CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECT

Presented By:
1. Kalyanaruba.M-K Ramakrishnan College of Engineering - AIML

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.

Github link: https://github.com/KALYANARUBA/Project-PMGSY

PROPOSED SOLUTION

- The proposed system aims automates the classification of rural infrastructure projects under PMGSY schemes (PMGSY-I, PMGSY-II, RCPLWEA, etc.) using machine learning on IBM Cloud, enabling faster, accurate, and scalable decision-making for planners and policymakers. The solution will consist of the following components:

Data Collection:

Use AI KOSH dataset on rural roads and bridges.

Collect physical and financial attributes of each project.

Data Preprocessing:

Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.

Feature engineering to extract relevant features from the data that might impact bike demand.

Machine Learning Algorithm:

Implement a machine learning algorithm, such as a Random Forest Classifier to predict the correct PMGSY scheme based on project attributes.

Applied Label Encoding to convert categorical variables and trained the model on structured physical and financial data.

Deployment:

Model developed and tested using IBM Cloud Jupyter environment and IBM Cloud Object Storage.

Can be integrated into a cloud-based interface or API for real-time classification of new project data.

Evaluation:

Evaluated using accuracy and detailed classification report (precision, recall, F1-score).Fine-tune the model based on feedback and continuous monitoring of prediction accuracy.

Result:

Automated classification improved efficiency and accuracy over manual processes.

edunet foundation

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the intelligent classification of rural infrastructure system. Here's a suggested structure for this section:

System requirements:

- **Platform: IBM Cloud (Lite Tier)**
- **Notebook Environment: jupyder Notebook (via IBM Watson Studion)**
- **Storage: IBM Cloud Object Storage**
- **Compute: IBM Cloud-based Python runtime model for execution.**
- **Data Format: CSV (PMGSY dataset from AI-KOSH)**

Library required to build the model:

- **Pandas, sklearn, ibm_boto3 and botocore**

ALGORITHM & DEPLOYMENT

- **Algorithm Selection**:
  A Random Forest Classifier was chosen for its accuracy, ability to handle multiclass classification, and robustness with both numerical and categorical data.

- **Data Input:**
  Input features include physical (e.g., road length, type) and financial (e.g., cost, expenditure) attributes, along with encoded state and district names.

- **Training Process:**
  The dataset was split (80/20), categorical columns were label-encoded, and the model was trained to recognize patterns for scheme classification.

- **Prediction Process**:
  The trained model predicts PMGSY schemes for new records and is evaluated using accuracy score and classification report metrics.

RESULT

- **The Random Forest Classifier achieved a high accuracy in classifying PMGSY projects into their correct schemes, as validated on the test dataset.**
- **Accuracy Score and Classification Report revealed strong model performance, with precision and recall values indicating reliable classification across multiple scheme categories.**
- **Visualizations like confusion matrix and bar plots (actual vs. predicted) further demonstrate the model's effectiveness and minimal misclassification.**

Accuracy: 0.8926940639269406

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	6
1	0.96	0.92	0.94	143
2	0.86	0.94	0.90	128
3	0.86	0.89	0.87	150
4	0.50	0.09	0.15	11
accuracy			0.89	438
macro avg	0.84	0.77	0.77	438
weighted avg	0.89	0.89	0.89	438

CONCLUSION

- **The system helps automatically identify which PMGSY scheme a project belongs to using machine learning.**
  The Random Forest model gave accurate results and worked well with different types of data.
  This system can help government officials make quicker and better decisions for planning and managing projects.
  Using IBM Cloud makes the system easy to access, reliable, and ready to be used on a large scale.

FUTURE SCOPE

- **The system can be improved by integrating more data sources, optimizing algorithms for better accuracy, and expanding coverage to more regions. Future work may explore edge computing and advanced ML models for real-time, scalable deployment.**

REFERENCES

- Research articles on infrastructure classification using ML (e.g., IEEE, Springer)
- Scikit-learn documentation for model building and evaluation
- IBM Cloud and Watson Studio official docs for deployment
- AI Kosh Dataset: https://aikosh.indiaai.gov.in
- Github link: https://github.com/KALYANARUBA/Project-PMGSY

IBM CERTIFICATIONS

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Kalyanaruba Mani

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/76a1d060-7363-4661-9ec9-e08d3342ceb3

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Kalyanaruba Mani

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 17 Jul 2025 (GMT)

Learning hours: 20 mins

edunet foundation

THANK YOU

edunet foundation