CAPSTONE PROJECT

IMPROVED SOURCE OF DRINKING WATER

Presented By:
Spandana H A
National Institute of Engineering
Department of Information Science and Engineering

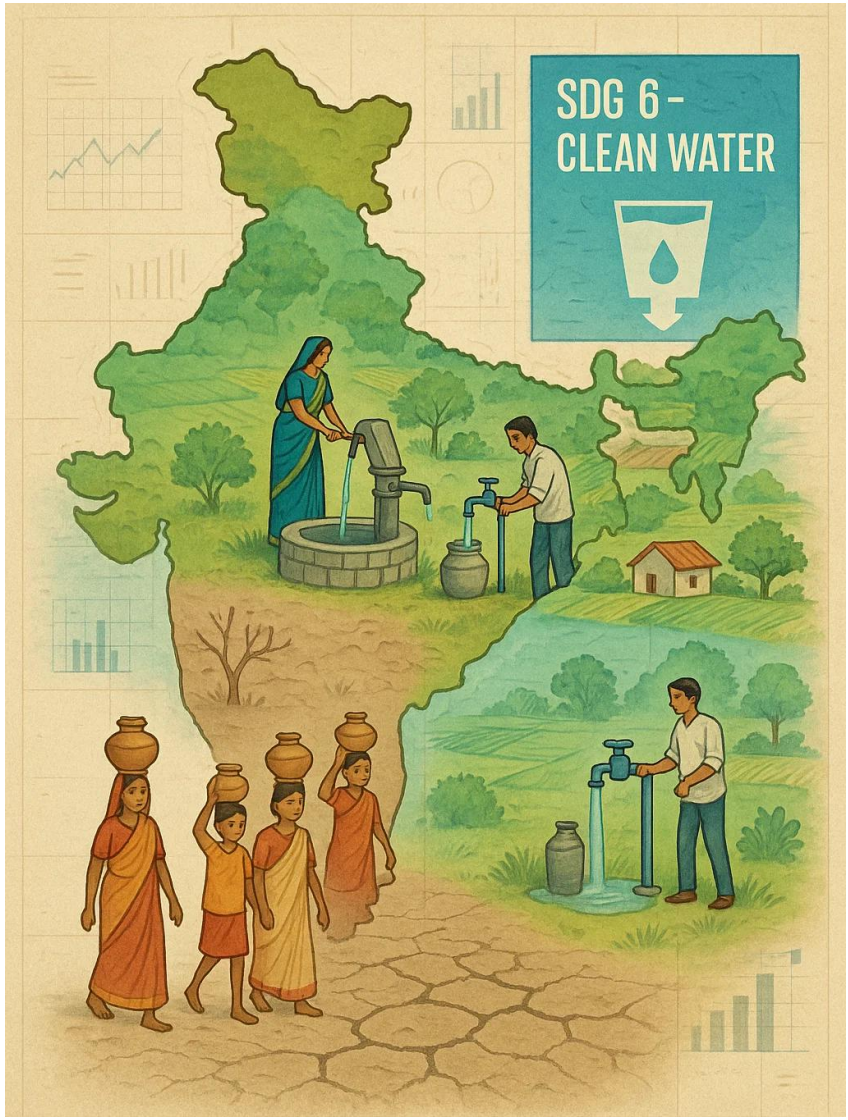
edunet foundation

OUTLINE

- Problem Statement
- Proposed System
- System Development Approach
- Algorithm & Deployment
- Result
- Conclusion
- Future Scope
- References
- Screenshots

PROBLEM STATEMENT

Access to safe and improved sources of drinking water remains a critical issue, especially in rural and underdeveloped regions of India. Despite government initiatives and efforts under Sustainable Development Goals (SDGs), significant disparities still exist across states and communities. There is a need for a data-driven approach to identify and predict areas lacking access to safe drinking water, enabling better planning and targeted interventions.



