CAPSTONE PROJECT

NSAP ELIGIBILITY PREDICTION USING IBM WATSONX.AI

Presented By:
1. Shubham Kumar Sharma-Engineering College, Ajmer
   Computer Science Department

edunet foundation

OUTLINE

- **Problem Statement** (Should not include solution)
- **Proposed System/Solution**
- **System Development Approach** (Technology Used)
- **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.

PROPOSED SOLUTION

- **To automate the classification of applicants into the correct NSAP sub-scheme using a machine learning model trained on demographic and socio-economic data, thereby reducing manual errors and accelerating benefit delivery**
- **Data Collection:**
  - Gathered synthetic or real applicant data with features such as:
  - Age , Gender, Maritial State, Disability Status etc.
- **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 relevant schems.
- **Machine Learning Algorithm:**
  - **Explored multiple machine learning algorithm :**
    - Snap Random Forest Classifier
    - Random Forest Classifier

edunet foundation

PROPOSED SOLUTION

- **Deployment:**
- Algorithm Chosen:
  - Decision Tree Classifier (with HPO and FE)
  - Achieved highest accuracy of 0.967
  - Snap Decision Tree Classifier also performed well
- Model deployed in Watsonx.ai
- Input interface supports table & JSON
- Predictions exported in table and JSON format
- **Training Process:**
  - Used supervised learning on labeled scheme data
  - Hyperparameter optimization (HPO-1 & HPO-2) and Feature Engineering (FE) applied

SYSTEM APPROACH

- System Requirements:
  - IBM Cloud (Lite – Mandatory)
  - IBM Watsonx.ai Studio
  - IBM Cloud Object Storage
- Steps Followed:
  - Data preprocessing and labeling
  - AutoAI experiment to train multiple pipelines
  - Model selection based on cross-validated accuracy
  - Real-time prediction using Watson UI

ALGORITHM & DEPLOYMENT

- Algorithm Selection:
  - Decision Tree Classifier (with HPO and FE)
  - Achieved highest accuracy of 0.967
  - Snap Decision Tree Classifier also performed well
- Data Input:
  - Age , Gender, Marital status, Aadhar number (optional), Mobile number , Category (obc, gen , sc and at).
  - These features were selected because they directly correlate with eligibility conditions across various NSAP sub-schemes like:
    - IGNOAPS (Old Age Pension)
    - IGNWPS (Widow Pension)
    - IGNDPS (Disability Pension)

