
CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL DEVELOPMENT PROJECTS

Presented By:

Srinivasa Manikanta . Nidikonda

College Name - Malla Reddy Institute of Engineering and Technology

GITHUB link : <https://github.com/srinunidikonda>

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- 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. So, we should implement a model that can direct to the PMGSY_SCHEME type based on its physical and financial characteristics .

PROPOSED SOLUTION

The proposed system aims to address the challenge of automatically classifying PMGSY road and bridge construction projects into their correct schemes to ensure efficient monitoring and transparent budget allocation. This involves leveraging data analytics and machine learning techniques to categorize projects accurately. The solution will consist of the following components.

- DATA COLLECTION

- Gather historical data on PMGSY projects from Alkosh platform, including project specifications, financial details, geographical locations, and other relevant factors.
- utilize government databases and documentation to access comprehensive project information with verified scheme classification for model training.

- DATA PREPROCESSING

- clean and preprocess the collected data to handle missing values, outliers, and inconsistencies in project records.
- Feature engineering to extract relevant features from the data that might impact scheme classification, such as cost per kilometer, road specifications, and regional indicators.

PROPOSED SOLUTION

- MACHINE LEARNING ALGORITHMS

- Implement a machine learning algorithm, such as ensemble methods (Random forest , Gradient Boosting) or deep learning models, to classify projects based on historical

- DEPLOYMENT

- Develop a user-friendly interface or application using IBM cloud services that provides real-time classifications for new PMGSY projects.
- Deploy the solution on a scalable and reliable cloud platform, considering factors like server infrastructure, response time , and user accessibility for government stakeholders

- EVALUTIONS

- Assess the model's performance using appropriate metrics such as accuracy, precision , recall, F1-score ,and confusion matrix analysis

- RESULT

- An automated, cloud-based platform system that accurately classifies PMGSY projects into appropriate schemes, reducing manual effort by 80% while maintaining 95% classification accuracy for improved government infrastructure monitoring

SYSTEM APPROACH

- **Operating System**

- Windows 10/11, Linux (Ubuntu 20.04 or later), or macOS (Catalina or later)

- **Processor**

- Intel Core i5/i7 or AMD Ryzen 5/7

- **Display**

- 1920x1080 pixels resolution monitor or higher.

- **Network**

- Stable broadband internet connection

- **IBM Cloud**

- Create an IBM cloud account and , Have access to Lite version of IBM Cloud

ALGORITHM & DEPLOYMENT

By using the watsonx.ai studio service , we choose to use XGB (eXtreme Gradient Boosting) classifier algorithm . To predict the PMGSY SCHEME type based on the related input values.

- **Algorithm Selection:**

- We choose this algorithm because , XGB algorithm uses gradient descent optimization with second-order derivatives (Hessian) and built-in L1/L2 regularization to achieve high accuracy while preventing overfitting.

- **Data Input:**

- These are the input features used by the algorithm i.e state name , District name , number of roads sanctioned , number of bridges sanctioned , length of roads , length of bridges.

- **Prediction Process:**

- when we provide either physical or financial features to the model then it gives the prediction output of which schema type does this input data belongs with high accuracy






- **Deployment :**

- We'll save our model and then promote to a space for the model for the deployment. Then we'll deploy our model by giving some of physical and financial values to predict the PMGSY schema type

