CAPSTONE PROJECT 34

PREDICTING ELIGIBILITY FOR SCHEMES USING MACHINE LEARNING

Presented By:

Vuppu Adithya Kumar
Anurag University
B. Tech CSE(Data Science)

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result (Output Image)
- Conclusion
- Future Scope
- References

edunet foundation

PROBLEM STATEMENT

The National Social Assistance Program (NSAP) is a flagship social security and welfare program by the Government of India. It aims to provide financial assistance to the elderly, widows, and persons with disabilities belonging to below-poverty-line (BPL) households. The program consists of several sub-schemes, each with specific eligibility criteria.

Manually verifying applications and assigning the correct scheme can be a time consuming and error-prone process. Delays or incorrect allocation can prevent deserving individuals from receiving timely financial aid.

Your task is to design, build, and evaluate a multi-class classification model that can accurately predict the most appropriate NSAP scheme for an applicant based on their demographic and socio-economic data. The goal is to create a reliable tool that could assist government agencies in quickly and accurately categorizing applicants, ensuring that benefits are delivered to the right people efficiently

edunet foundation

PROPOSED SOLUTION

- The proposed ML model aims to address the challenge of predicting the most appropriate NSAP schemes based on their economic and demographic (locality) data. This involves understanding data and set it to make prepared for machine learning techniques to forecast demand patterns accurately. The solution will consist of the following components:
- Data Collection:
  - Gathered data on districts wise, caste wise, gender wise people, schemes by NASP which suits them, and other relevant factors, and took Kaggle dataset which covers all requirements of my topic for project,
- Data Preprocessing:
  - Cleaned and preprocess the collected data to handle missing values, outliers, and inconsistencies and error free for predicting outcomes.
- Machine Learning Algorithm:
  - Implement a machine learning algorithm, a random forest ML model for multiclass classification, to predict the appropriate NSAP scheme for people based on parameters in dataset, and also considering other factors to improve prediction accuracy.
- Deployment:
  - Developed a user-friendly model or website that provides real-time predictions on Schemes based on factors in dataset for each applicant with their details with a scalable and reliable platform, considering factors like district, caste for eligibility.
- Evaluation:
  - Assess the model's performance using appropriate metrics such as caste, district and some details of candidate or other relevant metrics.

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the rental bike prediction system. Here's a suggested structure for this section:

System requirements:

- IBM CLOUD (MANDATORY)
- IBM WATSONX.AI STUDIO FOR MODEL DEVELOPMENT AND DEPLOYEMENT
- IBM CLOUD OBJECT STORAGE FOR DATASET HANDLING

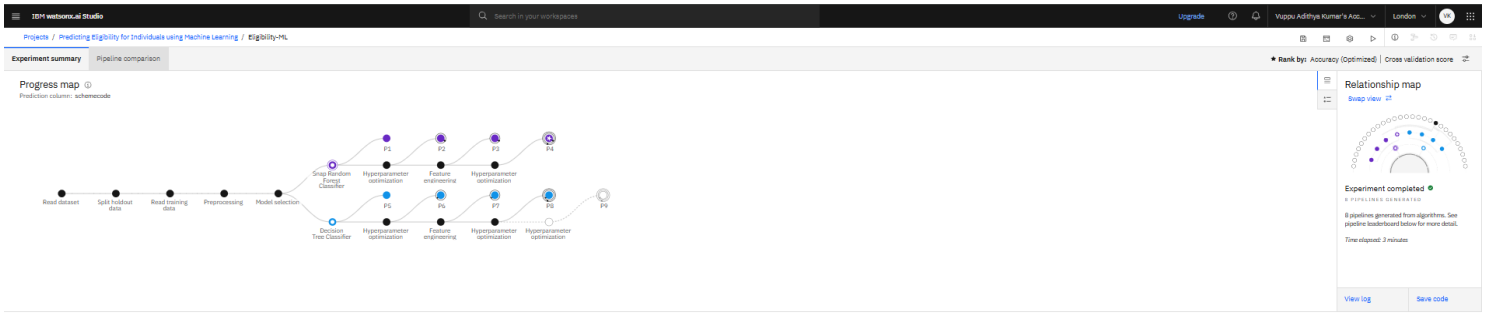
edunet foundation

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting bike counts. Here's an example structure for this section:
- **Algorithm Selection:**
- Here, I used Random Forest Classifier for ML development and for creating patterns by ML models.
- **Data Input:**
  - I have specified the input features used by the algorithm, such as historical bike rental data, like year, caste, locality and other parameters for eligibility.
- **Training Process:**
  - The algorithm is trained using historical data, creating 8 pipelines and out of that pipeline 4 is considered as highly preferred for deployment, with optimized accuracy of 0.984 in cross validation.
- **Prediction Process:**
  - The ML model successfully predicted schemes for each application based on applicant's details as mentioned earlier with >90% accuracy.