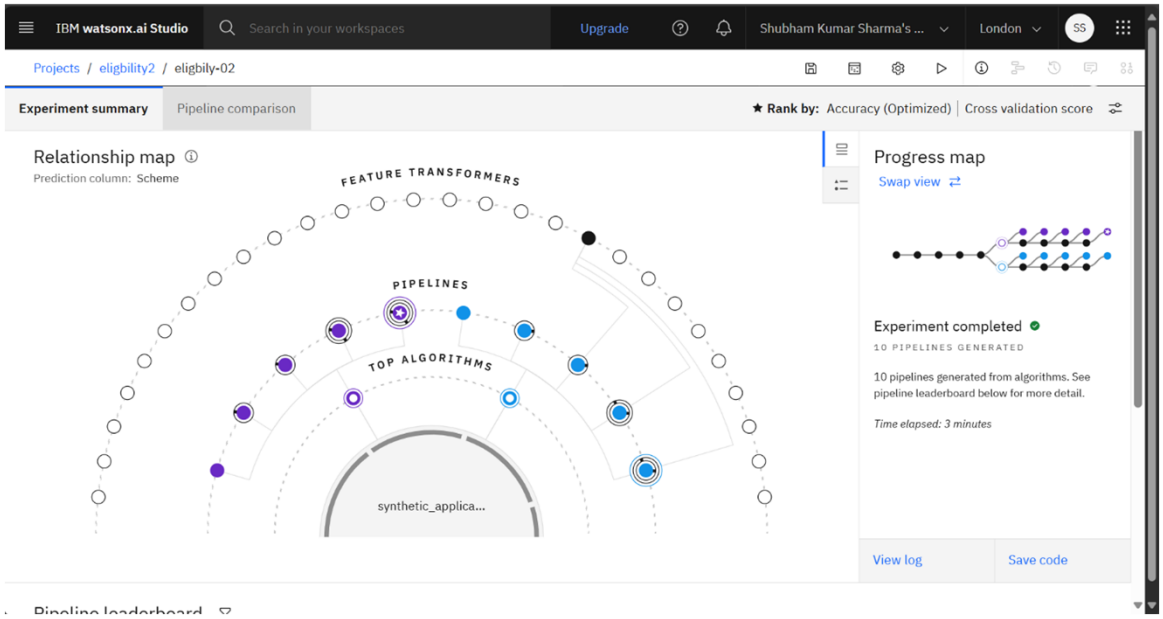
ALGORITHM & DEPLOYMENT

- Training Process :
  - **Platform Used:** IBM Watsonx.ai AutoAI
  - **Dataset Split:** AutoAI handled training/validation split internally (typically 80:20 or with K-fold cross-validation)
  - Preprocessing:
    - Categorical encoding (e.g., gender, pension type)
    - Handling of missing values
    - Automatic feature transformation by AutoAI

Model Selection: The pipeline with highest **accuracy** (96.7%) using **Decision Tree Classifier** was selected

RESULT

- 8 ML Pipelines trained using AutoAI in 3 minutes.
- Best Model: Batched Tree Ensemble Classifier (Snap Random Forest Classifier)
- Example Prediction: IGNDPS with probability [1.0, 0.0, 0.0]



