# CAPSTONE PROJECT

# PREDICTING ELIGIBILITY FOR NATIONAL SOCIAL ASSISTANCE PROGRAMME (NSAP) SCHEMES USING MACHINE LEARNING

**Presented By:**
1. BOOMIKA S – Vellore Institute of Technology – Department of Computer Science and Engineering

edunet foundation

# OUTLINE

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

# PROBLEM STATEMENT

- The National Social Assistance Programme (NSAP) is a flagship welfare program by the Government of India. It provides financial support to elderly individuals, widows, and persons with disabilities from below-poverty-line (BPL) households through sub-schemes like IGNOAPS, IGNWPS, and IGNDPS.
- Currently, manual verification of applications and deciding the correct scheme is time-consuming and error-prone. This leads to delays or incorrect allocations, preventing timely support to truly deserving individuals.
- There is a need to automate the process of categorizing applicants into the most appropriate NSAP scheme based on their demographic and socio-economic data.

# PROPOSED SOLUTION

- **The project builds a multi-class classification model to automatically predict the most suitable NSAP sub-scheme (IGNOAPS, IGNWPS, IGNDPS) for applicants based on district-wise demographic and socio-economic data.**
- **Data Collection:**
  - **Use AI Kosh's district-wise NSAP pension dataset with demographic & socio-economic data.**
  - **Select key features: total beneficiaries, male/female counts, caste-wise counts, Aadhaar coverage.**
- **Data Preprocessing:**
  - **Handle missing values and encode categorical variables.**
  - **Engineer meaningful features (e.g., SC/ST percentage, gender ratio).**
- **Machine Learning Algorithm:**
  - **Build a multi-class classification model to predict schemecode.**
  - **Use IBM AutoAI to automatically train and pick the best model.**
- **Deployment:**
  - **Deploy as an online REST API on IBM Cloud Lite.**
  - **Enable real-time predictions for new applicant data.**
- **Evaluation:**
  - **Evaluate using Accuracy and F1-score.**
  - **Continuously monitor and fine-tune model performance.**
  - Result: Automated scheme prediction improves speed and consistency.

# SYSTEM APPROACH

The “System Approach” section outlines the overall strategy and methodology for developing and implementing the NSAP scheme eligibility prediction system.

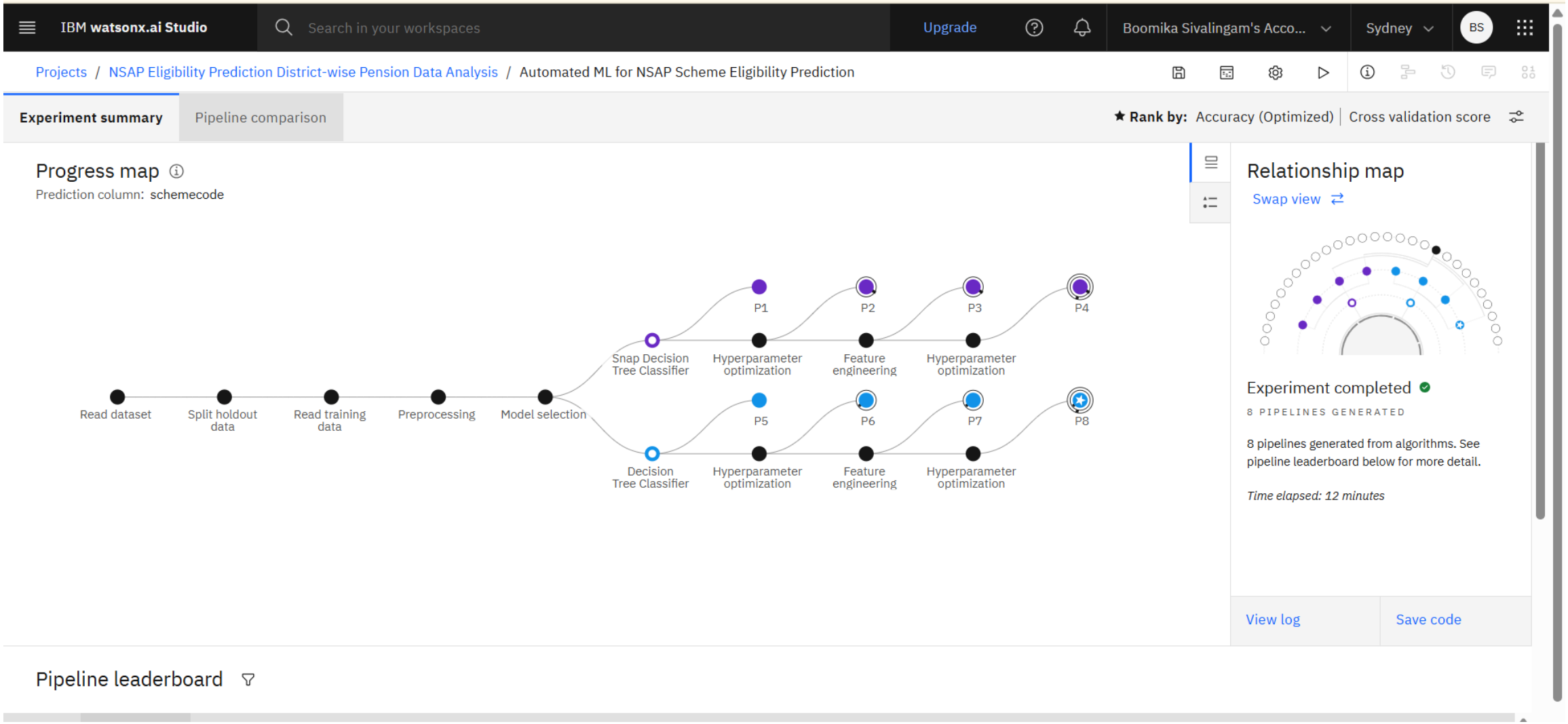
- System requirements
- IBM Cloud Lite account with Watson Machine Learning service.
- Storage: IBM Cloud Object Storage to store datasets and trained models.