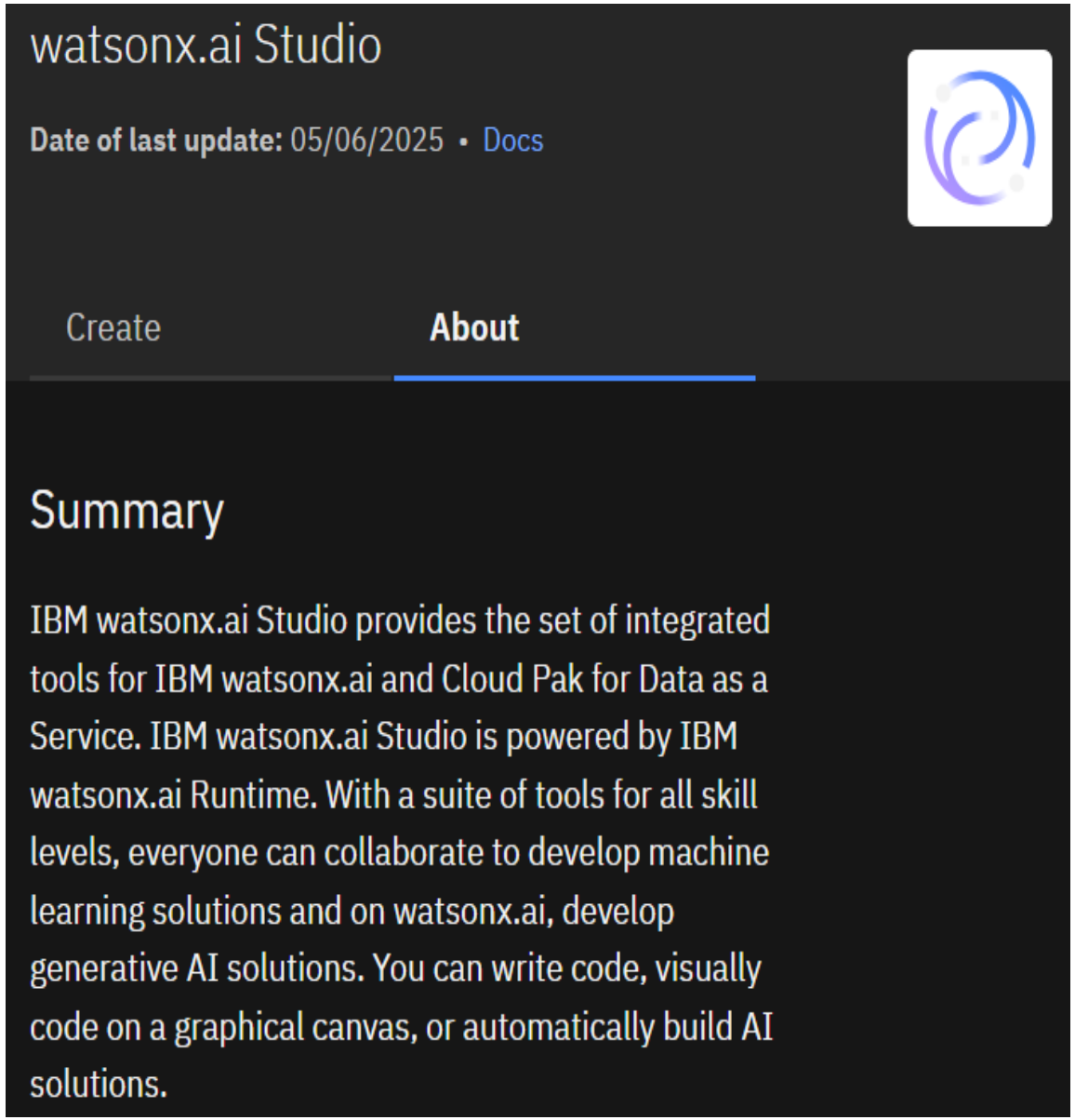
PROPOSED SOLUTION

- The system aims to **predict access to improved sources of drinking water** using data analytics and machine learning. This enables policymakers to identify areas at risk and make informed decisions.
- **Data Collection:**
  - Data from the **78th Round of NSSO survey** on household drinking water sources.
  - Includes location-wise data (state/UT), socio-economic indicators, and access levels.
- **Data Preprocessing:**
  - Handle missing values, normalize data, and perform feature selection for impactful attributes.
  - Performed feature selection and encoding to prepare the dataset for modeling.
- **Machine Learning Algorithm:**
  - Use AutoAI in IBM watsonx.ai to automatically train and select the best regression model (Snap Boosting Machine Regressor).
  - Model trained to predict the *Indicator* column — a numeric representation of water source status.
- **Deployment:**
  - Model deployed using **IBM Watsonx.ai Studio**.
  - Accessible via REST API for real-time prediction across regions.
- **Evaluation:**
  - Performance assessed using **R^2 score** and **RMSE**
  - Model showed high accuracy in predicting improved source availability across states..