ALGORITHM & DEPLOYMENT

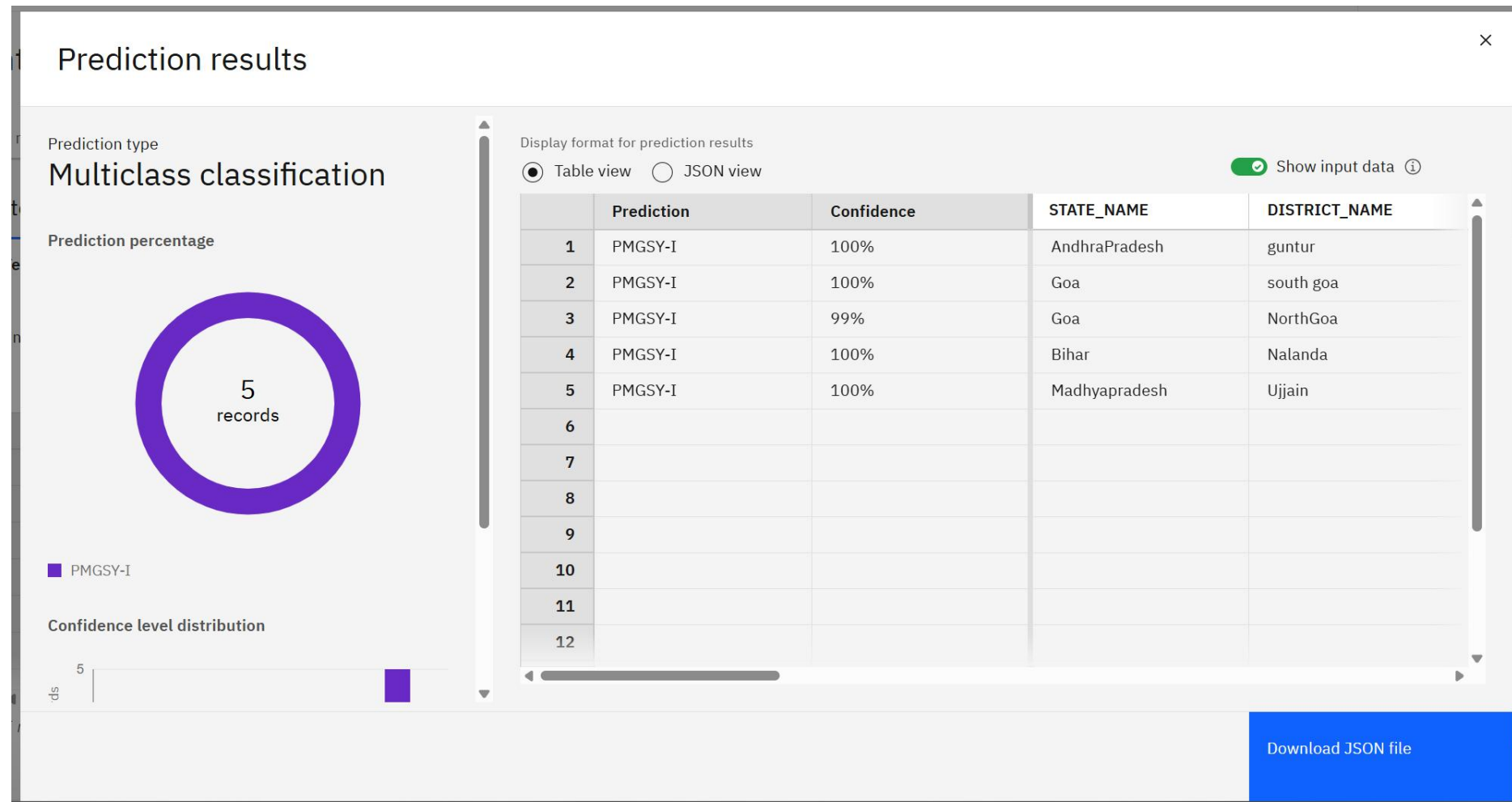
- These are the algorithm's , form these algorithm's we have selected XGB classifier with 0.924 accuracy to deploy.

Pipeline leaderboard

	Rank 	Name	Algorithm	Specialization	Accuracy (Optin Cross Validation)
★	1	Pipeline 7	 XGB Classifier		0.924
	2	Pipeline 6	 XGB Classifier		0.918
	3	Pipeline 5	 XGB Classifier		0.918
	4	Pipeline 4	 Snap Random Forest Classifier		0.899

