

---

# **CAPSTONE PROJECT**

## **INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS USING MACHINE LEARNING**

**Presented By:**  
**VISHESHA NERALI - S.G BALEKUNDRI INSTITUTE OF TECHNOLOGY- CE**

# 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

Develop a machine learning model that classifies The Pradhan Mantri Gram Sadak Yojana (PMGSY) through different phases or schemes (PMGSY-I, PMGSY-II, RCPLWEA, etc.) using data set provided, the model will process infrastructure planners, and policy analysts, efficiently categorizing thousands of ongoing and completed projects it will classify road or bridge construction project in less time and less error.

## ■ Key components

- Data Collection: AI Kosh dataset .
- Pre-processing: Clean and normalize the data set.
- Machine learning algorithm: Train a classification model using P8-XGB Classifier.
- Deployment: Deploy the solution on a scalable and reliable platform IBM Cloud.
- Evaluation: Validate the model using accuracy and precision.
- Result: Assist government agencies to categorize PMGSY quickly and accurately.

---

# SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing Intelligent Classification of Rural Infrastructure Projects (PMGSY) .Here's a suggested structure for this section:

- **System requirements**

- IBM Cloud
- watsonx.ai studio for model development and deployment.
- watsonx.ai Runtime
- Cloud storage object for dataset handling.

# ALGORITHM & DEPLOYMENT

In the Algorithm section, describes the machine learning algorithm chosen for predicting PMGSY. Here's an example structure for this section:

- **Algorithm Selection:**

P8-XGB Classifier

- **Data Input:**

State Name, District Name, PMGSY Scheme, No of road worked sanctioned, Length of road worked sanctioned, No of bridges sanctioned, Cost of work sanctioned, No Of road works completed, Length of road work completed, No of bridges completed, Expenditure occurred, No of road works balance, Length of road work balance, No of bridge balance.