- Library required to build the model
- IBM AutoAI for automated ML model building and deployment.
- Python (pandas, scikit-learn) for data preprocessing & testing.
- AI Kosh dataset: District-wise Pension Data under NSAP

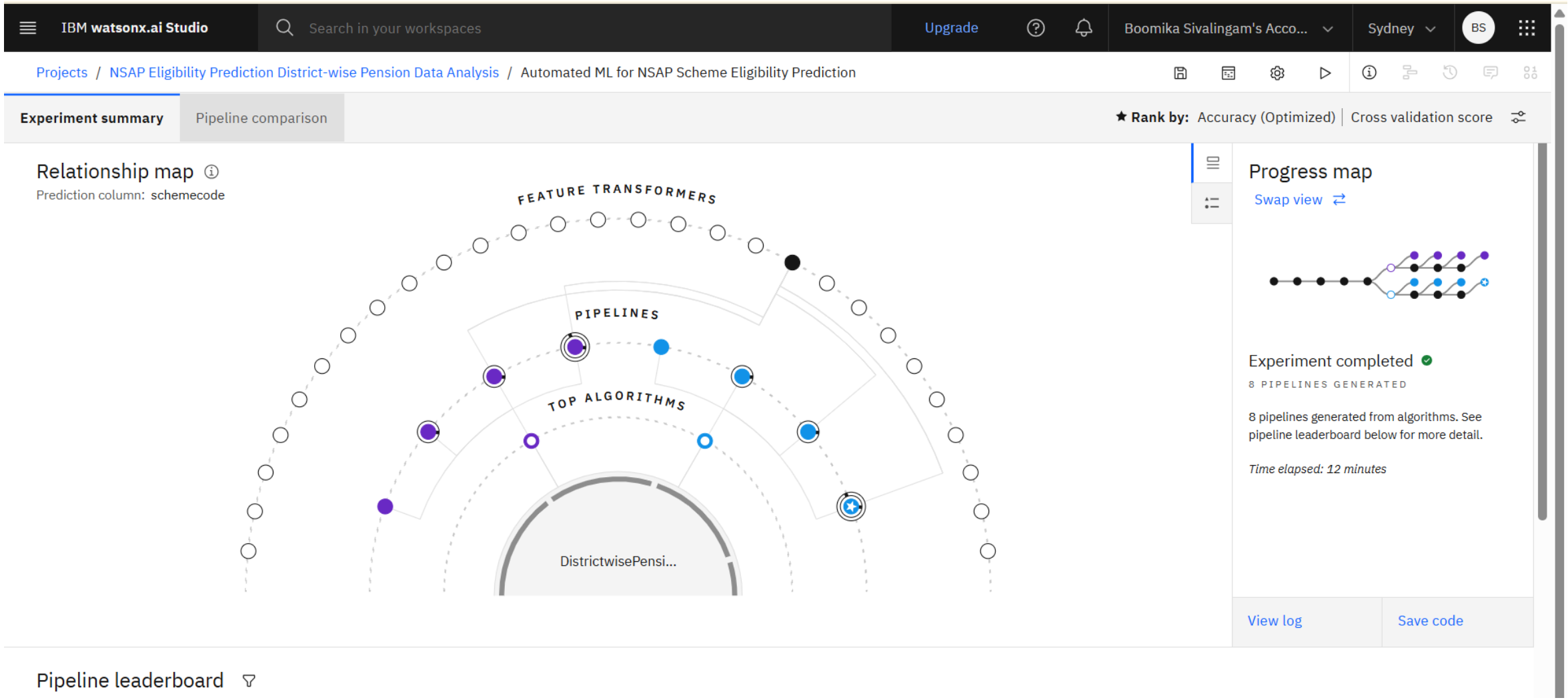
# ALGORITHM & DEPLOYMENT

- For this project, a Decision Tree Classifier was chosen as the core machine learning algorithm within IBM AutoAI.
- Algorithm Selection:
  - Chose Decision Tree Classifier via AutoAI for its interpretability and suitability for multi-class classification.
- Data Input:
  - Used district, gender counts, caste counts, Aadhaar/mobile coverage as input features.
- Training Process:
  - AutoAI trained and evaluated multiple pipelines, selecting the best based on accuracy.
- Prediction Process:
  - Deployed as on online REST API to predict schemecode for new applicant data in real-tome.

