CAPSTONE PROJECT

INTELLIGENT CLASSIFICATION OF RURAL INFRASTRUCTURE PROJECTS

Presented By :

1. Student Name - Preetham H S
2. College Name - Dayananda Sagar University
3. Department – Computer Science And Engineering

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result 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 is designed to automate the classification of road and bridge projects under various PMGSY schemes using a machine learning model. The solution eliminates manual errors, accelerates classification, and enhances project monitoring and decision-making for policy makers.

Key Components of the Solution:

- Dataset Source: AI Kosh's official PMGSY dataset, containing details of thousands of infrastructure projects with attributes like project length, cost, construction year, location, and target scheme.
- ML Objective: Build a multi-class classification model that predicts the appropriate PMGSY_SCHEME (e.g., PMGSY-I, PMGSY-II, RCPLWEA) using project attributes.

PROPOSED SOLUTION

- ML Pipeline Automation: Use IBM Watsonx.ai AutoAI to automatically:
  - Clean and preprocess the data
  - Identify the most influential features
  - Train, test, and tune multiple classification algorithms
  - Rank models based on performance metrics
- Deployment: Host the trained model on IBM Cloud as an online service accessible for real-time scheme predictions.
- Use Case: Government agencies can use the deployed model to classify upcoming or completed projects quickly for audits, budget distribution, or analytics.

SYSTEM APPROACH

- Platform: IBM Watsonx.ai Studio
- Storage: IBM Cloud Object Storage
- Dataset: PMGSY_DATASET.csv (AI Kosh)
- Columns: Length, Cost, Year, State, District, etc.
- Target: PMGSY_SCHEME
- No-code ML: AutoAI automatically performs preprocessing, model training & tuning

ALGORITHM & DEPLOYMENT

Algorithm: AutoAI handles end-to-end model generation using a collection of classic classification algorithms:

- Models Evaluated:
  - Logistic Regression
  - Decision Tree Classifier
  - Random Forest
  - Gradient Boosted Trees
- AutoAI evaluates models using metrics like accuracy, precision, recall, and F1-score, and automatically ranks them in a leaderboard.

Training Process:

- Dataset is split into training and test sets internally by AutoAI
- Automated feature scaling and encoding are performed
- AutoAI uses cross-validation and scoring to select the optimal pipeline

ALGORITHM & DEPLOYMENT

Prediction Process:

- The best model is saved and deployed as an online REST API
- The deployed model takes project characteristics as input and returns the predicted PMGSY scheme
- Interface available through IBM Watsonx.ai deployment testing tab or via API integration

Deployment Architecture:

- Model → Promoted to Deployment Space → Deployed Online
- Once deployed, predictions can be made via web console or integrated into government dashboards or web portals for policy analysis

RESULT

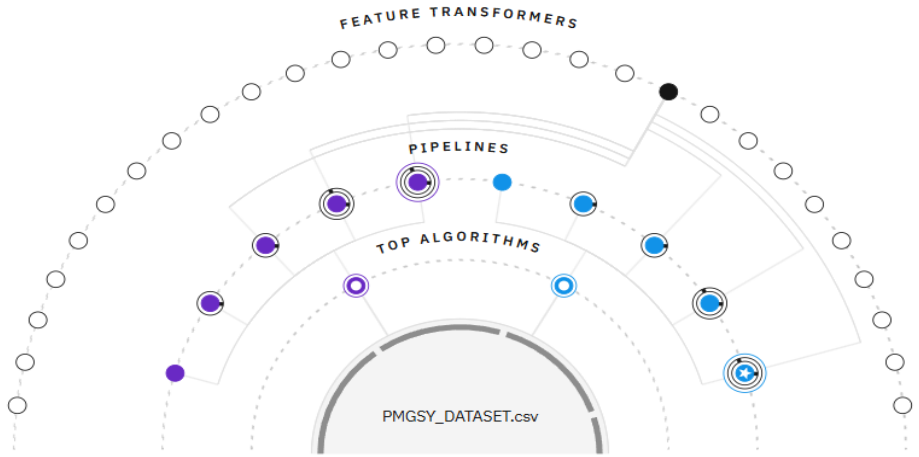
IBM watsonx.ai Studio | Search in your workspaces | Upgrade | Preetham H S's Account | Sydney | PH

Projects / Pradhan Mantri Gram Sadak Yojana / Intelligent Classification of Rural Infrastructure Projects

Experiment summary | Pipeline comparison | ★ Rank by: Accuracy (Optimized) | Cross validation score

