CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS (PMGSY)

(Problem Statement No. 35)

Presented By:
1. IYSHWARYA S – PUDUCHERRY TECHNOLOGICAL UNIVERSITY - C.S.E.

edunet
foundation

OUTLINE

- **Problem Statement**

- **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

- The proposed system aims to build an intelligent model to automatically classify PMGSY projects.

- Data Collection:

  - Utilize the historical project data from the Pradhan Mantri Gram Sadak Yojana (PMGSY) dataset available on AI Kosh.

- Data Preprocessing:

  - The collected data will be cleaned and preprocessed to handle any missing values or inconsistencies.

  - This step is automated using the capabilities of IBM's AutoAI tool.

- Machine Learning Algorithm:

  - Implement a multi-class classification algorithm to predict the PMGSY_SCHEME based on project characteristics.

  - IBM AutoAI will be used to automatically train and compare multiple algorithms to select the best-performing one.

- Deployment:

  - The final trained model will be deployed as a real-time web service on the IBM Cloud (Watsonx.ai) platform.

- Evaluation:

  - The model's performance will be assessed using the 'Accuracy' metric to determine how correctly it classifies the projects.

edunet
foundation

SYSTEM APPROACH

- System Requirements:

  - A cloud-based platform: IBM Watsonx.ai Studio on IBM Cloud.

  - An automated tool for model development: IBM AutoAI.

- Library Required to build the model:

  - ibm-watsonx-ai: For interacting with the IBM platform.

  - scikit-learn: For machine learning components.

  - autoai-libs: For running AutoAI generated pipelines.

  - pandas: For data manipulation.

ALGORITHM & DEPLOYMENT

- Algorithm Selection:
  - IBM's AutoAI was chosen to automate the algorithm selection process. It trained and evaluated several classification models like Logistic Regression, Decision Trees, and XGBoost.
  - The XGBoost Classifier was automatically selected as the best model based on its high accuracy for this problem.