# RESULT

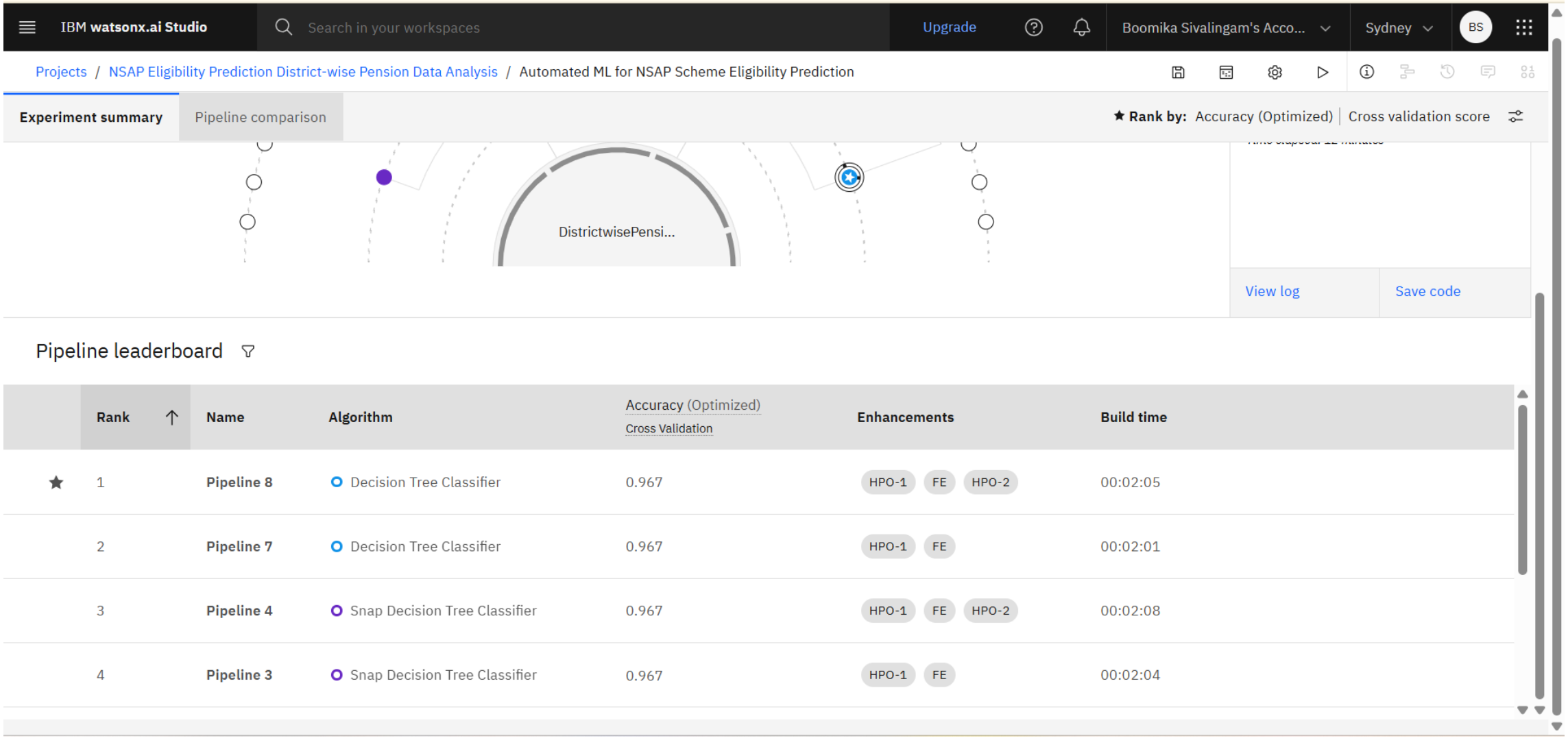


# RESULT



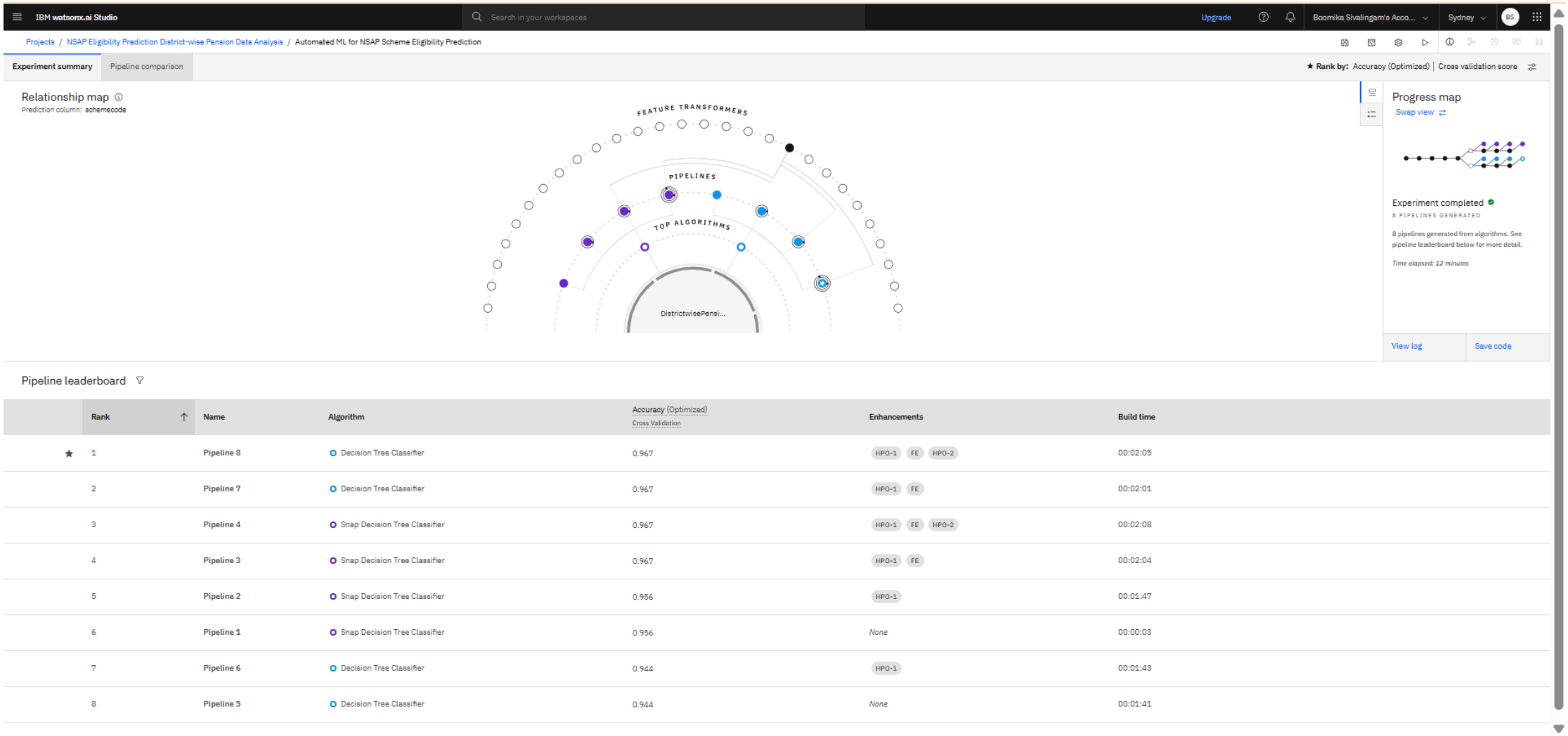

edunet foundation

# RESULT




| | Rank | Name | Algorithm | Accuracy (Optimized) Cross Validation | Enhancements | Build time |
|---|---|---|---|---|---|---|
| ★ | 1 | **Pipeline 8** | Decision Tree Classifier | 0.967 | HPO-1 FE HPO-2 | 00:02:05 |
| | 2 | **Pipeline 7** | Decision Tree Classifier | 0.967 | HPO-1 FE | 00:02:01 |
| | 3 | **Pipeline 4** | Snap Decision Tree Classifier | 0.967 | HPO-1 FE HPO-2 | 00:02:08 |
| | 4 | **Pipeline 3** | Snap Decision Tree Classifier | 0.967 | HPO-1 FE | 00:02:04 |