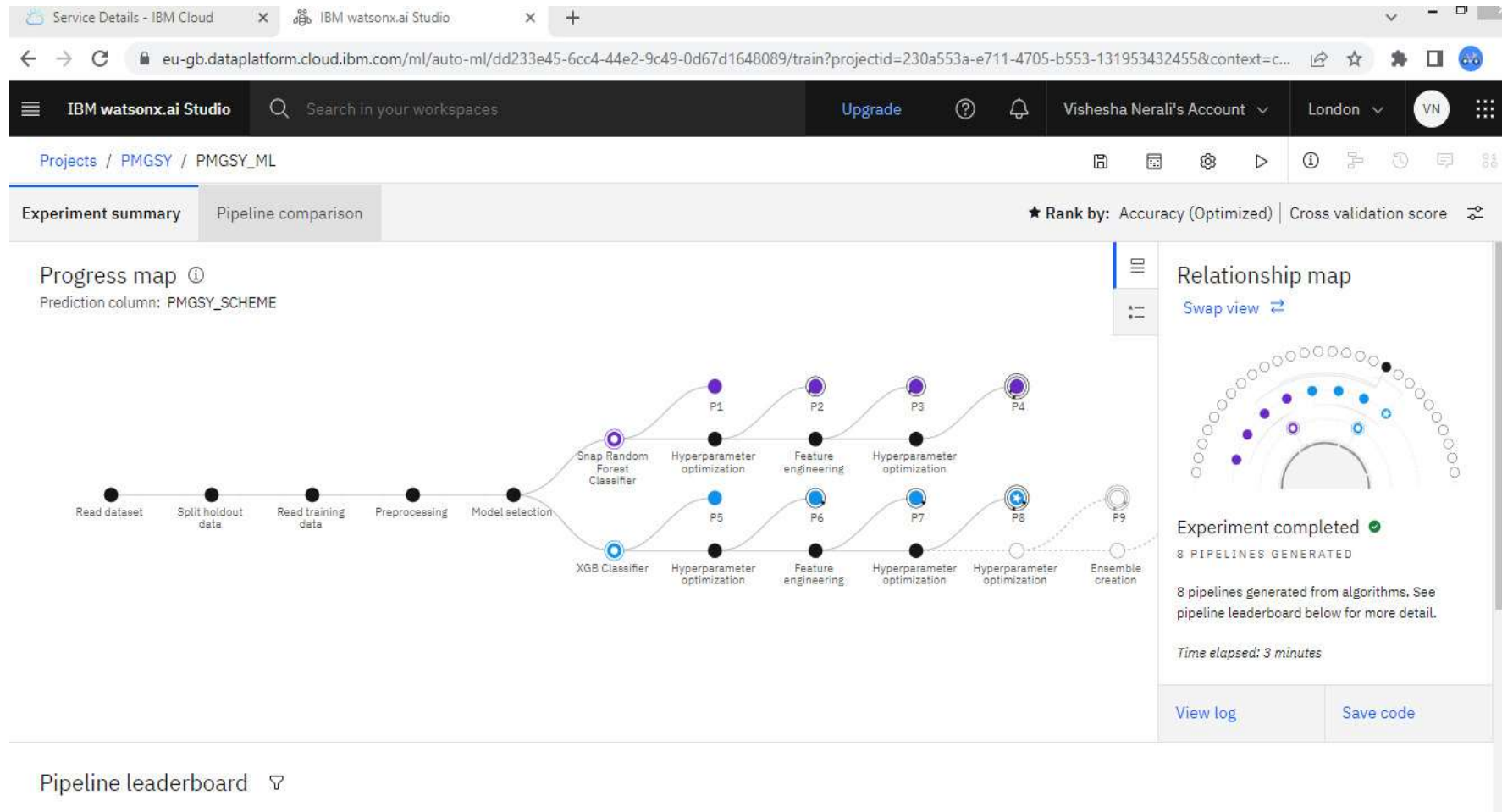
- **Training Process:**

Supervised learning using labelled PMGSY types.