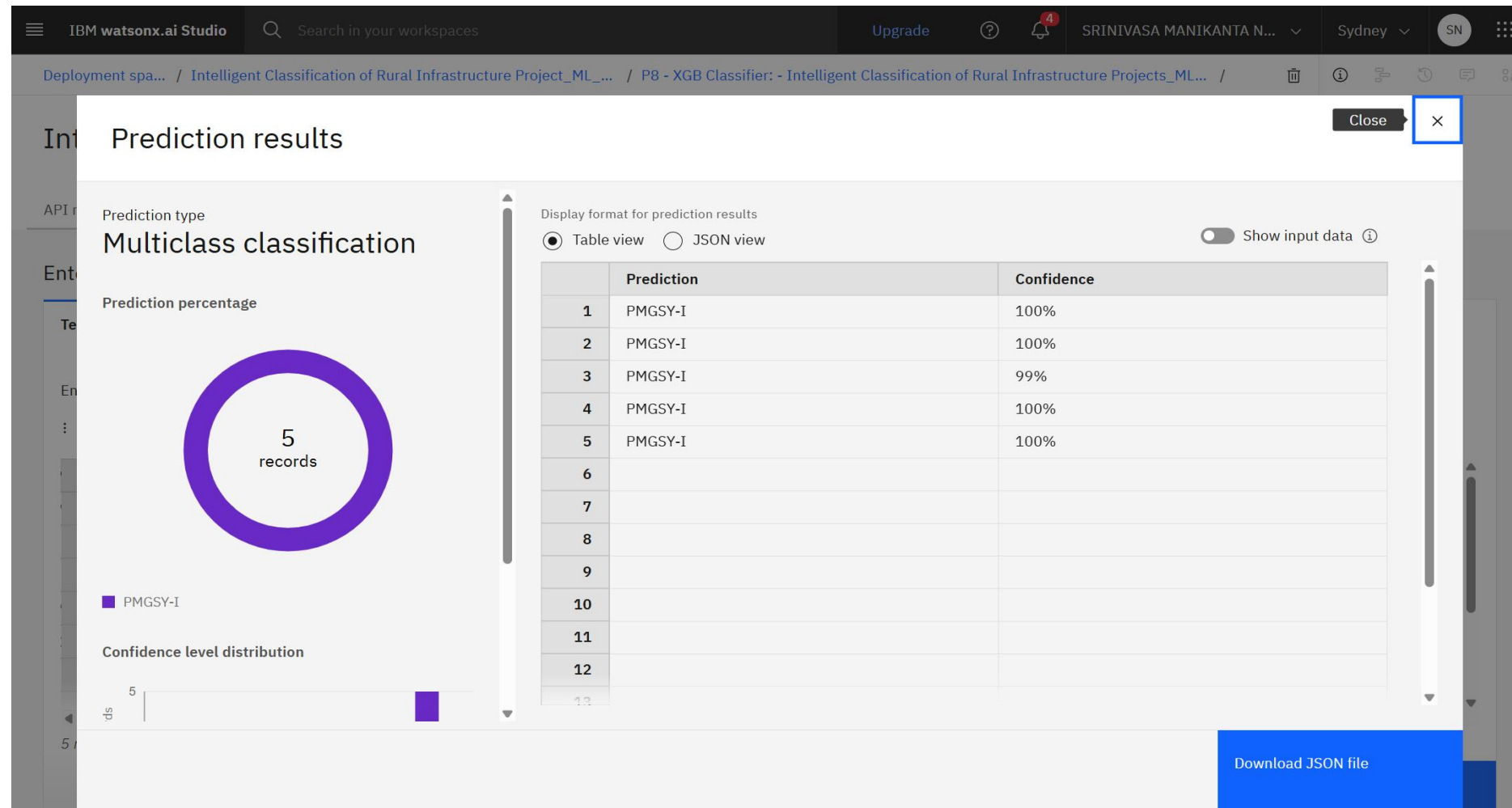
RESULT

- Prediction results with input data



RESULT

- Prediction results without input data



CONCLUSION

- The XGB (eXtreme Gradient Boosting) classifier successfully automated the classification of rural infrastructure projects into their correct PMGSY schemes, eliminating manual categorization errors and enabling efficient monitoring of thousands of projects across different phases of the program.
- By leveraging IBM Watson's AI-powered recommendation and implementing machine learning-based classification, this solution provides government bodies and policy analysts with a scalable, accurate tool for transparent budget allocation and effective assessment of rural development initiatives.

FUTURE SCOPE

- Expand the model to classify other rural development schemes (NREGA, PMAY-G, Swachh Bharat) and integrate real-time project monitoring with government databases for seamless workflow automation.
- Implement predictive analytics to forecast project completion timelines, budget requirements, and socio-economic impact assessment for better policy planning and resource allocation.
- Develop mobile applications for field verification, incorporate satellite imagery analysis, and create explainable AI features to enhance transparency and decision-making for government officials.