Relationship map

Prediction column: PMGSY_SCHEME



Progress map

Swap view

Experiment completed

10 PIPELINES GENERATED

10 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 13 minutes

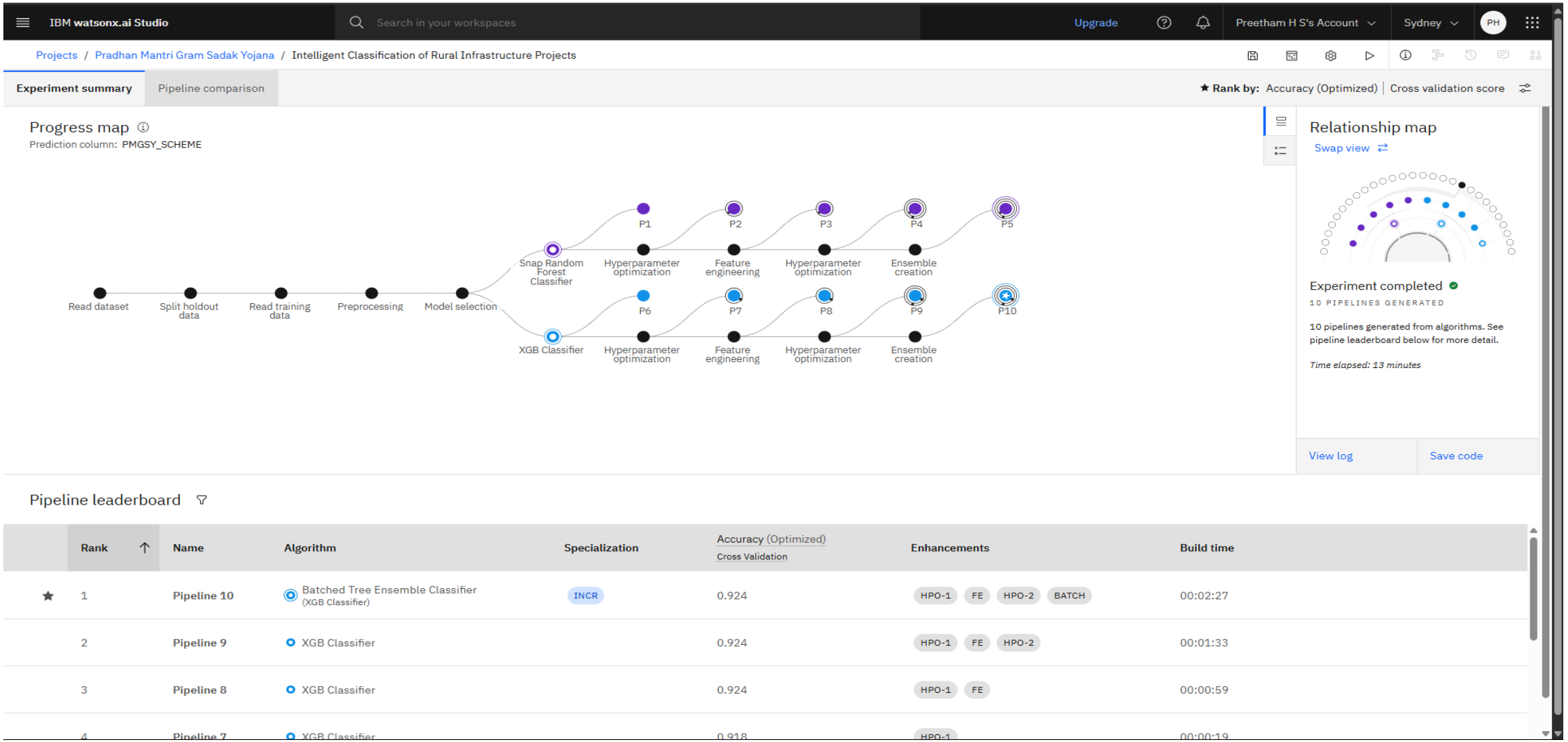
View log | Save code

Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 10	Batched Tree Ensemble Classifier (XGB Classifier)	INCR	0.924	HPO-1 FE HPO-2 BATCH	00:02:27
	2	Pipeline 9	XGB Classifier		0.924	HPO-1 FE HPO-2	00:01:33
	3	Pipeline 8	XGB Classifier		0.924	HPO-1 FE	00:00:59
	4	Pipeline 7	XGB Classifier		0.918	HPO-1	00:00:19