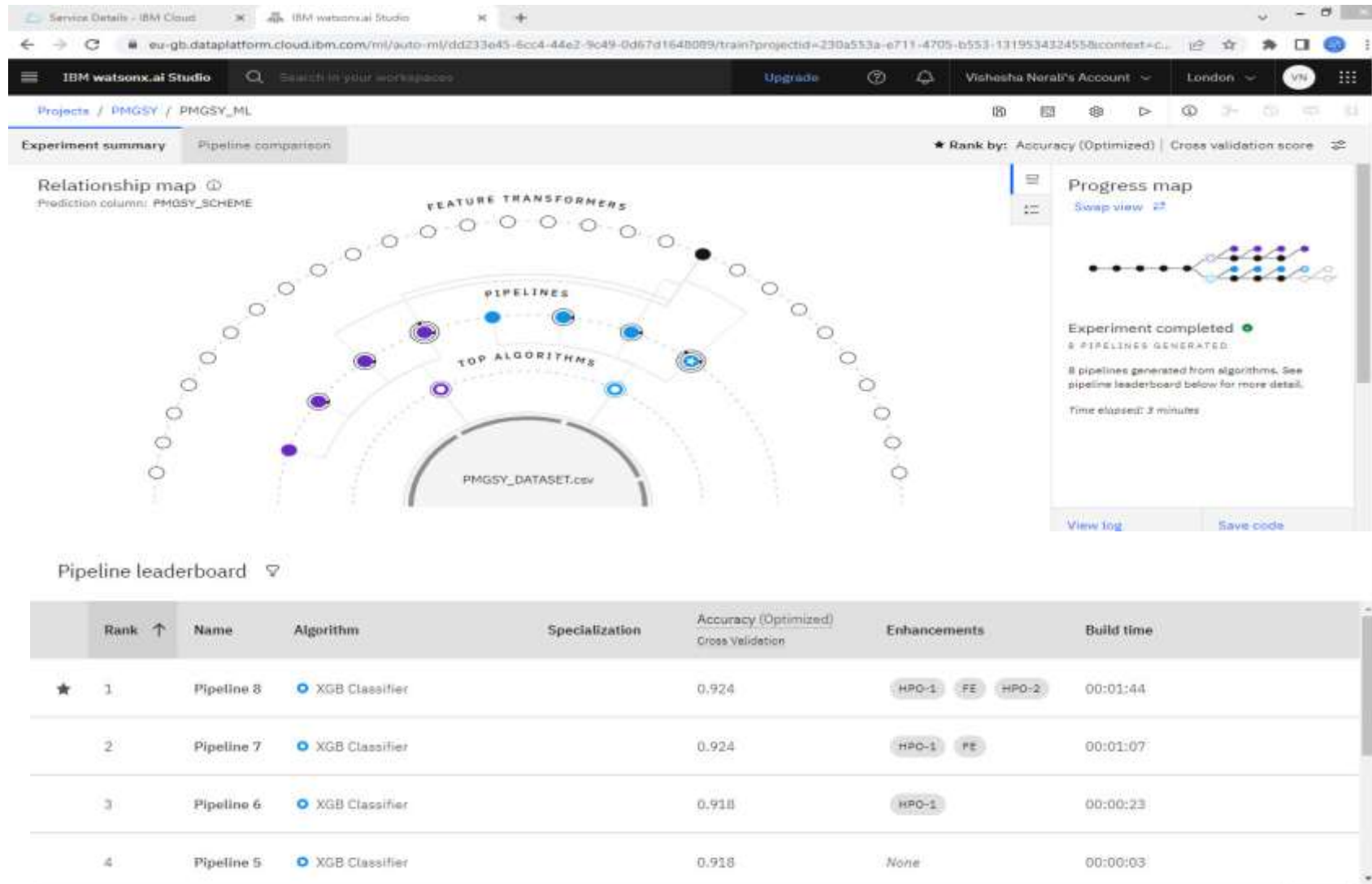
- **Prediction Process:**

Model deployed on IBM watsonax.ai Studio with API endpoint real-time predictions.

# PROGRESS MAP



# RELATIONSHIP MAP





# INPUT DATA

Service Details - IBM Cloud x PMGSY\_DEPLOY2 — PMGSY\_DEP x +

eu-gb.dataplatform.cloud.ibm.com/ml-runtime/deployments/73e02a0a-6b89-4a53-8ec2-a5d50a6e3c8c/test?space\_id=fe0e9cac-e817-4cdf-bf84-3ba54de74c...

IBM watsonx.ai Studio Search in your workspaces Upgrade ? 1 Vishesha Nerall's Account London VN

Deployment spaces / PMGSY\_DEPLOY / P8 - XGB Classifier: PMGSY\_ML /

PMGSY\_DEPLOY2 Deployed Online

API reference **Test**

Enter input data

Text JSON

Enter data manually or use a CSV file to populate the spreadsheet. Max file size is 50 MB.

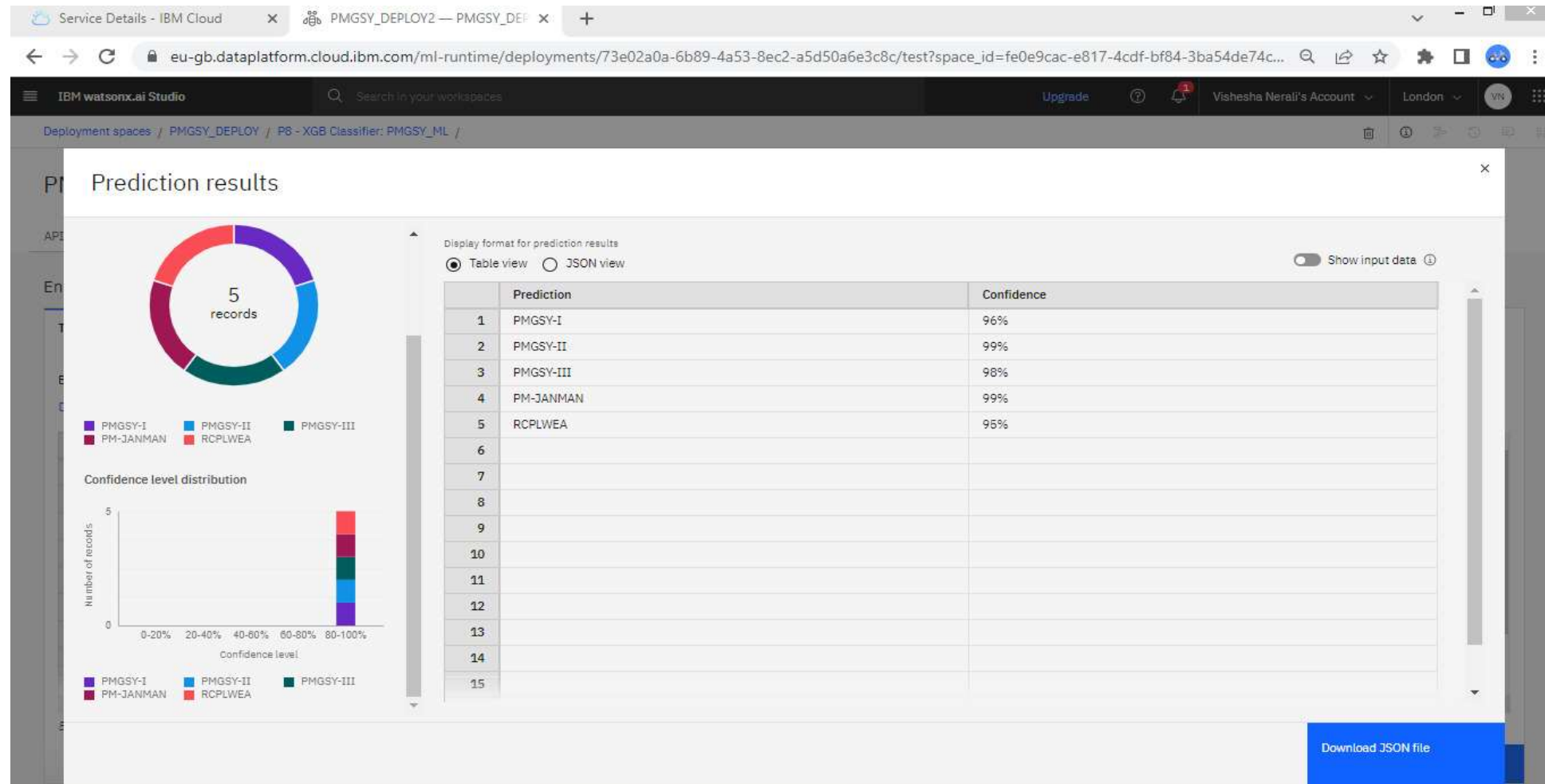
[Download CSV template](#) [Browse local files](#) [Search in space](#) [Clear all](#) x