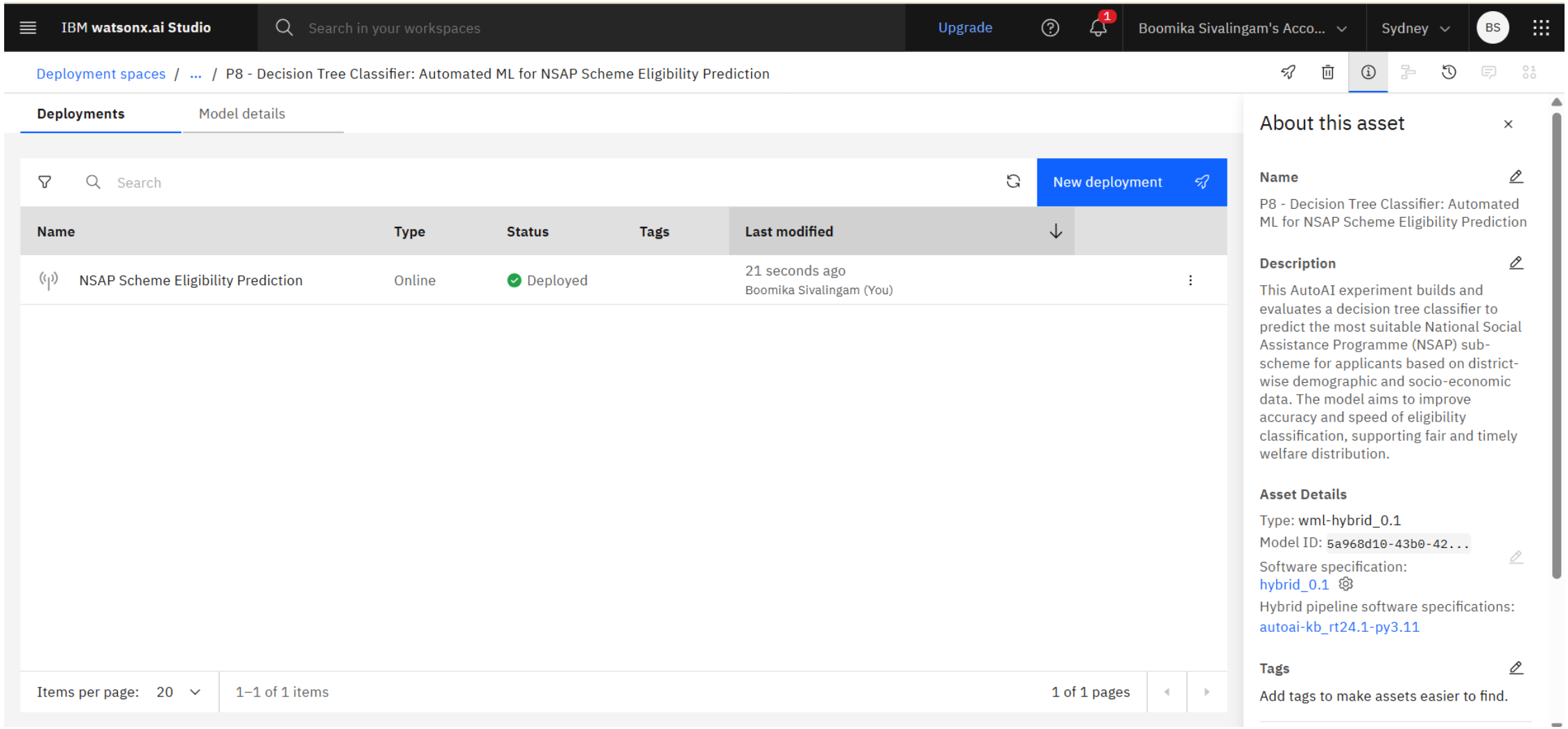
# RESULT




Pipeline leaderboard

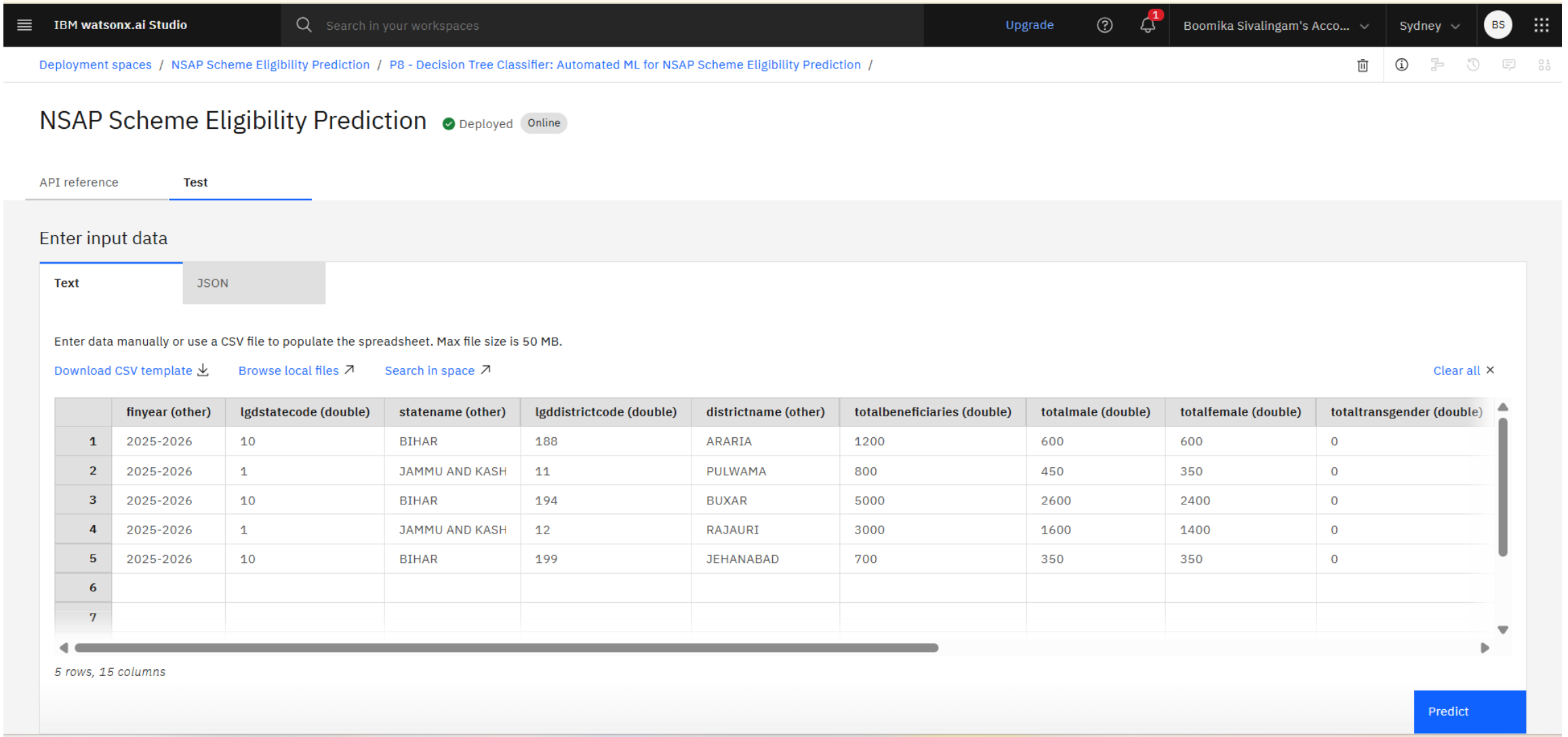
| | Rank | Name | Algorithm | Accuracy (Optimized) Cross Validation | Enhancements | Build time |
|---|---|---|---|---|---|---|
| ★ | 1 | Pipeline 8 | Decision Tree Classifier | 0.967 | HPO-1 FE HPO-2 | 00:02:05 |
| | 2 | Pipeline 7 | Decision Tree Classifier | 0.967 | HPO-1 FE | 00:02:01 |
| | 3 | Pipeline 4 | Snap Decision Tree Classifier | 0.967 | HPO-1 FE HPO-2 | 00:02:08 |
| | 4 | Pipeline 3 | Snap Decision Tree Classifier | 0.967 | HPO-1 FE | 00:02:04 |
| | 5 | Pipeline 2 | Snap Decision Tree Classifier | 0.956 | HPO-1 | 00:01:47 |
| | 6 | Pipeline 1 | Snap Decision Tree Classifier | 0.956 | None | 00:00:03 |
| | 7 | Pipeline 6 | Decision Tree Classifier | 0.944 | HPO-1 | 00:01:43 |
| | 8 | Pipeline 5 | Decision Tree Classifier | 0.944 | None | 00:01:41 |