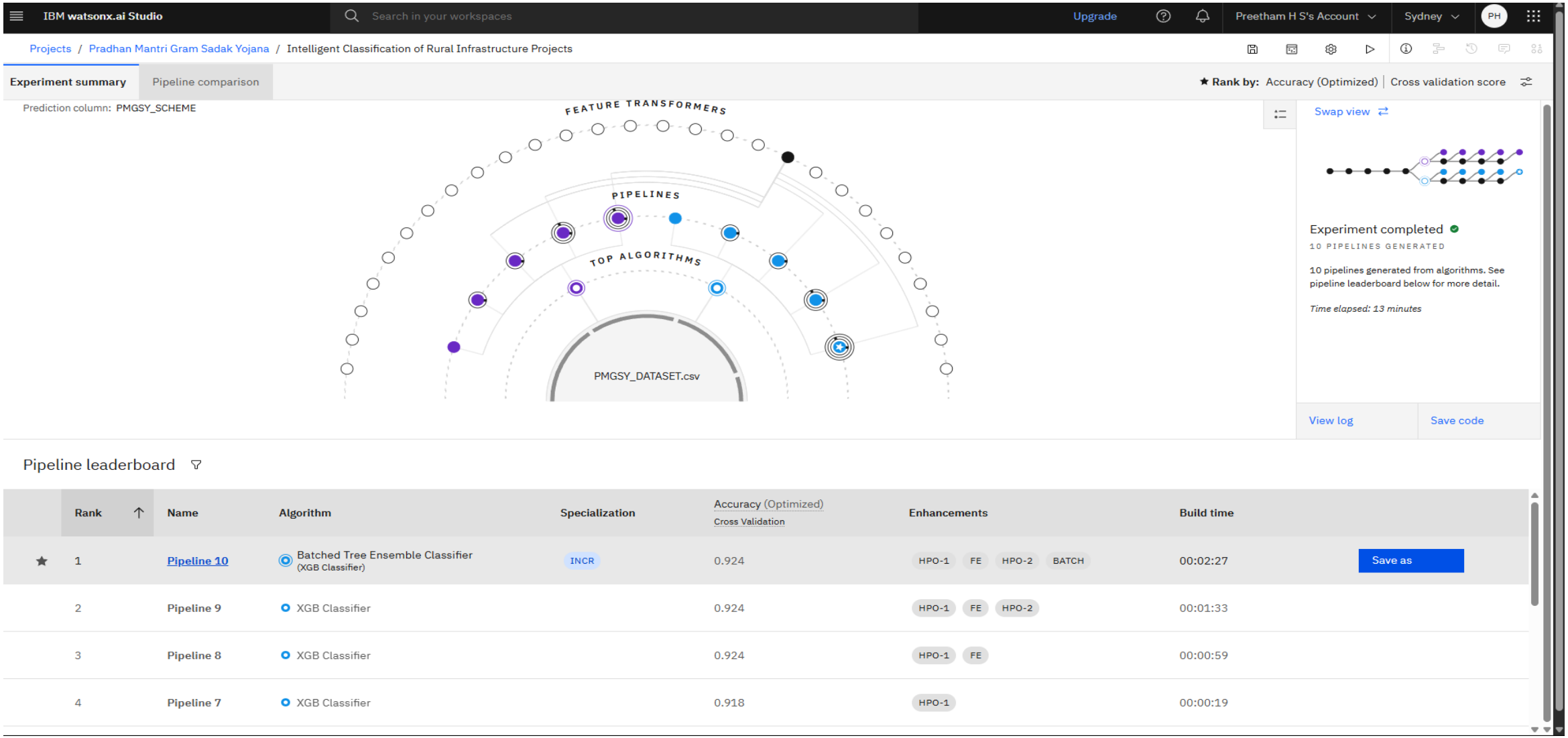
edunet foundation

RESULT



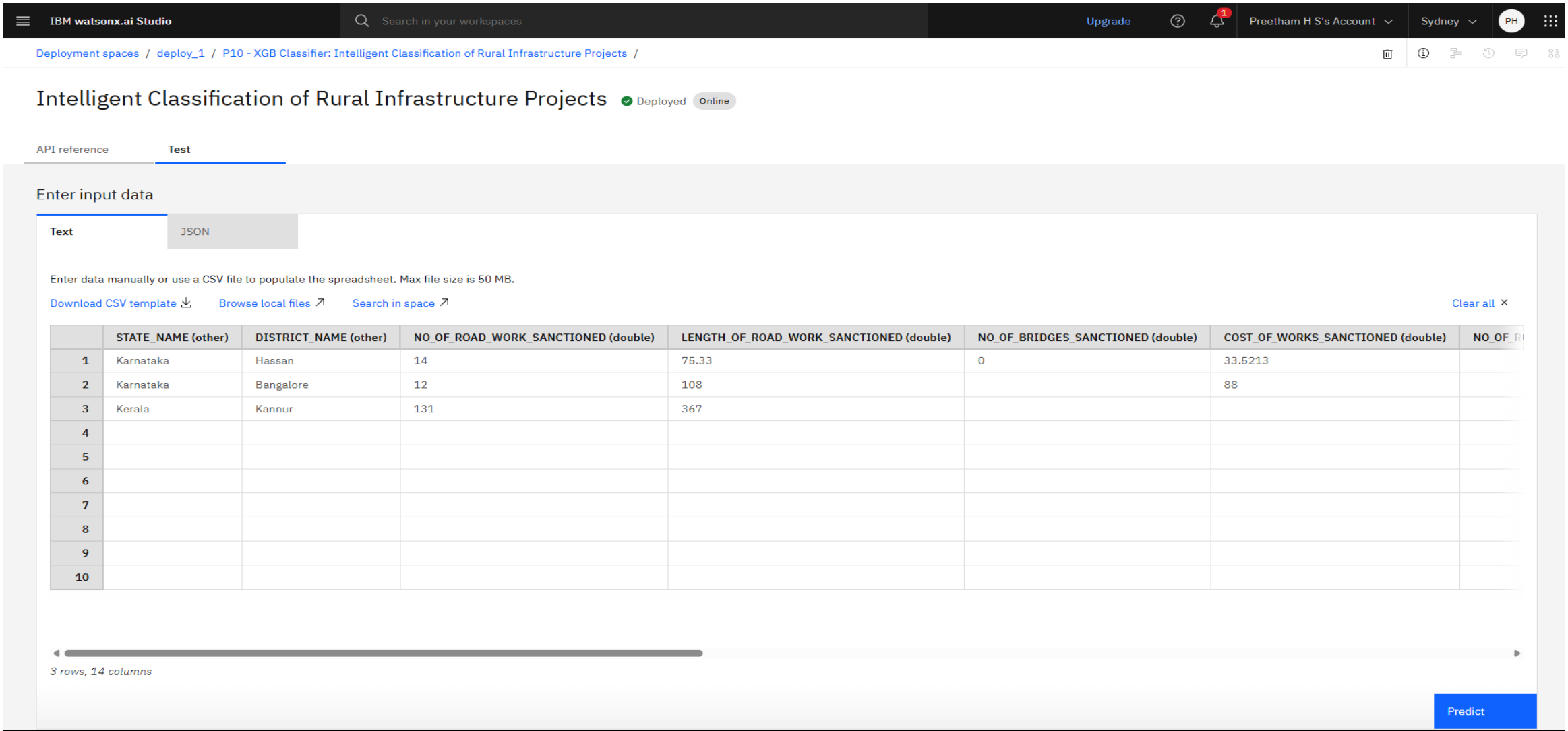
	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 10	Batched Tree Ensemble Classifier (XGB Classifier)	INCR	0.924	HPO-1 FE HPO-2 BATCH	00:02:27
	2	Pipeline 9	XGB Classifier		0.924	HPO-1 FE HPO-2	00:01:33
	3	Pipeline 8	XGB Classifier		0.924	HPO-1 FE	00:00:59
	4	Pipeline 7	XGB Classifier		0.918	HPO-1	00:00:19

