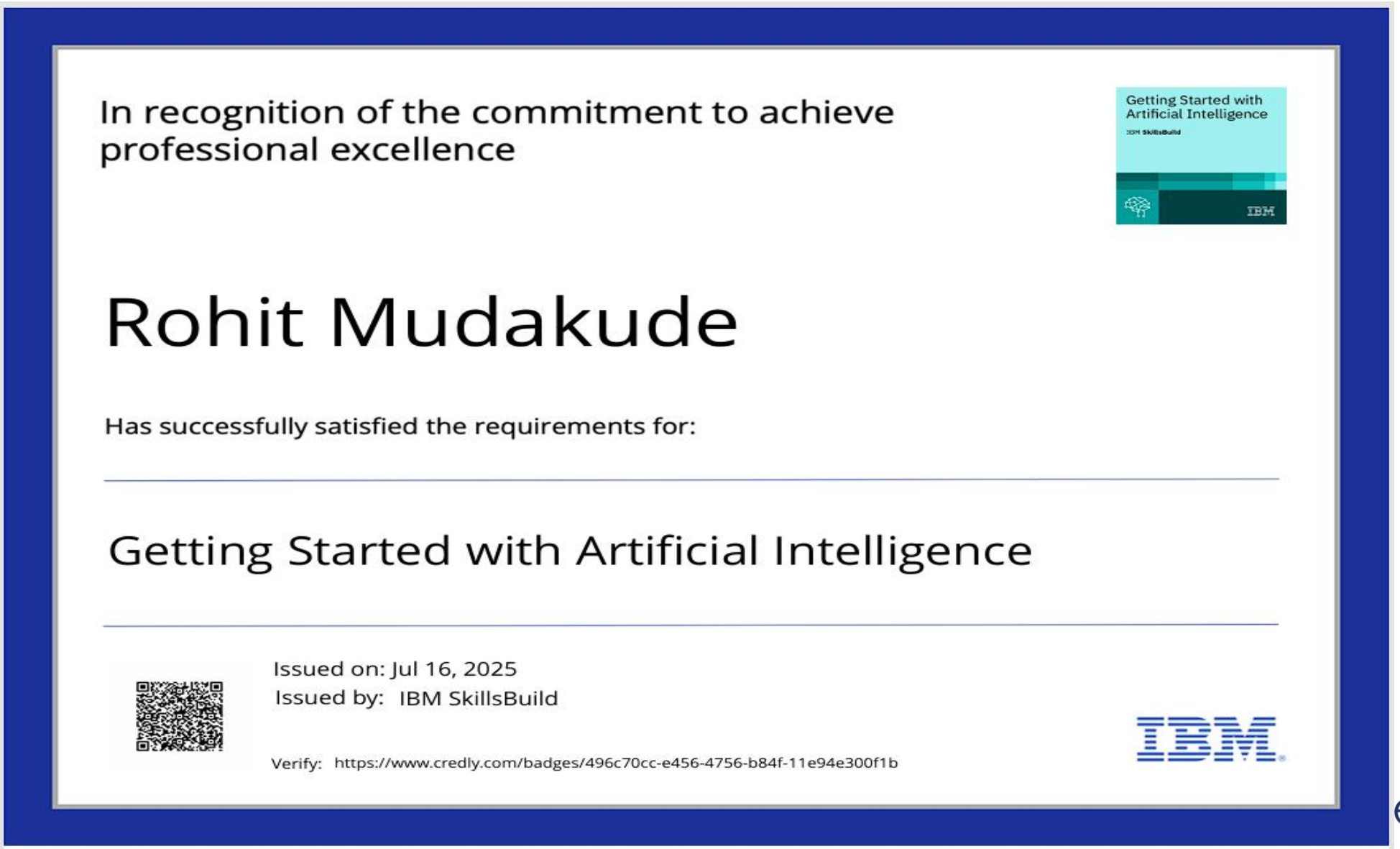
FUTURE SCOPE

- Integrate the model into a **centralized dashboard** for real-time classification
- Extend to other schemes beyond PMGSY, improving generalizability
- Enhance model accuracy using **deep learning** or **ensemble approaches**
- Enable feedback loop from government users for active learning
- Build **API endpoints** for third-party applications to access predictions

REFERENCES

- AI Kosh Dataset –https://aikosh.indiaai.gov.in/account/login
- IBM Cloud – https://cloud.ibm.com/
- IBM Watson Studio – https://cloud.ibm.com/catalog/services/watsonxai-studio
- GitHub Repository Link - https://github.com/rohit3110-pro/-Intelligent-Classification-of-Rural-Infrastructure-Projects/upload

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

IBM

Rohit Mudakude

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 16, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/496c70cc-e456-4756-b84f-11e94e300f1b

IBM

edunet foundation

IBM CERTIFICATIONS


In recognition of the commitment to achieve professional excellence

Rohit Mudakude

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/3cd60ad1-f329-479b-b1ee-9f0182eb4bcf

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** | Completion Certificate

This certificate is presented to

Rohit Mudakude

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