

# CAPSTONE PROJECT

## PREDICTING NSAP ELIGIBILITY USING MACHINE LEARNING (PROBLEM STATEMENT - 34 )

Presented By:

1. BOLLINENI PREM SAI – KL University - CSE

# OUTLINE

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

# PROBLEM STATEMENT

The National Social Assistance Program(NSAP) is a Government of India welfare initiative aimed at providing financial assistance to vulnerable groups such as elderly, widows, and disabled individuals living below the poverty line (BPL). Eligibility is typically based on criteria such as age (e.g., 60 years or above for old age pension), BPL status, disability level, and widowhood status. The **challenge** is to predict an individual's eligibility for NSAP benefits using machine learning techniques applied to socio-economic and demographic data, such as the AI\_KOSH dataset.

# PROPOSED SOLUTION

Designing a Machine Learning-based automated system that predicts the correct applicant scheme (IGNOAPS, IGNWPS, IGNDPS) for a given citizen using personal attributes. This system enhances accuracy and efficiency in determining eligibility. :

- **Data Pre-processing:**

- Irrelevant columns were dropped and demographic features were cleaned and normalized for model training.

- **Machine Learning Algorithm:**

- IBM Watson AutoAI was used to generate and select the best Decision Tree classifier model for multi-class prediction.

- **Deployment:**

- The trained model was deployed on IBM Cloud as an online service, allowing real-time eligibility predictions via web interface.

- **Evaluation:**

- The model achieved high confidence (100%) on test predictions across multiple NSAP schemes, ensuring strong classification accuracy.

# SYSTEM APPROACH

**Platform Used:**

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

**Dataset Source**

AI-KOSH (Open Government Data Portal) – District-wise pension data under NSAP.

**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

- **Algorithm Selection:**

A Multiclass Classification approach was used. IBM Watson AutoAI automatically selected the best model pipeline, with a Decision Tree Classifier chosen based on high accuracy and interpretability.

- **Data Input:**

Input features included demographic and socio-economic fields such as gender, age, disability status, caste, marital status, and BPL status. Etc. The target variable (schemecode) classified applicants into one of several NSAP schemes.

