CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS

Presented By:
1. BHOJANAPALLI NAGA VENKATA DURGA OM JNAANESH – KL University - CSE

edunet foundation

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.

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 effectively automate the classification of PMGSY road and bridge projects into their respective schemes, a comprehensive approach involving data preparation and model deployment must be followed.

Data Pre-processing:

1. Clean and normalize project datasets, handling missing values and outliers.
2. Encode categorical variables and balance class distribution to improve model learning.

Machine Learning Algorithm:

1. Select and train a suitable classifier (e.g., Random Forest or Logistic Regression) based on data characteristics.
2. Tune hyperparameters and perform cross-validation to optimize model performance.

Deployment:

1. Deploy the trained model on IBM Cloud with a REST API for real-time project classification.
2. Set up automated data pipelines to feed new project data into the model for continuous prediction.

Evaluation:

1. Measure model accuracy, precision, recall, and F1-score to assess classification quality.
2. Analyze confusion matrix and conduct error analysis to identify and address model weaknesses.

SYSTEM APPROACH

Platform Used:
IBM Watson Studio – Used for AutoAI pipeline creation, model training, and deployment.

Dataset Source:
AI-KOSH dataset – Pradhan Mantri Gram Sadak Yojna (PMGSY).

Data Handling:
IBM Watson's built-in data flow managed preprocessing, including column selection and cleaning.

Model Selection & Training:
AutoAI in IBM Watson automatically tested multiple algorithms and selected the best-performing one with hyperparameter tuning.

Deployment Environment:
IBM Cloud Lite – Model deployed as a REST API for real-time access.

Interface:
Prediction results are tested using the Watson with options to upload CSV, JSON, or manual input.

ALGORITHM & DEPLOYMENT

1. **Algorithm Selection:**

- Analyze the dataset's features (e.g., numeric, categorical, size, class distribution) to choose an appropriate machine learning algorithm such as Random Forest, Logistic Regression, or other ensemble/classification algorithms.
- Consider the trade-offs: For example, Random Forest is robust and handles diverse data types well, which is suitable for project classification tasks.

2**. Data Input:**

- Prepare the raw project data by cleaning, encoding categorical variables, and normalizing numeric attributes.
- Format the data into a suitable structure (CSV) that aligns with the input requirements of the chosen ML model.