edunet foundation

RESULT

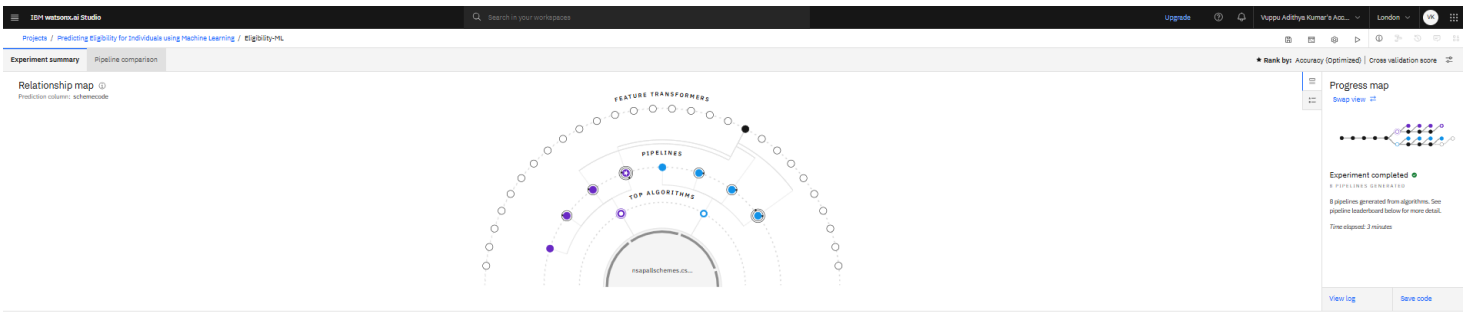


Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 4	Snap Random Forest Classifier		0.984	HPO-1 FE HPO-2	00:00:42
	2	Pipeline 3	Snap Random Forest Classifier		0.984	HPO-1 FE	00:00:32
	3	Pipeline 2	Snap Random Forest Classifier		0.981	HPO-1	00:00:09
	4	Pipeline 1	Snap Random Forest Classifier		0.981	None	00:00:04
	5	Pipeline 8	Decision Tree Classifier		0.980	HPO-1 FE HPO-2	00:00:32
	6	Pipeline 7	Decision Tree Classifier		0.980	HPO-1 FE	00:00:27
	7	Pipeline 6	Decision Tree Classifier		0.980	HPO-1	00:00:04
	8	Pipeline 5	Decision Tree Classifier		0.978	None	00:00:01