RESULT



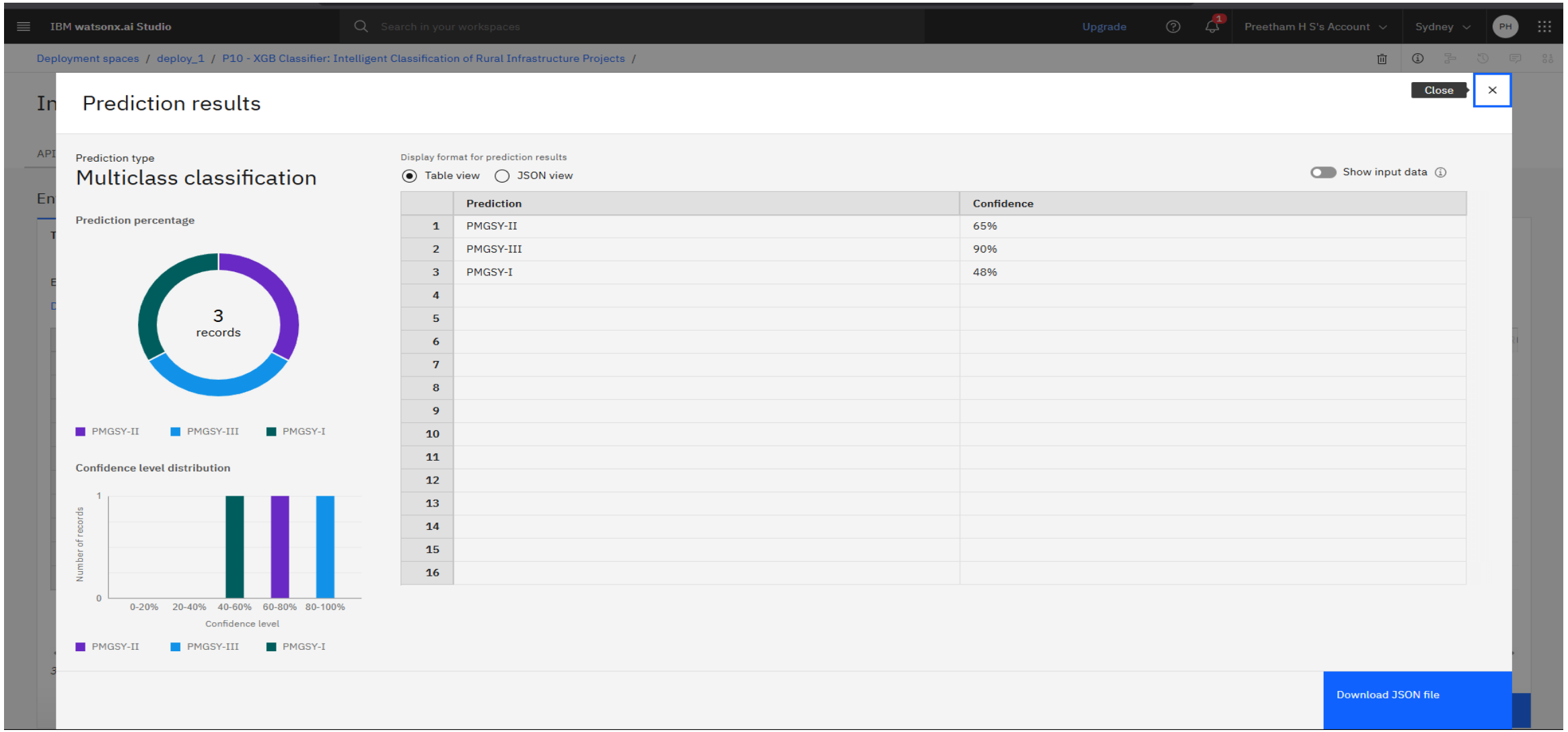
	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time	
★	1	Pipeline 10	Batched Tree Ensemble Classifier (XGB Classifier)	INCR	0.924	HPO-1 FE HPO-2 BATCH	00:02:27	Save as
	2	Pipeline 9	XGB Classifier		0.924	HPO-1 FE HPO-2	00:01:33	
	3	Pipeline 8	XGB Classifier		0.924	HPO-1 FE	00:00:59	
	4	Pipeline 7	XGB Classifier		0.918	HPO-1	00:00:19	

RESULT



	STATE_NAME (other)	DISTRICT_NAME (other)	NO_OF_ROAD_WORK_SANCTIONED (double)	LENGTH_OF_ROAD_WORK_SANCTIONED (double)	NO_OF_BRIDGES_SANCTIONED (double)	COST_OF_WORKS_SANCTIONED (double)	NO_OF_R
1	Karnataka	Hassan	14	75.33	0	33.5213	
2	Karnataka	Bangalore	12	108		88	
3	Kerala	Kannur	131	367			
4							
5							
6							
7							
8							
9							
10							

RESULT



	Prediction	Confidence
1	PMGSY-II	65%
2	PMGSY-III	90%
3	PMGSY-I	48%
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		

CONCLUSION

- This project helped us understand how AI can be used to solve real-world government problems.
- By automating the classification of schemes, we saved time and reduced the chance of human errors.
- Watsonx.ai's AutoAI made model building and deployment simple, even without coding.
- The final system is useful for rural development departments who want fast and reliable predictions.
- Overall, this project was a great learning experience in applying cloud-based AI to real data.

FUTURE SCOPE

- We can add more project details like terrain, road type, or contractor name to make predictions even better.
- The model can be made available through an API so that it connects with government portals.
- We could also train it on new project data regularly so it keeps getting smarter.
- This idea can be used for other schemes like Bharat Mala, AMRUT, or urban development projects too.
- Finally, we could design a simple front-end tool in local languages so that officers across states can easily use it.

REFERENCES

- AI Kosh PMGSY Dataset: https://aikosh.indiaai.gov.in
- IBM Watsonx.ai: https://cloud.ibm.com
- IBM AutoAI Documentation
- Government of India PMGSY Portal

THANK YOU

edunet foundation