- Data Input:
  - The algorithm uses input features like STATE_NAME, DISTRICT_NAME, LENGTH_OF_ROAD_WORK_SANCTIONED, COST_OF_WORKS_SANCTIONED, and other financial/physical parameters.
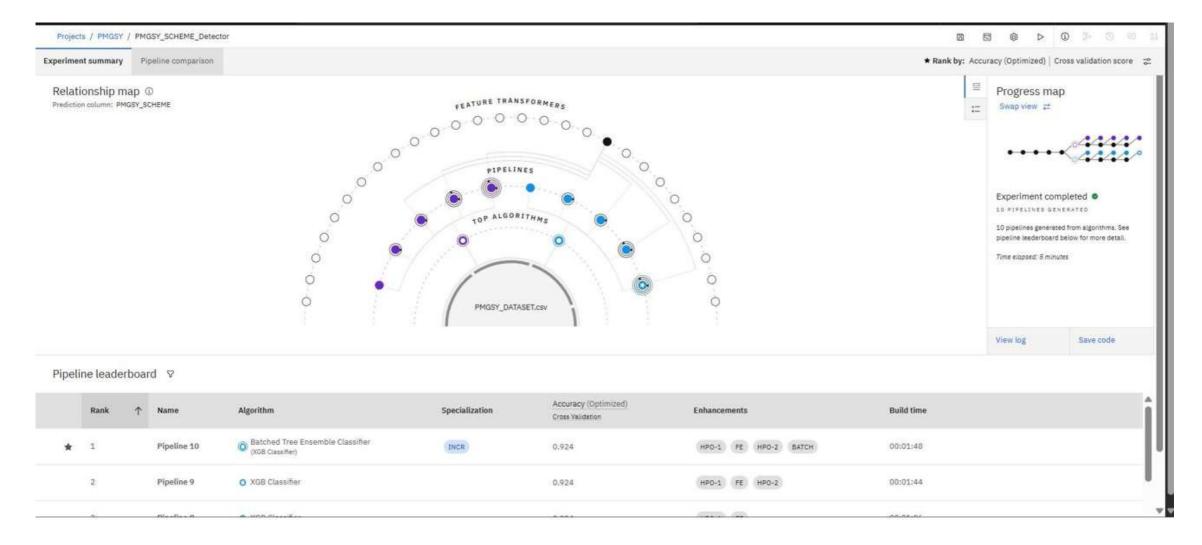
- Training Process:
  - AutoAI automated the entire training pipeline, including feature engineering, hyperparameter tuning, and model validation.
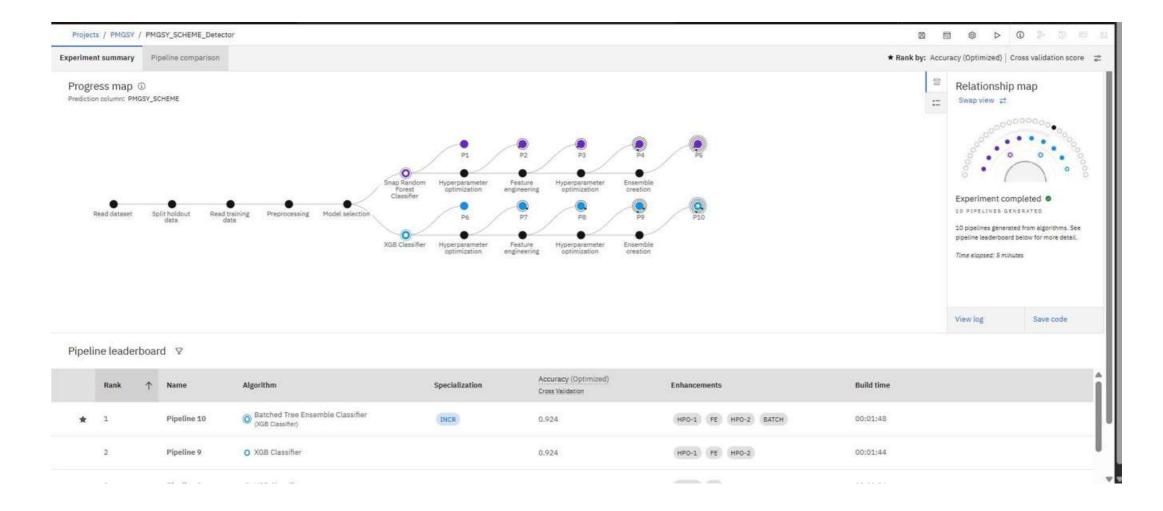
- Deployment:
  - The best pipeline was saved and deployed as an Online Deployment (Web Service) on IBM Watsonx.ai. This creates a REST API endpoint for getting real-time predictions.
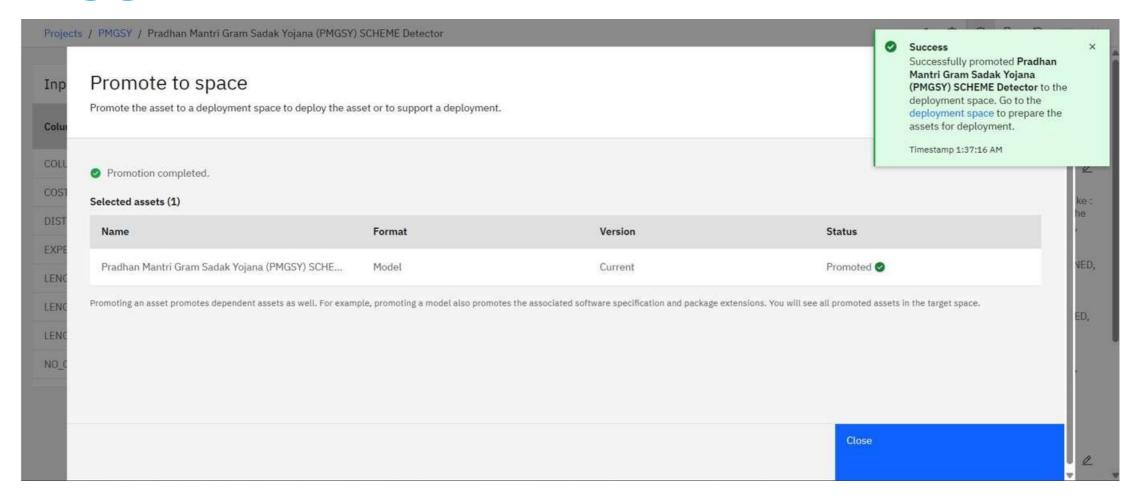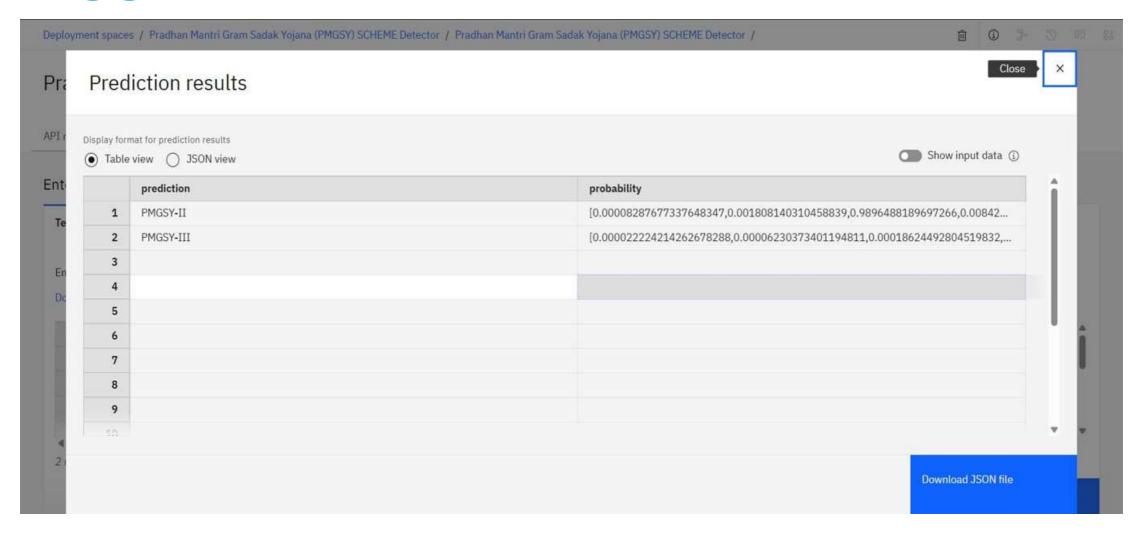
RESULT