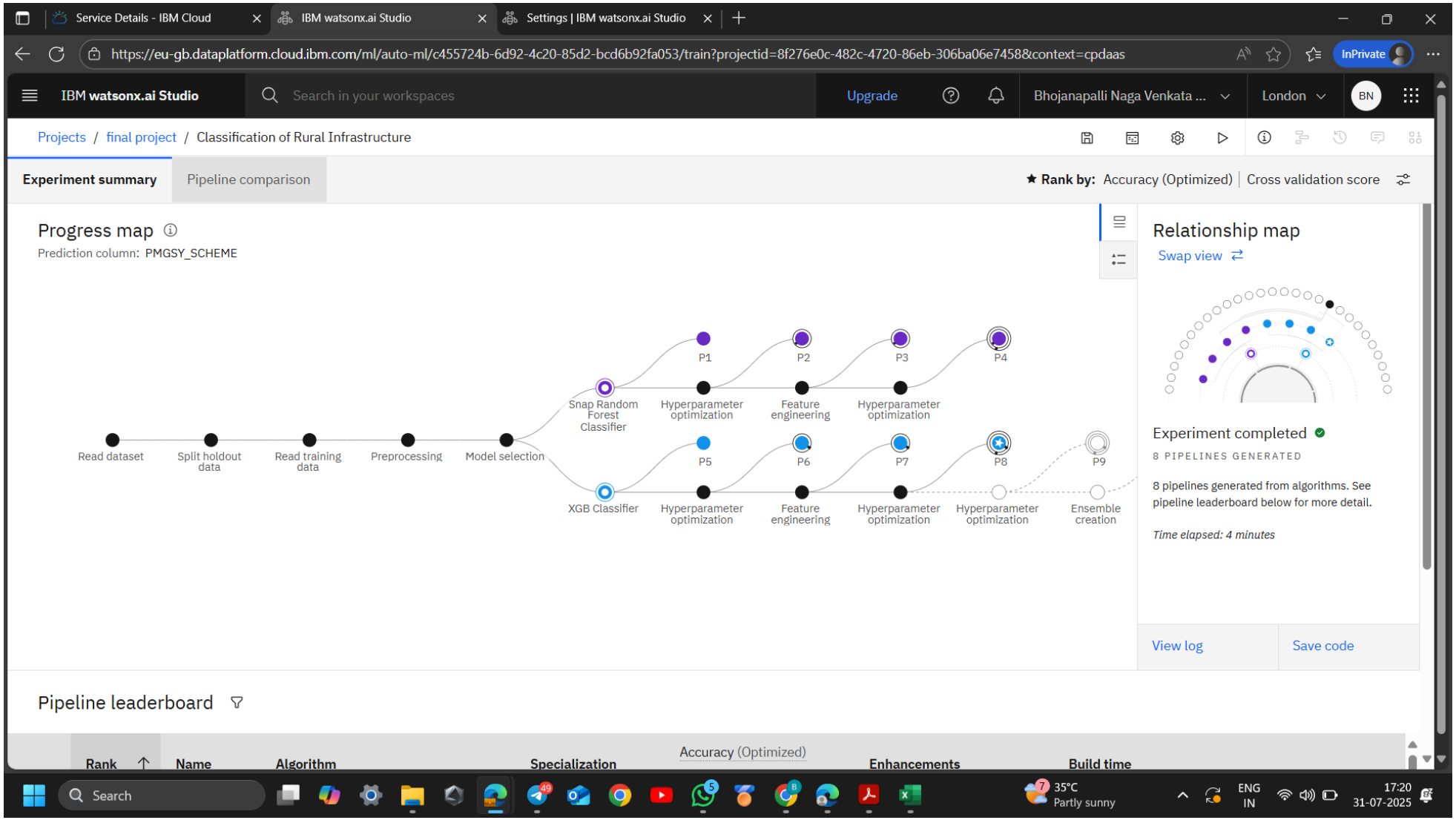
3. **Training Process:**

- Split the dataset into training and validation sets.
- Train the selected algorithm on the training portion, using techniques like cross-validation and hyperparameter tuning to optimize for accuracy.
- Evaluate the model using validation data, adjusting parameters and pre-processing steps as needed.