SYSTEM APPROACH

- **System requirements:**
  - **Platform:** IBM Watsonx.ai
  - **Dataset:** NSSO 78th Round (Drinking Water Survey)
  - **Prediction Target: Indicator** (Improved Water Access)

- **Library required to build the model**
  - **AutoAI:** Automated model selection and training
  - **Runtime:** 8 CPU, 32 GB RAM (Watsonx.ai)



ALGORITHM & DEPLOYMENT

- **Algorithm Selection:**
  - AutoAI on **IBM Watsonx.ai** was used to automatically select and train the best machine learning model for predicting water access indicators.
- **Data Input:**
  - Features: Household size, location, caste group, religion, water source type, distance to water source
  - Target: **Indicator** (Improved Drinking Water Access)
- **Training Process:**
  - Data split into training and testing sets
  - AutoAI performed preprocessing, feature engineering, and model tuning
  - Best model selected based on performance metrics like accuracy and ROC AUC
- **Prediction Process:**
  - Model deployed in Watson Studio
  - Accepts input features and returns predicted status of improved water access
  - Can be used for real-time policy planning and analysis

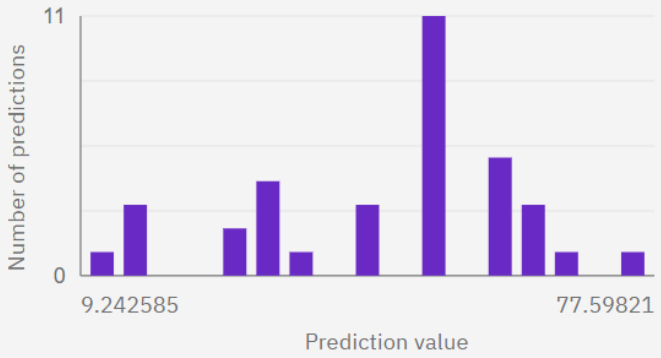
RESULT

Prediction results

Prediction type

Regression

Prediction distribution



Display format for prediction results

(•) Table view ( ) JSON view

Show input data

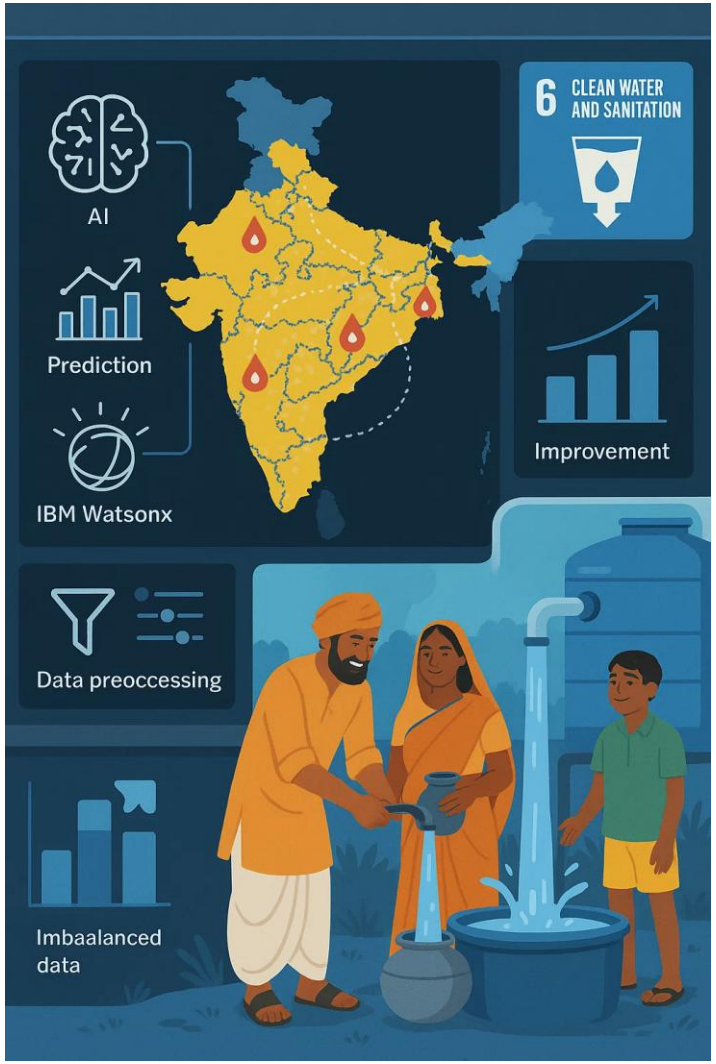
	Prediction	State	Sector	Indi
1	38.60203255531677	Lakshadweep	Rural	Perc
2	60.519622931769526	Madhya Pradesh	Urban	Perc
3	53.24522187324783	Karnataka	Rural	Perc
4	77.59821717848183	Ladakh	All	Perc
5	27.760366204490087	Manipur	Rural	Perc
6	71.57215436564029	Meghalaya	Urban	Perc
7	9.242585224809305	Meghalaya		Perc
8	29.088659866505896	Meghalaya		