RESULT

RESULT

RESULT

RESULT

CONCLUSION

- The project successfully demonstrated an end-to-end solution for automatically classifying rural infrastructure projects using IBM's AutoAI platform.

- The automated system significantly reduces the manual effort, time, and errors involved in project classification.

- This leads to more efficient monitoring and transparent financial planning for government bodies.

- The deployment on IBM Cloud makes the solution scalable and accessible for real-world use.

FUTURE SCOPE

- Real-time Data Integration: The system can be enhanced by incorporating real-time project progress data from the field.

- Performance Dashboard: Develop a user-friendly dashboard that visualizes the classification results and project statuses across different regions.

- Predictive Analytics: Expand the model's capability to not just classify, but also predict potential project delays or budget overruns.

- Advanced AI Techniques: Integrate advanced techniques like Natural Language Processing (NLP) to analyze project reports or notes for more nuanced classification.

REFERENCES

- Dataset: Pradhan Mantri Gram Sadak Yojana (PMGSY) Dataset - Provided by AI Kosh, Government of India.

    - AI Kosh dataset link – https://aikosh.indiaai.gov.in/web/datasets/details/pradhan_mantri_gram_sadak_yojna_pmgsy.html

- Platform Documentation: IBM Watsonx.ai Official Documentation.

- Research Papers: Relevant papers on machine learning applications in public policy and infrastructure management.

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with
Artificial Intelligence
IBM SkillsBuild

IYSHWARYA S

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

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

Verify: https://www.credly.com/badges/a940016c-0756-4dce-be8d-0ff2689d829c

IBM.

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud:
Envisioning
Your Solution

IBM SkillsBuild

IYSHWARYA S

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/ce32e94e-efcf-413a-8776-3608742523d3

IBM.

edunet
foundation

IBM CERTIFICATIONS

IBM **SkillsBuild**   Completion Certificate

This certificate is presented to

IYSHWARYA S

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 26 Jul 2025 (GMT)   **Learning hours:** 20 mins

edunet
foundation

THANK YOU