4. **Prediction Process:**

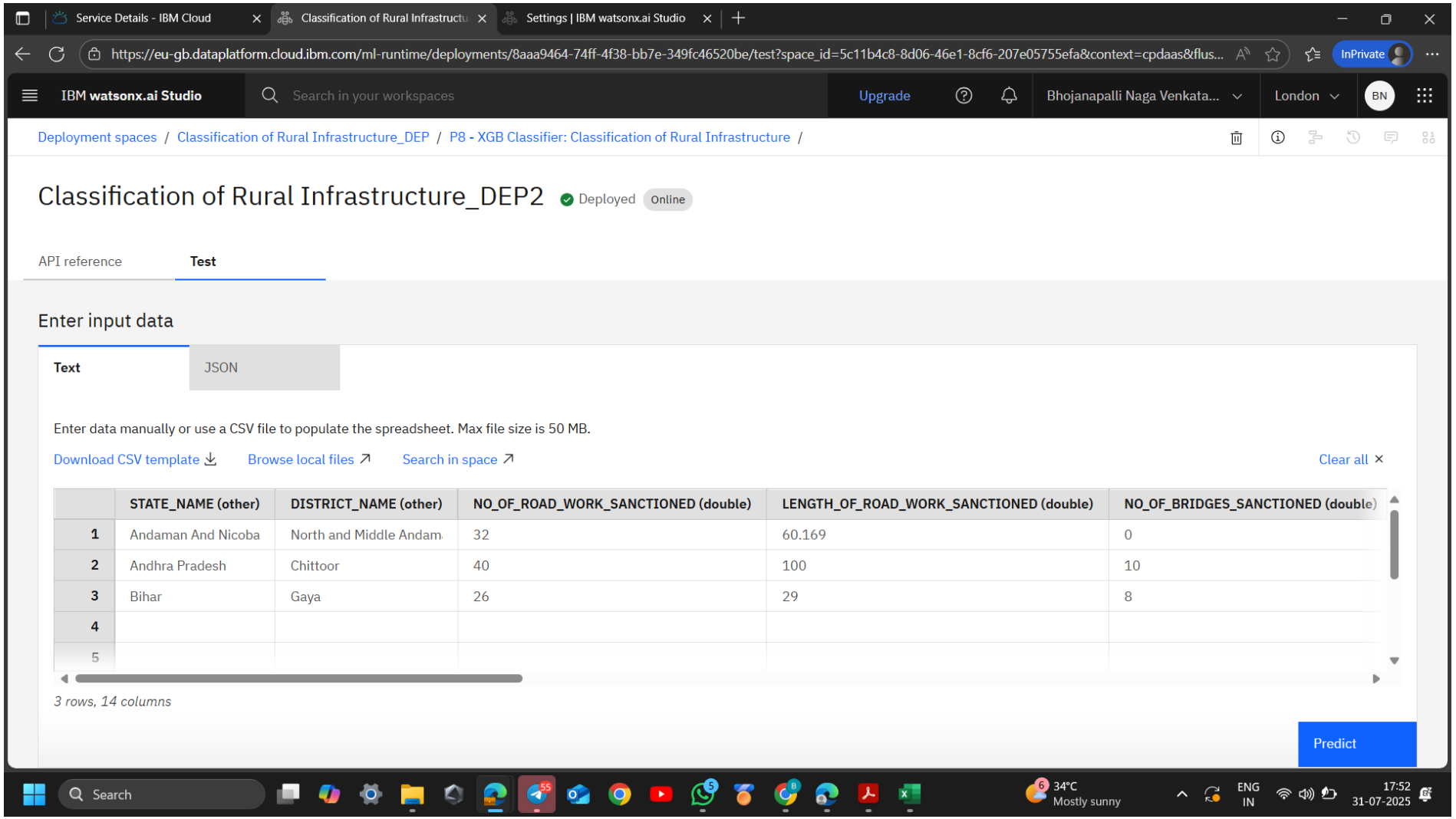
- Deploy the trained model (e.g., on IBM Cloud) as a RESTful API or web service.
- For real-time predictions: New project data is sent to the API, which preprocesses the input, applies the trained model, and returns the predicted PMGSY scheme.
- Continuously monitor performance by tracking prediction results and updating the model as new data becomes available.