edunet foundation

RESULT

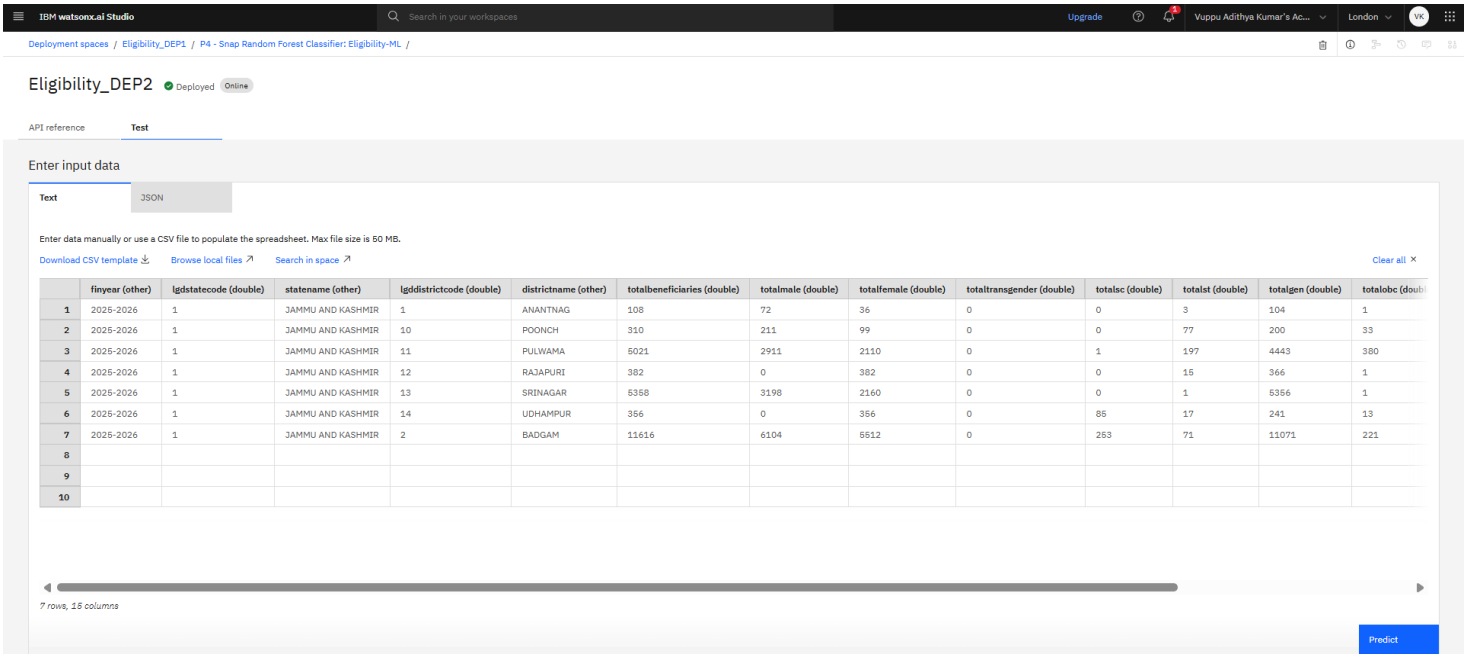


Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 4	Snap Random Forest Classifier		0.984	HPO-1 FE HPO-2	00:00:42
	2	Pipeline 3	Snap Random Forest Classifier		0.984	HPO-1 FE	00:00:32
	3	Pipeline 2	Snap Random Forest Classifier		0.981	HPO-1	00:00:09
	4	Pipeline 1	Snap Random Forest Classifier		0.981	None	00:00:04
	5	Pipeline 8	Decision Tree Classifier		0.980	HPO-1 FE HPO-2	00:00:32
	6	Pipeline 7	Decision Tree Classifier		0.980	HPO-1 FE	00:00:27
	7	Pipeline 6	Decision Tree Classifier		0.980	HPO-1	00:00:04
	8	Pipeline 5	Decision Tree Classifier		0.975	None	00:00:01