- The model effectively predicts whether a household has access to improved drinking water, aiding in data-driven decision-making.

edunet foundation

CONCLUSION

- The project successfully leveraged IBM Watsonx and cloud-based AI tools to predict access to improved drinking water with high accuracy.
- The solution helps identify regions and communities lacking proper access, enabling focused planning and resource allocation.
- Challenges included data imbalance and limited real-time features, which were mitigated through preprocessing and feature selection.
- This approach proves effective in supporting data-driven decisions for achieving SDG targets related to clean water access.



FUTURE SCOPE

- Integrate real-time IoT sensor data (e.g., water quality, flow rate) for more dynamic and localized predictions.
- Expand the system to cover more states and rural regions with higher water scarcity levels.
- Apply advanced ML models like Gradient Boosting or Deep Learning for increased prediction accuracy.
- Utilize edge computing for faster decision-making in remote areas with low connectivity.
- Partner with local authorities and NGOs to drive real-world impact through targeted interventions.

REFERENCES

- **AI Kosh Dataset** – Improved Source of Drinking Water (78th Round Survey)

  https://aikosh.indiaai.gov.in/web/datasets/details/improved_source_of_drinking_water_multiple_indicator_survey_78th_round.html

- **IBM Cloud Lite Documentation** –Use of IBM Watson Studio, IBM Cloud Object Storage, and IBM Machine Learning Services – https://cloud.ibm.com