	STATE_NAME (other)	DISTRICT_NAME (other)	NO_OF_ROAD_WORK_SANCTIONED (double)	LENGTH_OF_ROAD_WORK_SANCTIONED (double)	NO_OF_BRIDGES_SANCTIONED (double)	COST_OF_WORKS_SANCTIONED (double)
1	Andaman And Nicoba	North and Middle Andam	32	60.169	0	24.6908
2	Andhra Pradesh	Anantapur	14	125.7	0	56.0711
3	Andhra Pradesh	Visakhapatnam	18	134.545	5	81.59809
4	Chhattisgarh	Gariaband	35	65.295	0	
5	Chhattisgarh	Kondagaon	72	549.97	11	392.957
6						
7						
8						
9						

5 rows, 14 columns

Predict

# RESULT



---

# CONCLUSION

Summarize the findings and discuss the effectiveness of the proposed solution. Highlight any challenges encountered during the implementation and potential improvements. Emphasize the importance of Use of Machine Learning in the project minimizes errors and saves time and gives clear idea of project.

---

# FUTURE SCOPE

Discuss potential enhancements and expansions for the system. This could include incorporating additional data sources, optimizing the algorithm for better performance, and expanding the system to cover multiple cities or regions. Consider the integration of emerging technologies such as edge computing or advanced machine learning techniques.

# REFERENCES

Intelligent Classification of Rural Infrastructure Projects PMGSY dataset was collected from aikosh:

[https://aikosh.indiaai.gov.in/web/datasets/details/pradhan\\_mantri\\_gram\\_sadak\\_yojna\\_pmgsy.html](https://aikosh.indiaai.gov.in/web/datasets/details/pradhan_mantri_gram_sadak_yojna_pmgsy.html)

IBM cloud, watsonx.ai Studio, machine learning algorithms, best practices in data pre-processing, model evaluation, and articles were instrumental in developing the proposed solution. This could include academic papers on PMGSY prediction,

# IBM CERTIFICATIONS

In recognition of the commitment to achieve  
professional excellence



**VISHESHA NERALI**

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/bd1f594c-f0f3-4f90-99f8-a782c21b98dd>



# IBM CERTIFICATIONS



# IBM CERTIFICATIONS

**IBM SkillsBuild**

Completion Certificate



This certificate is presented to

VISHESHA NERALI

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





**THANK YOU**