edunet foundation

Prediction results

Prediction type

Multiclass classification

Prediction percentage

IGNDPS IGNWPS

Display format for prediction results

Table view JSON view

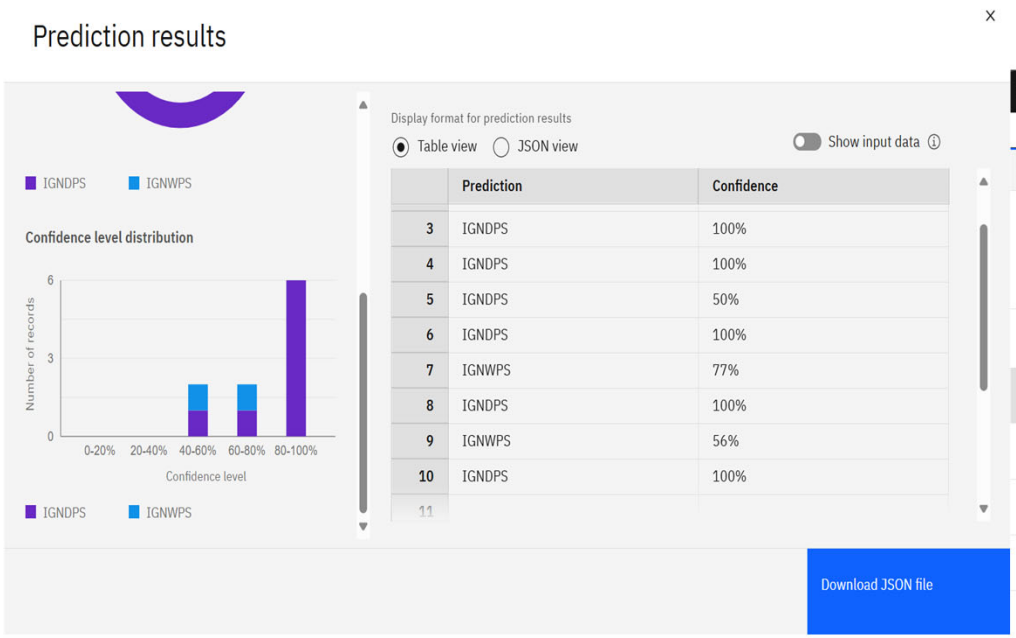
Show input data

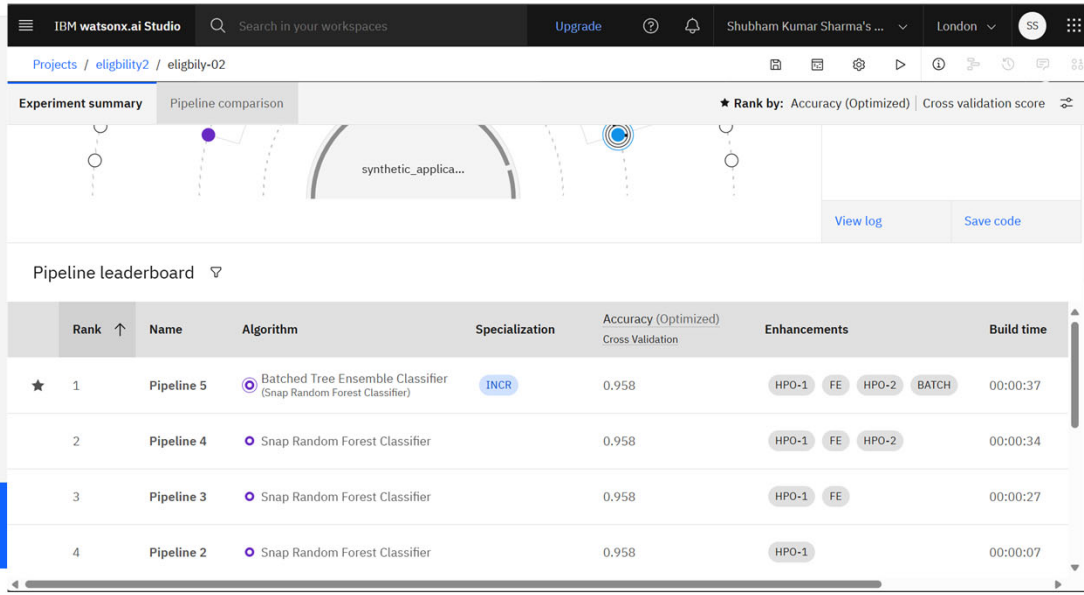
	Prediction	Confidence	Age
1	IGNDPS	100%	24
2	IGNDPS	70%	65
3	IGNDPS	100%	37
4	IGNDPS	100%	23
5	IGNDPS	50%	85
6	IGNDPS	100%	52
7	IGNWPS	77%	55
8	IGNDPS	100%	44

Download JSON file

edunet foundation

RESULT





edunet foundation

CONCLUSION

- Automated prediction improves efficiency and reduces manual errors.
- Watsonx.ai supports scalable deployment and real-time predictions.
- Model helps ensure timely scheme allocation to eligible citizens.