- **Python Libraries** – *Pandas, NumPy*

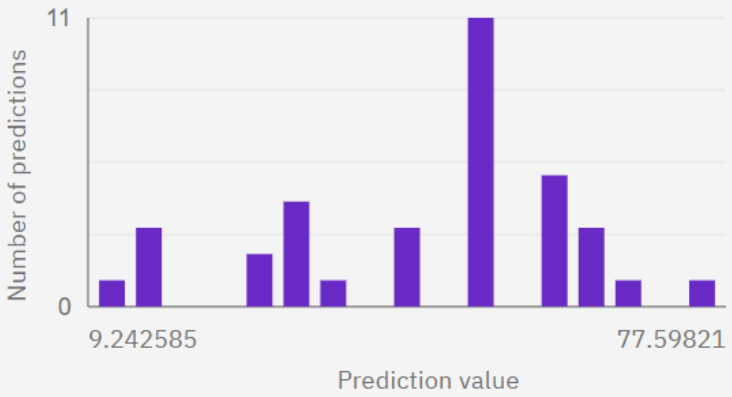
SCREENSHOTS OF PROJECT

Prediction results

Prediction type

Regression

Prediction distribution



Display format for prediction results

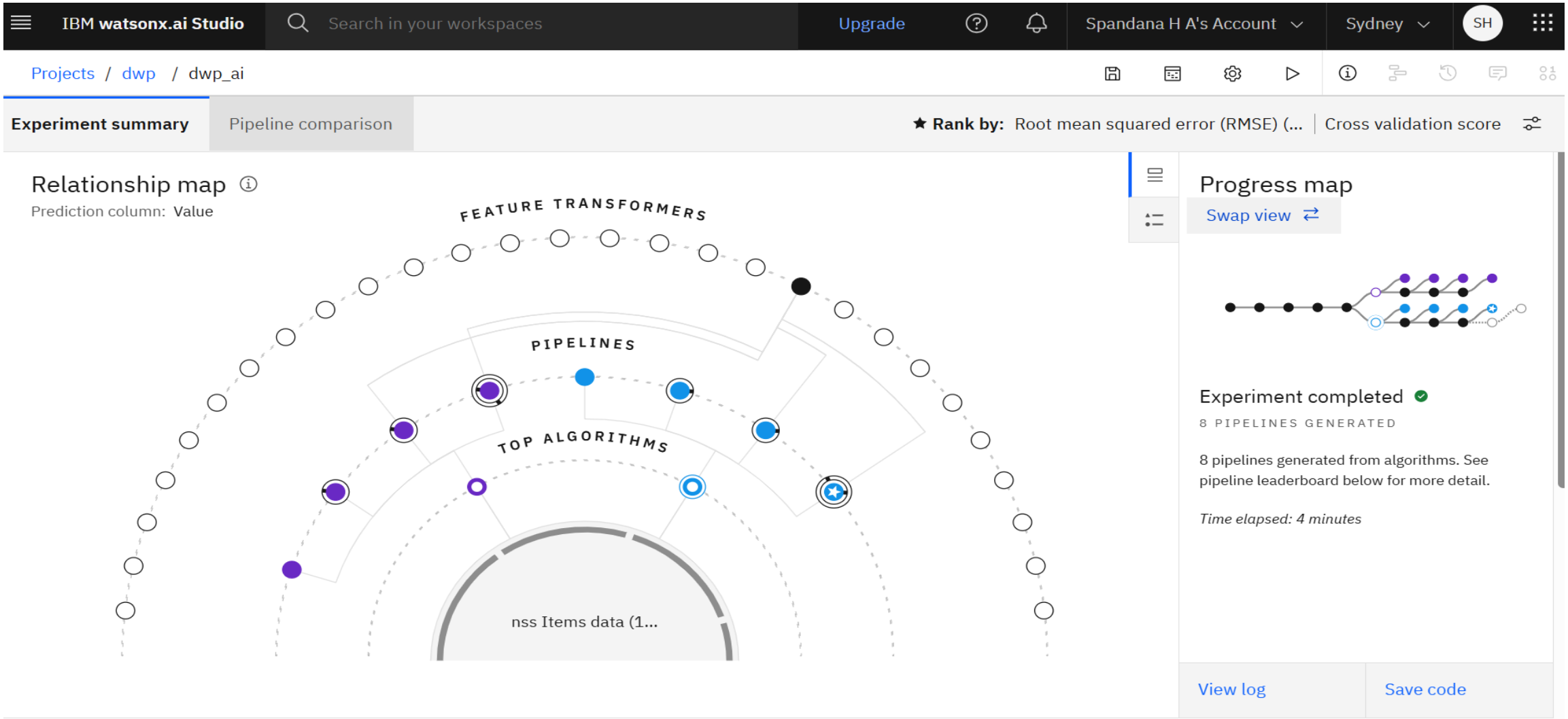
(•) Table view ( ) JSON view

Show input data

	Prediction	State	Sector	Indic
1	38.60203255531677	Lakshadweep	Rural	Perc
2	60.519622931769526	Madhya Pradesh	Urban	Perc
3	53.24522187324783	Karnataka	Rural	Perc
4	77.59821717848183	Ladakh	All	Perc
5	27.760366204490087	Manipur	Rural	Perc
6	71.57215436564029	Meghalaya	Urban	Perc
7	9.242585224809305	Meghalaya		Perc
8	29.088659866505896	Meghalaya		