RESULT



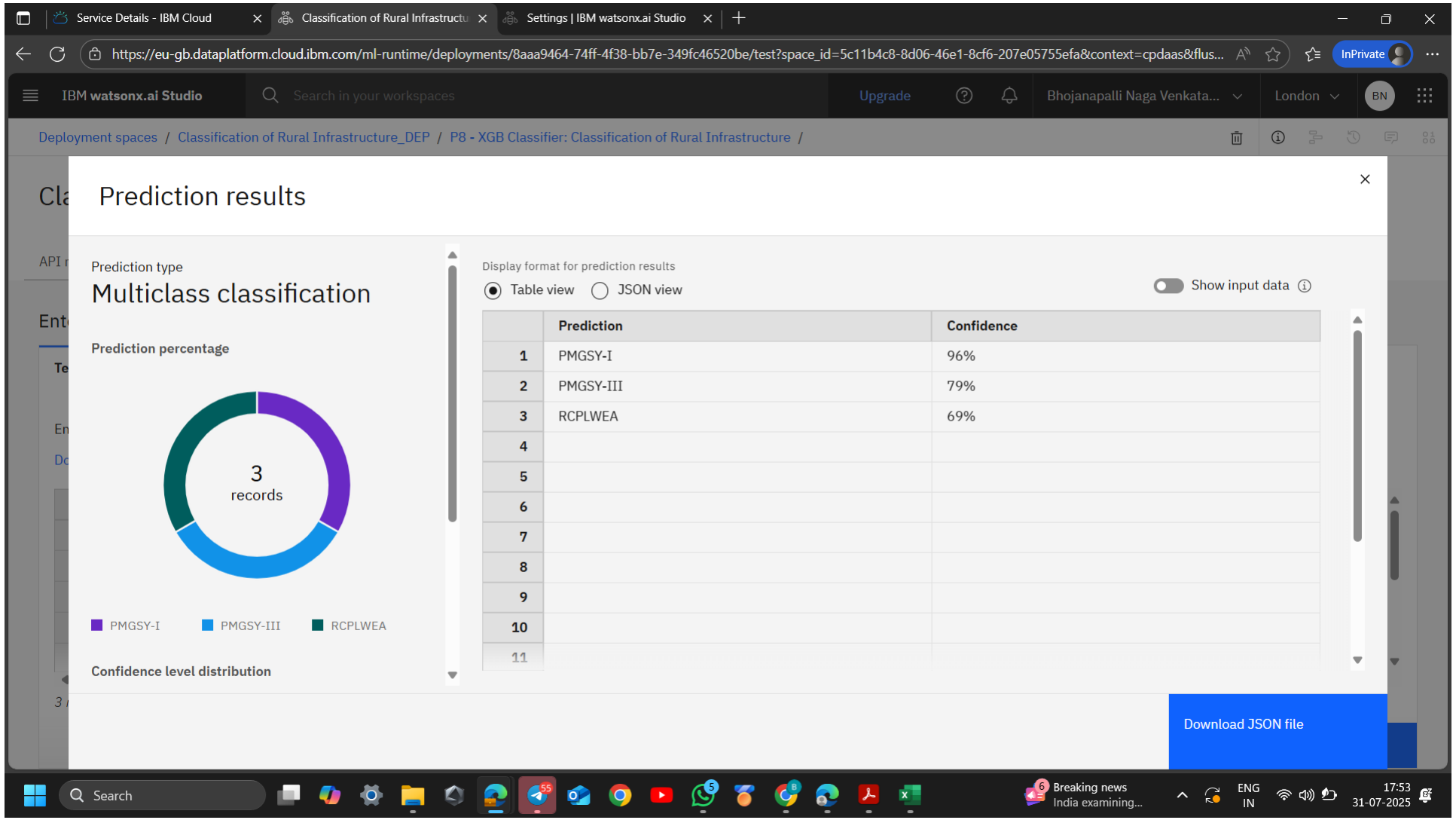
edunet foundation

RESULT



edunet foundation

RESULT



edunet foundation

CONCLUSION

- Automated classification of PMGSY road and bridge projects was successfully developed using machine learning based on physical and financial data.
- Data preprocessing, algorithm selection, and IBM Cloud deployment improved accuracy, scalability, and efficiency over manual methods.
- The system helps streamline project monitoring, budget allocation, and supports real-time prediction for ongoing projects.
- The framework allows continuous updates and can adapt as more data becomes available.
- Future enhancements may include integrating more data sources, applying advanced algorithms, and expanding deployment features to better support rural infrastructure development.

FUTURE SCOPE

1. **Use More Data:**

- Add geographic and satellite data to make predictions better.
- Use real-time data from sensors or mobile apps to track projects live.

2. **Improve the Algorithm**

- Try advanced models like deep learning to get more accurate results.
- Automatically pick the best features to make the model faster and smarter.

3. **Expand to More Areas**

- Scale the system to cover multiple cities, states, or all over India.
- Customize the model based on local conditions and construction practices.

4. **Use New Technologies**

- Make the model's decisions easy to understand for officials using explainable AI.
- Set up automatic model updates as new data comes in.

5. **Help Policymakers and Collaborate**

- Create dashboards to visualize results and aid decision-making.
- Work with government and researchers to share data and improve the system.

REFERENCES

- AI Kosh Dataset: https://aikosh.indiaai.gov.in/web/datasets/details/pradhan_mantri_gram_sadak_yojna_pmgsy.html
- IBM Watson Studio and IBM cloud resources.

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

Bhojanapalli Naga Venkata Durga Om Jnaanesh

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/239f3389-d7fa-4c66-a829-5e671ca15aba

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild
IBM

Bhojanapalli Naga Venkata Durga Om Jnaanesh

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/42ba153f-9e72-4f46-bb6a-5f9364672b9d

IBM

edunet foundation

IBM CERTIFICATIONS

7/24/25, 7:59 PM Completion Certificate | SkillsBuild

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Bhojanapalli Naga Venkata Durga Om Jnaanesh

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

edunet foundation