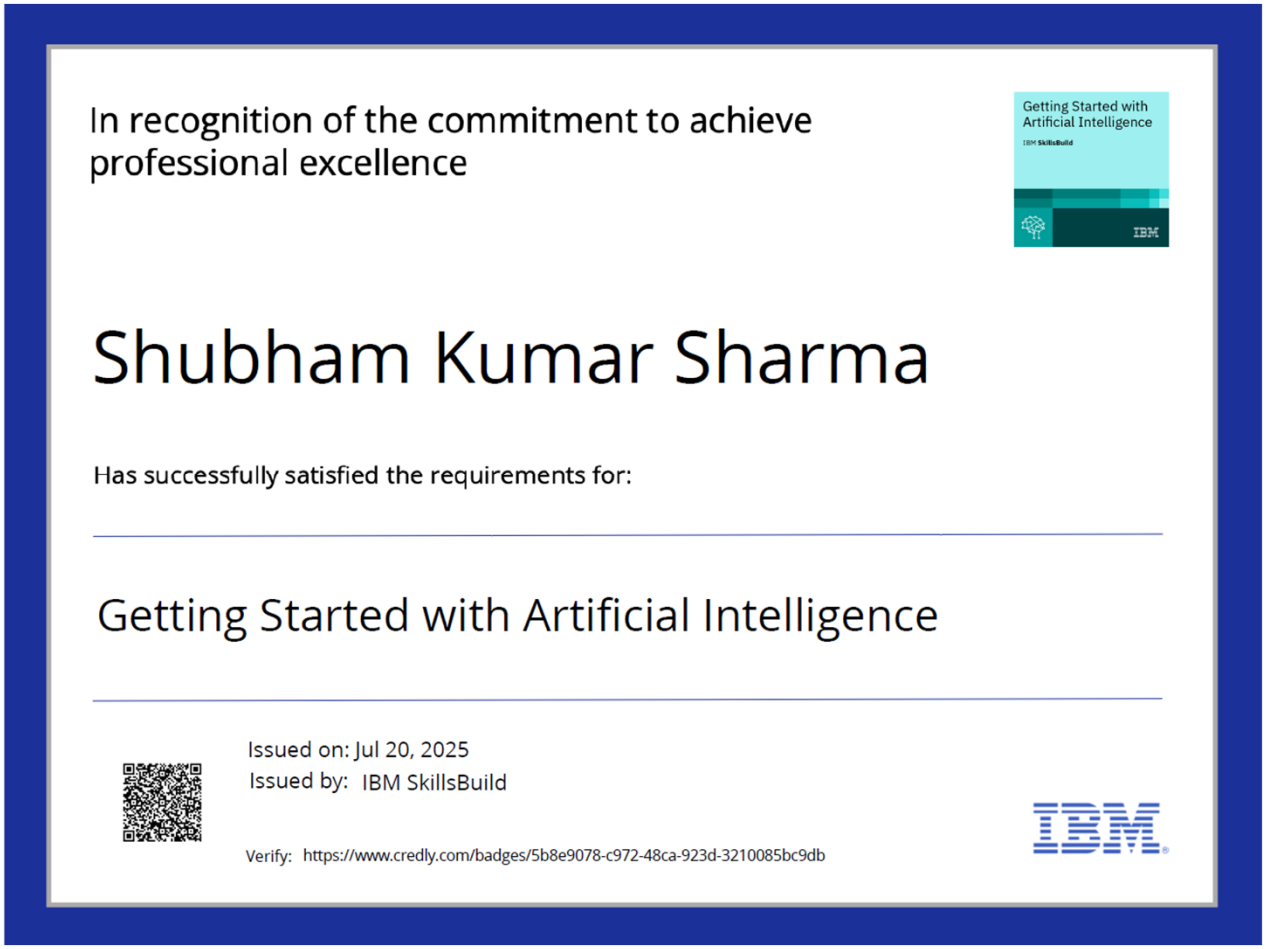
FUTURE SCOPE

- Expand to cover more NSAP schemes and complex eligibility rules.
- Hybrid ML + Rule-based logic integration.
- Aadhaar and real-time ID verification integration.
- API-based bulk processing and feedback loop for retraining.

REFERENCES

- AI Kosh Dataset: https://aikosh.indiaai.gov.in
- IBM Watson Studio & Cloud Documentation
- NSAP Guidelines – Ministry of Rural Development
- IBM AutoAI Technical Docs

IBM CERTIFICATIONS



In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild

IBM

Shubham Kumar Sharma

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/5b8e9078-c972-48ca-923d-3210085bc9db