edunet foundation

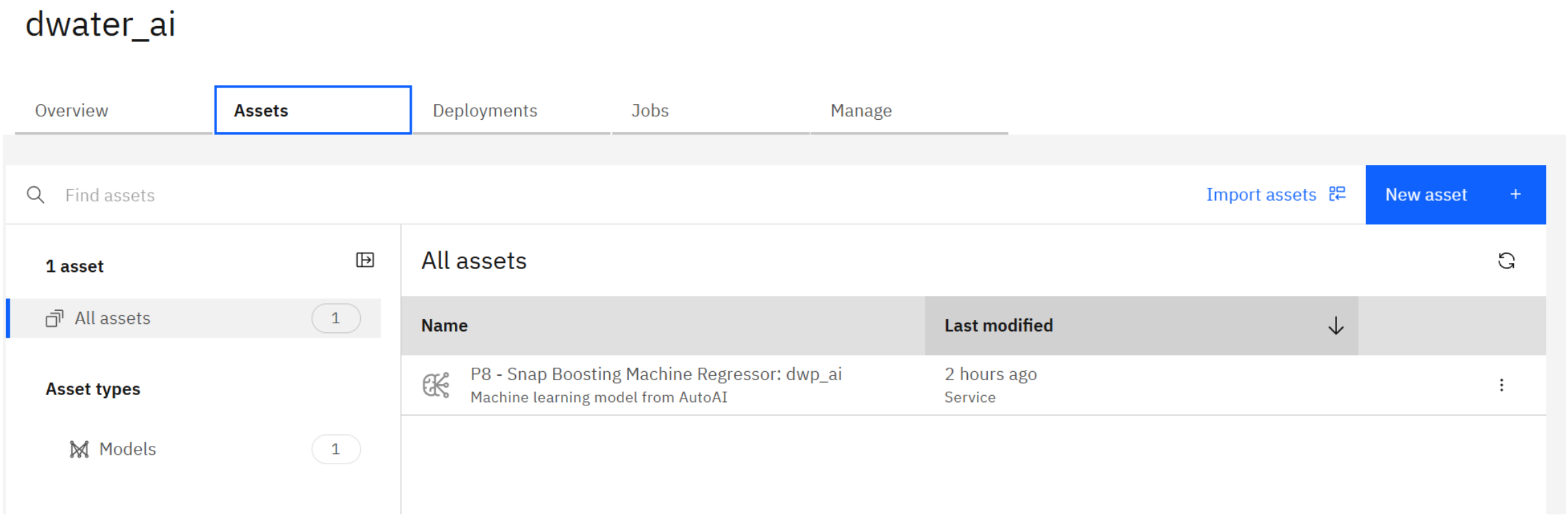
SCREENSHOTS OF PROJECT



SCREENSHOTS OF PROJECT

	A	B	C	D	E
1	State	Sector	Indicator	Sub Indicator	Value
2	All India	Rural	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	22.5
3	All India	Urban	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	58.2
4	All India	All	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	32.9
5	All India	Rural	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	56.3
6	All India	Urban	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	76.3
7	All India	All	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	62.1
8	All India	Rural	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	59.6
9	All India	Urban	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	80.2
10	All India	All	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	65.6
11	All India	Rural	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	47.7
12	All India	Urban	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	63.1
13	All India	All	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	52.2
14	All India	Rural	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	50.6
15	All India	Urban	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	66.5
16	All India	All	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	55.3
17	Andaman & Nicobar Islands	Rural	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	56.7
18	Andaman & Nicobar Islands	Urban	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	70.2
19	Andaman & Nicobar Islands	All	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	62
20	Andaman & Nicobar Islands	Rural	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	58
21	Andaman & Nicobar Islands	Urban	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	70.3
22	Andaman & Nicobar Islands	All	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	62.8
23	Andaman & Nicobar Islands	Rural	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	70.7
24	Andaman & Nicobar Islands	Urban	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	98.3
25	Andaman & Nicobar Islands	All	Percentage of Persons Reported Access to Drinking Water,	Improved Source of Drinking Water Located in the Household Premises	81.5
26	Andaman & Nicobar Islands	Rural	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	53.8
27	Andaman & Nicobar Islands	Urban	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	69.2
28	Andaman & Nicobar Islands	All	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	59.8
29	Andaman & Nicobar Islands	Rural	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	66.6
30	Andaman & Nicobar Islands	Urban	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	96.4
31	Andaman & Nicobar Islands	All	Percentage of Persons Reported Access to Drinking Water,	Exclusive Access to Improved Source of Drinking Water Located in the H	78.2
32	Andhra Pradesh	Rural	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	21.9
33	Andhra Pradesh	Urban	Percentage of Persons Reported Access to Drinking Water,	Piped Water into Dwelling or Yard/plot Which was Sufficiently Available	48

SCREENSHOTS OF PROJECT



SCREENSHOTS OF PROJECT

Deployment spaces / dwater_ai / P8 - Snap Boosting Machine Regressor: dwp_ai

Deployments | **Model details**

Input (1)

Column	Type
Indicator	other
Sector	other
State	other
Sub Indicator	other

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Spandana H A

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/c9456f92-5ca3-4d6c-b323-3d687890c716

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Spandana H A

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 17, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/c9456f92-5ca3-4d6c-b323-3d687890c716

IBM

edunet foundation

IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild

IBM

Spandana H A

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 21, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/dfb69242-bdb5-40fd-b736-9227ca8c3316

IBM

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Spandana H A

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 21 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

THANK YOU

edunet foundation