

---

# **CAPSTONE PROJECT**

## **INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECT**

**Presented By:**  
**Siva Prasad M L - College of Engineering, Munnar**  
**Computer science and Engineering.**

# 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

- The proposed system aims to automate the classification of rural road and bridge projects into their correct PMGSY scheme. The solution will consist of the following components:
- Data Collection:
  - Collected real world rural project data from the AI Kosh Government Open Dataset Portal.
  - It includes features such as state , district , length of the road and bridge , cost of work sanctioned , expenditure , work completion status and the target column PMGSY SCHEME.
- Data Preprocessing:
  - Handled missing values and removed empty values
  - Encoded categorical variables using Auto AI's built-in transformations.
  - Normalized numerical features where required.
- Machine Learning
  - Tested multiple classification models like Logistic Regression , Random Forest Classifier, Gradient Boosted Trees. Performed automatic hyperparameter turning.
- Deployment:
  - Final model was developed using IBM Watson Deployment Space.
  - An API endpoint was generated for future integration.
- Evaluation:
  - Evaluated with Auto-AI leadership metrics like Accuracy , Precision , Recall and Confusion matrix.

# SYSTEM APPROACH

**The System Approach section shows the overall strategy and methodology for developing and implementing the rural infrastructure project.**

- **System requirements**

- Cloud-based ML platform(IBM Watson Studio)

- Dataset from India AI Kosh - PMGSY

- **Library required to build the model**

- IBM Watsonx.ai Studio (Auto-AI)
  - Python
  - GitHub (project repository)

# ALGORITHM & DEPLOYMENT

- **The machine learning algorithm chosen for Intelligent Classification of Rural Infrastructure are specified below :**
- **Algorithm Selection:**
  - Auto- AI automatically selected classification pipelining using Random Forest , Gradient Boosted Trees and Logistic Regression.
- **Data Input:**
  - Road/bridge length , Number of sanctioned works , Cost of works , State/District information etc..
- **Training Process:**
  - Trained the model by preprocessing data, testing multiple algorithms , turning parameters and selecting best pipeline based on accuracy and F1 score for deployment.
- **Prediction Process:**
  - Model takes project details as input (like length, cost, completion status) and predicts the corresponding PMGSY scheme using the trained classification pipeline.

# RESULT

Deployment spaces / Rural Infrastructure / PB • XGB Classifier: Rural Infrastructure Project /

Classification of Rural Infrastructure 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

	STATE_NAME (other)	DISTRICT_NAME (other)	NO_OF_ROAD_WORK_SANCTIONED (double)	LENGTH_OF_ROAD_WORK_SAN
1	Tamil Nadu	Kanniyakumari	106	140.055
2	Kerala	Idukki	12	58.557
3	Maharashtra	Pune	32	216.72
4				
5				
6				

## Prediction results

Prediction type

Multiclass classification

Prediction percentage



PMGSY-I

PMGSY-II

PMGSY-III

Confidence level distribution



PMGSY-I

PMGSY-II

PMGSY-III

Display format for prediction results

☒ Table view ☐ JSON view

Show input data ⓘ

	Prediction	Confidence
1	PMGSY-I	100%
2	PMGSY-II	99%
3	PMGSY-III	100%
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

---

# CONCLUSION

- Successfully built and deployed a multi-class classifier using Auto-AI.
- The solution can assist government bodies in Automatic scheme tagging.
- Reduces manual errors and boosts monitoring efficiently.



# FUTURE SCOPE

- Expand model to include more features like geography and contractor data.
- Scale model to include real-time project updates.
- Integrate with government dashboards via API.

# REFERENCES

- India AI Kosh : PMGSY Dataset.
  - IBM Watson Studio Documentation.
  - Auto-AI Toolchain Guides.
  - Government of India PMGSY Guidelines.
- 
- **GitHub repository Link :** <https://github.com/Sivaprasad2k/rural-infra-autoai>

# IBM CERTIFICATIONS

Screenshot/credly certificates (Getting started with AI)



# IBM CERTIFICATIONS

- Screenshot/credly certificate( Journey to Cloud)



# IBM CERTIFICATIONS

- Screenshot/ credly certificate( RAG Lab)





**THANK YOU**