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild

IBM

Shubham Kumar Sharma

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

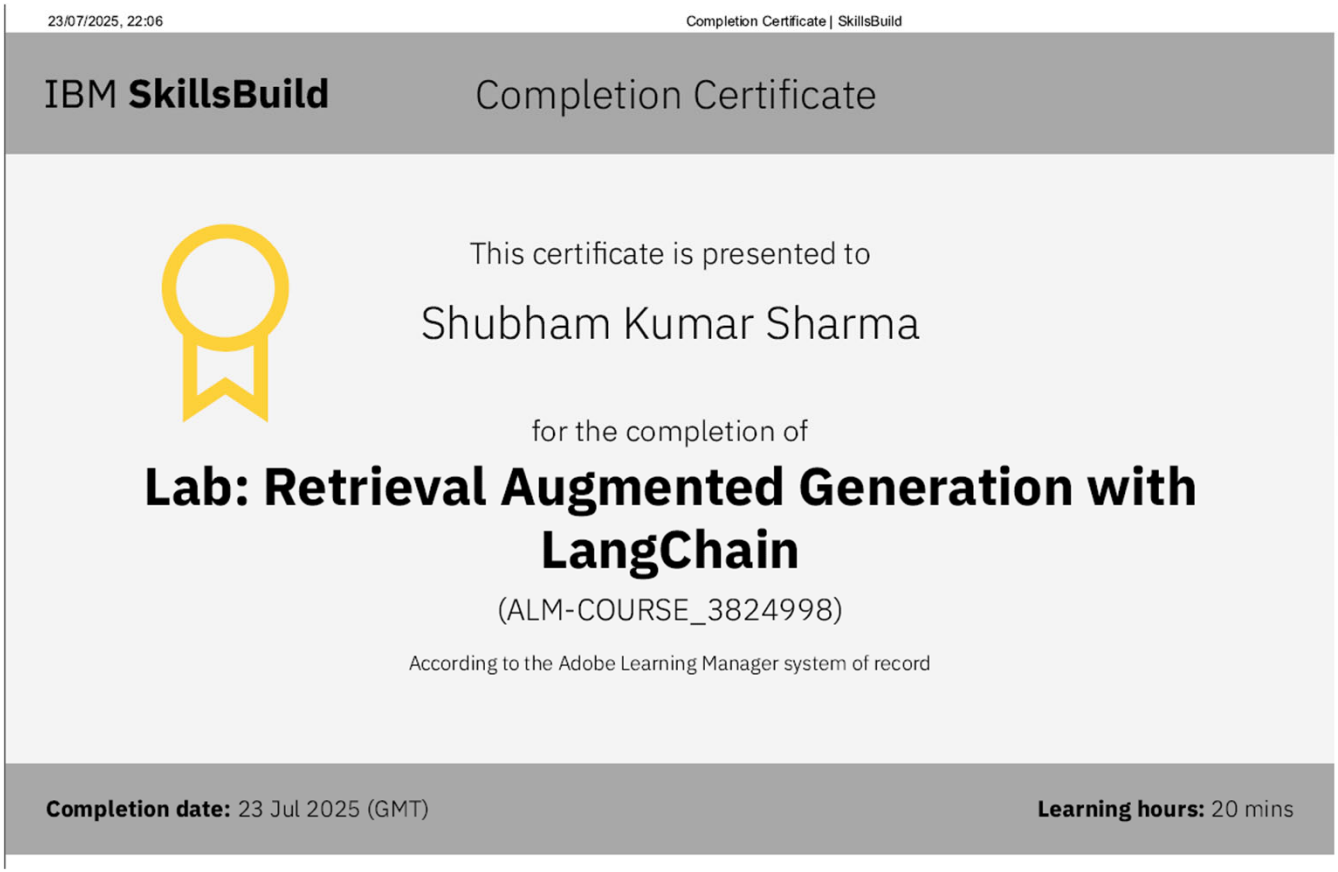
Issued on: Jul 20, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/ae2ebdcd-ac08-4f66-87af-030fcd5204f2

IBM

edunet foundation

IBM CERTIFICATIONS



23/07/2025, 22:06 Completion Certificate | SkillsBuild

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Shubham Kumar Sharma

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