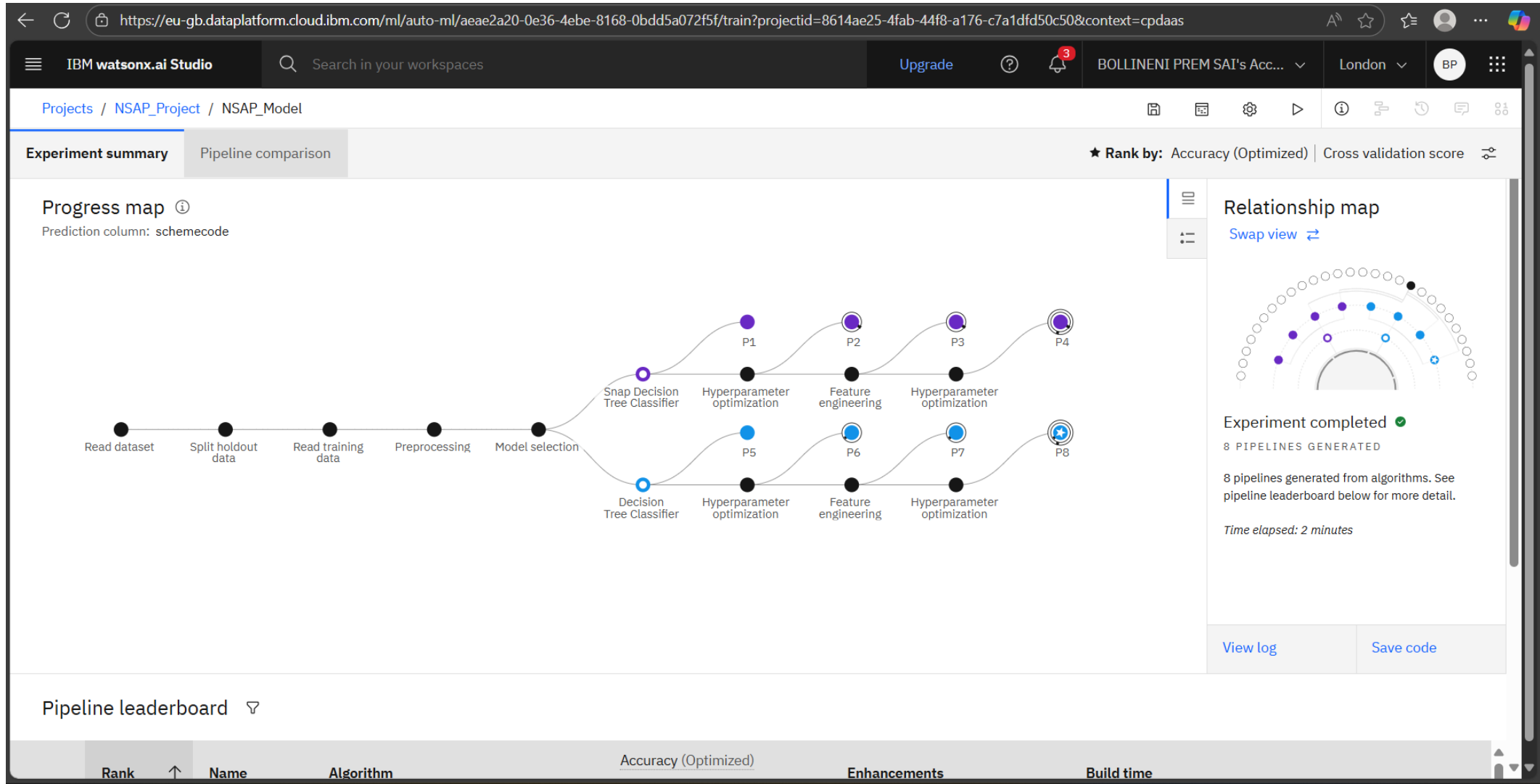
- **Training Process:**

AutoAI split the data into training and test sets, applied automated feature selection, and ran hyperparameter optimization across multiple algorithms and pipelines. The best-performing pipeline was selected for deployment.

- **Prediction Process:**

The trained Decision Tree model predicts the most suitable NSAP scheme for new applicants. The deployed model accepts input in CSV, JSON, or manual entry formats and returns predictions through a cloud-based API.

# RESULT



# RESULT

←↻🔒https://eu-gb.dataplatform.cloud.ibm.com/ml-runtime/deployments/3445b801-c309-4728-bbdf-f30b9d1cac39/test?space\_id=60590999-6597-449e-9d8e-3a3651ceca98&context=cpdaas&flus...🔍🌟🌟👤⋮🎨

IBM watsonx.ai Studio🔍Search in your workspacesUpgrade?🔔4BOLLINENI PREM SAI's Ac...LondonBP⋮

Deployment spaces / NSAP\_Deploy / P8 - Decision Tree Classifier: NSAP\_Model /

NSAP\_Deploy✔️ DeployedOnline

API referenceTest

Enter input data

TextJSON

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](#) ×

	finyear (other)	lgdstatecode (double)	statename (other)	lgddistrictcode (double)	districtname (other)	totalbeneficiaries (double)	totalmale (double)	totalfemale (double)
1	2025-2026	1	JAMMU AND KASH	1	ANANTNAG	107	71	36
2	2025-2026	1	JAMMU AND KASH	13	SRINAGAR	5357	3198	2159
3	2025-2026	1	JAMMU AND KASH	14	UDHAMPUR	356	0	356
4	2025-2026	10	BIHAR	192	BHAGALPUR	125795	61196	64589
5								

4 rows, 15 columns

Predict

BP

BOLLINENI PR...  
Profile and settings

Log out



# RESULT

## Prediction results

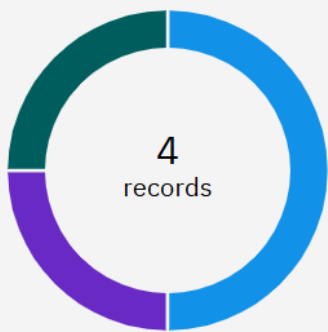
Close



Prediction type

Multiclass classification

Prediction percentage



■ IGNDPS ■ IGNOAPS ■ IGNWPS

Confidence level distribution

Display format for prediction results

☒ Table view ☐ JSON view

☐ Show input data ⓘ

	Prediction	Confidence
1	IGNDPS	100%
2	IGNOAPS	100%
3	IGNWPS	100%
4	IGNOAPS	100%
5		
6		
7		
8		
9		
10		
11		

Download JSON file

# CONCLUSION

- A machine learning-based system was successfully developed to predict the eligibility of applicants for various NSAP schemes using demographic and socio-economic data.
- Used IBM Watson service, simplified model selection and ensured high accuracy in multiclass classification.
- The deployed model can assist government agencies by automating scheme assignment, reducing manual workload, and improving the efficiency of benefit delivery.
- This solution ensures that deserving individuals receive timely financial aid by minimizing errors in scheme allocation.

# FUTURE SCOPE

- **Integration with Government Portals:**

The model can be integrated with real-time government application systems to instantly classify and suggest the appropriate NSAP scheme for applicants.

- **Expansion to More Schemes:**

The model can be extended to cover other welfare schemes beyond NSAP by retraining with additional labeled data.

- **Improved Data Collection:**

Incorporating more granular and updated socio-economic indicators (like income, health status, or location-based poverty indices) can improve accuracy.

- **Explainable AI Integration:**

Adding explainability tools (e.g., SHAP, LIME) would help justify predictions, making the system more transparent and trusted by officials.

---

# REFERENCES

- AI Kosh Dataset
- IBM Watson Studio and AutoAI

# IBM CERTIFICATIONS

In recognition of the commitment to achieve  
professional excellence



## BOLLINENI PREM SAI

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/f3e7c90a-2769-4b91-839e-e5901ef8efd6>



# IBM CERTIFICATIONS

In recognition of the commitment to achieve  
professional excellence



## BOLLINENI PREM SAI

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/e4ea900e-bd75-471b-b288-190513535236>



# IBM CERTIFICATIONS

7/24/25, 9:23 PM

Completion Certificate | SkillsBuild

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to  
**BOLLINENI PREM SAI**

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**