edunet foundation

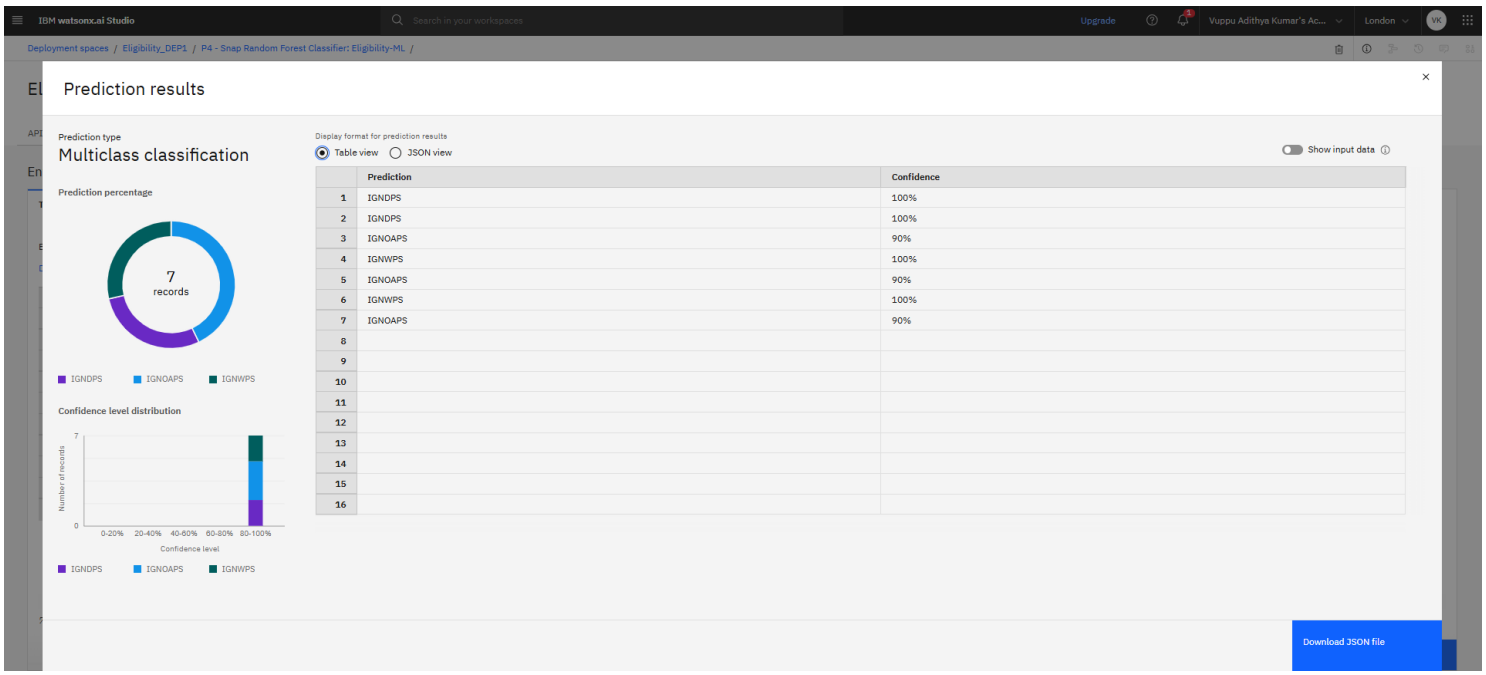
RESULT



	finyear (other)	lgdstatecode (double)	statename (other)	lgddistrictcode (double)	districtname (other)	totalbeneficiaries (double)	totalmale (double)	totalfemale (double)	totaltransgender (double)	totalsc (double)	totalst (double)	totalgen (double)	totalobc (doubl
1	2025-2026	1	JAMMU AND KASHMIR	1	ANANTNAG	108	72	36	0	0	3	104	1
2	2025-2026	1	JAMMU AND KASHMIR	10	POONCH	310	211	99	0	0	77	200	33
3	2025-2026	1	JAMMU AND KASHMIR	11	PULWAMA	5021	2911	2110	0	1	197	4443	380
4	2025-2026	1	JAMMU AND KASHMIR	12	RAJAPURI	382	0	382	0	0	15	366	1
5	2025-2026	1	JAMMU AND KASHMIR	13	SRINAGAR	5358	3198	2160	0	0	1	5356	1
6	2025-2026	1	JAMMU AND KASHMIR	14	UDHAMPUR	356	0	356	0	85	17	241	13
7	2025-2026	1	JAMMU AND KASHMIR	2	BADGAM	11616	6104	5512	0	253	71	11071	221
8													
9													
10													

edunet foundation

RESULT



	Prediction	Confidence
1	IGNDPS	100%
2	IGNDPS	100%
3	IGNOAPS	90%
4	IGNWPS	100%
5	IGNOAPS	90%
6	IGNWPS	100%
7	IGNOAPS	90%
8		
9		
10		
11		
12		
13		
14		
15		
16		

CONCLUSION

- There are many people in Kashmir who are needful for the schemes implemented by NSAP and to fulfil all those queries, using this model, things made to be easy to check each application easily and give a scheme which is based on eligibility of people on which scheme with accuracy in model.

FUTURE SCOPE

•Once validated, the model can be integrated into state-level welfare portals.
•Enables uniform and automated scheme allocation across diverse regions.
•Can be embedded into mobile or web-based application systems.
•Offers instant feedback to applicants and officials, reducing delays.
•Enhances accuracy by cross-verifying socio-economic data.
•Reduces fraudulent claims and improves targeting of genuine beneficiaries.
•Model can be retrained periodically using new application data.
•Improves prediction accuracy and adapts to changing eligibility norms.

edunet
foundation

REFERENCES

- AI Kosh dataset link –
- https://aikosh.indiaai.gov.in/web/datasets/details/district_wise_pension_data_under_the_national_social_assistance_programme_nsap_1.html
- Technology –
- IBM cloud lite services

edunet foundation

IBM CERTIFICATIONS

edunet foundation

IBM CERTIFICATIONS

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Vuppu Adithya Kumar

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 23 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

THANK YOU

edunet
foundation