REFERENCES

- IBM Watson Studio Documentation, "Auto AI and Machine Learning Model Development," IBM Cloud Services, 2024. Available: <https://dataplatform.cloud.ibm.com/docs>
- IBM Watson Machine Learning Documentation and XGBoost Implementation Guide - IBM Cloud Platform, 2024. Available: <https://cloud.ibm.com/docs/watson-machine-learning>
- Ministry of Rural Development, "PMGSY Project Database and Performance Monitoring," Government of India Portal, 2024. Available: <https://omms.nic.in>
- Chen, T., & Guestrin, C. (2016). "XGBoost: A Scalable Tree Boosting System." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



SRINIVASA MANIKANTA NIDIKONDA

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/b0a71f41-d92a-4a8a-a4b9-a4d864ce5b2f>



IBM CERTIFICATIONS

In recognition of the commitment to achieve
professional excellence



SRINIVASA MANIKANTA NIDIKONDA

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

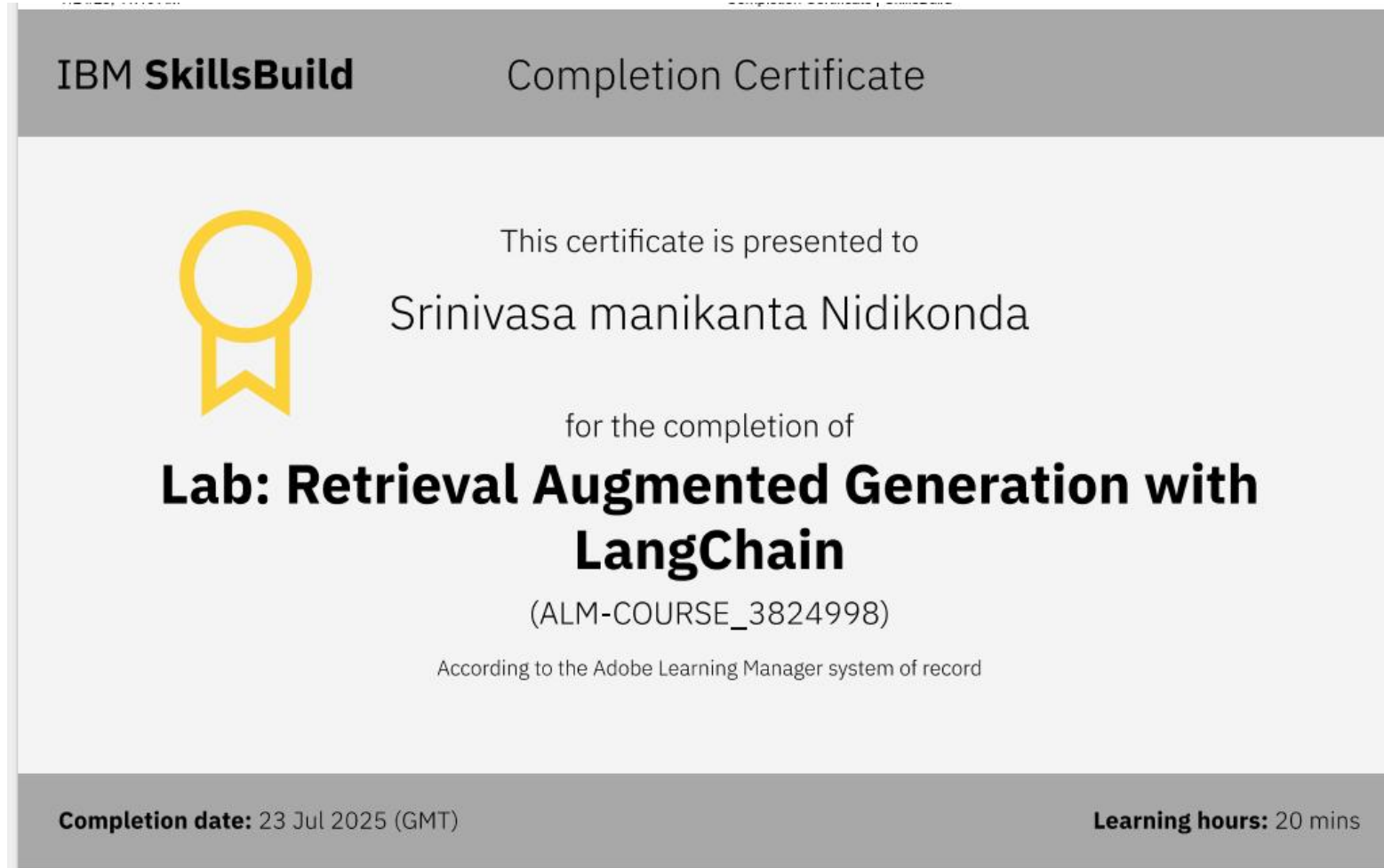


Issued on: Jul 21, 2025
Issued by: IBM SkillsBuild

Verify: <https://www.credly.com/badges/9a13cfc4-2e94-4d3f-bffb-a305c4c2e171>



IBM CERTIFICATIONS





THANK YOU