# RESULT



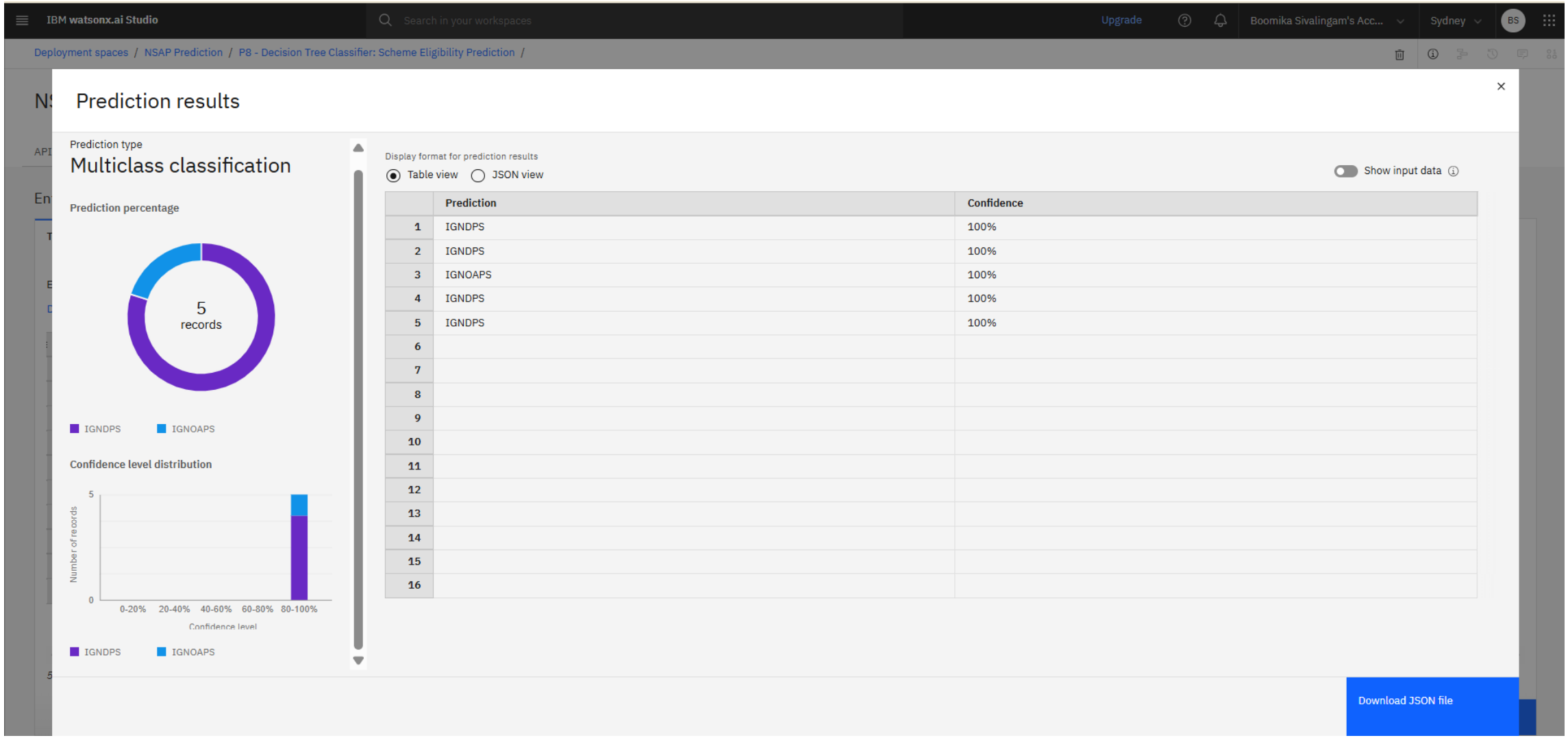

edunet foundation

# RESULT




| | finyear (other) | lgdstatecode (double) | statename (other) | lgddistrictcode (double) | districtname (other) | totalbeneficiaries (double) | totalmale (double) | totalfemale (double) | totaltransgender (double) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 2025-2026 | 10 | BIHAR | 188 | ARARIA | 1200 | 600 | 600 | 0 |
| 2 | 2025-2026 | 1 | JAMMU AND KASH | 11 | PULWAMA | 800 | 450 | 350 | 0 |
| 3 | 2025-2026 | 10 | BIHAR | 194 | BUXAR | 5000 | 2600 | 2400 | 0 |
| 4 | 2025-2026 | 1 | JAMMU AND KASH | 12 | RAJAURI | 3000 | 1600 | 1400 | 0 |
| 5 | 2025-2026 | 10 | BIHAR | 199 | JEHANABAD | 700 | 350 | 350 | 0 |
| 6 | | | | | | | | | |
| 7 | | | | | | | | | |

# RESULT




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

edunet foundation

# CONCLUSION

- The project successfully demonstrates how machine learning can automate the process of NSAP scheme eligibility prediction.
  By using demographic and district-wise data, the model reduces manual verification efforts and ensures timely benefit delivery.
- The deployed API enables real-time predictions, supporting transparent and efficient social welfare distribution.

# FUTURE SCOPE

- Incorporate applicant-level data for more personalized prediction.
- Add more features: age, income, disability percentage, etc.
- Improve model using advanced algorithms like ensemble methods or deep learning.
- Extend system to other welfare schemes beyond NSAP.

# REFERENCES


- AI Kosh dataset: District-wise Pension Data under NSAP.
- IBM Cloud documentation and AutoAI user guides.
- Machine Learning texts and online resources on multi-class classification.

# IBM CERTIFICATIONS - GETTING STARTED WITH AI


edunet foundation

# IBM CERTIFICATIONS - JOURNEY TO CLOUD


edunet foundation

# IBM CERTIFICATIONS - RAG LAB



IBM **SkillsBuild** Completion Certificate



This certificate is presented to

Boomika S

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

edunet foundation