CAPSTONE PROJECT

VOICE4ALL: AUTOAI-POWERED PREDICTION OF TELE-LAW ENGAGEMENT

Presented By:
Student Name: Boddu Durga Ram Sashank
College Name: Lovely Professional University
Department: Computer Science and Engineering

edunet foundation

OUTLINE

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

PROBLEM STATEMENT

- The problem focuses on the **inequality in access to Tele-Law services** across India. While the initiative has expanded to many states and districts, there's a noticeable lack of understanding about **how different demographic groups and regions are using these services.**
- Certain **marginalized communities**—such as **Scheduled Castes (SC), Scheduled Tribes (ST), and Other Backward Classes (OBC)**—appear to be **underrepresented** in the data. Similarly, some **districts show very low levels of outreach or engagement**, which raises questions about how fairly the legal aid services are reaching the population.
- Another complication is that the **number of CSCs (Common Service Centres)** varies from region to region, making it difficult to make direct comparisons between areas. This variation creates complexity in assessing whether legal aid is being distributed equitably across all districts and communities.

PROPOSED SOLUTION

In this project, a machine learning model is developed to **predict the total number of Tele-Law case registrations** (Total) for each district. The prediction is based on input features such as:

- **Number of CSCs** in the district
- **Gender-wise data** (Female, Male)
- **Caste-wise data** (SC, ST, OBC, General)
- **Geographic identifiers** (State/UT, District)

The target variable is the **Total case registrations**, which represents the volume of legal aid usage in each region. The model learns how demographic and infrastructural factors influence service utilization, allowing for forecasting of future case registration volumes and identifying patterns of underutilization.

SYSTEM APPROACH

This section outlines the overall strategy, tools, and methodology used to develop and deploy a machine learning model that predicts the **total number of Tele-Law case registrations** across districts using IBM Watsonx.ai's AutoAI.

System Requirements:

Component	Specification
Device	Laptop/Desktop with minimum 4GB RAM (8GB+ recommended for smoother operation)
Internet Speed	Stable internet connection (minimum 10 Mbps recommended)
Browser	Google Chrome / Mozilla Firefox (latest version recommended)
Cloud Environment	IBM Cloud Academic Portal
Account Access	IBM Cloud account with access to Watsonx.ai and Cloud Object Storage
Platform	IBM Watsonx.ai (AutoAI)
Runtime	Watsonx.ai Runtime service (provisioned for model building and deployment)
Deployment	AutoAI-generated model deployed in a Watsonx.ai Deployment Space
Dataset Format	CSV (Tele-Law Case Registrations.csv)

SYSTEM APPROACH

Although IBM AutoAI handles model creation without manual coding, the backend utilizes several essential libraries and frameworks for data processing, model training, and evaluation:

Libraries Required:

Library/Tool	Purpose
pandas	Data manipulation and analysis
numpy	Numerical computations
scikit-learn	Machine learning models, preprocessing, pipeline creation
xgboost / lightgbm	Advanced tree-based boosting algorithms (AutoAI uses automatically)
watson-machine-learning-client	Interacting with deployed models on IBM Cloud

Note: All of these are handled automatically within the AutoAI pipeline; no manual installation or coding is required unless additional customization is needed outside the AutoAI interface.

ALGORITHM & DEPLOYMENT

ALGORITHM OVERVIEW

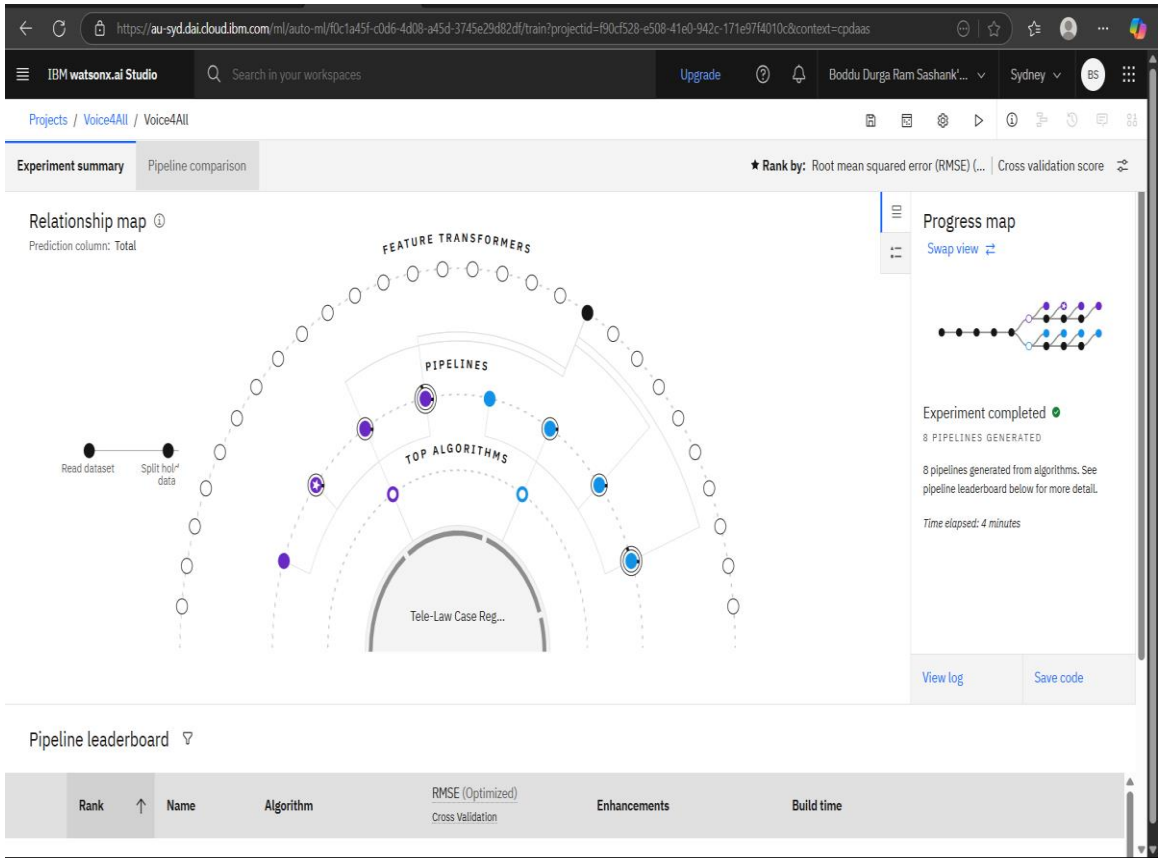
- This project uses **IBM Watsonx.ai AutoAI**, which automates model training and selection. It performs:
- **Automatic preprocessing**: Handles missing values, data types, and feature engineering.
- **Model generation**: Builds multiple pipelines using algorithms like:
  - Random Forest Regressor
  - Gradient Boosting (XGBoost/LightGBM)
  - Decision Tree, Linear, and Ridge/Lasso Regression
- **Evaluation**: Ranks models using metrics like **R^2**, **MAE**, and **RMSE**.
- **Selection**: Chooses the best-performing pipeline for deployment.

ALGORITHM & DEPLOYMENT

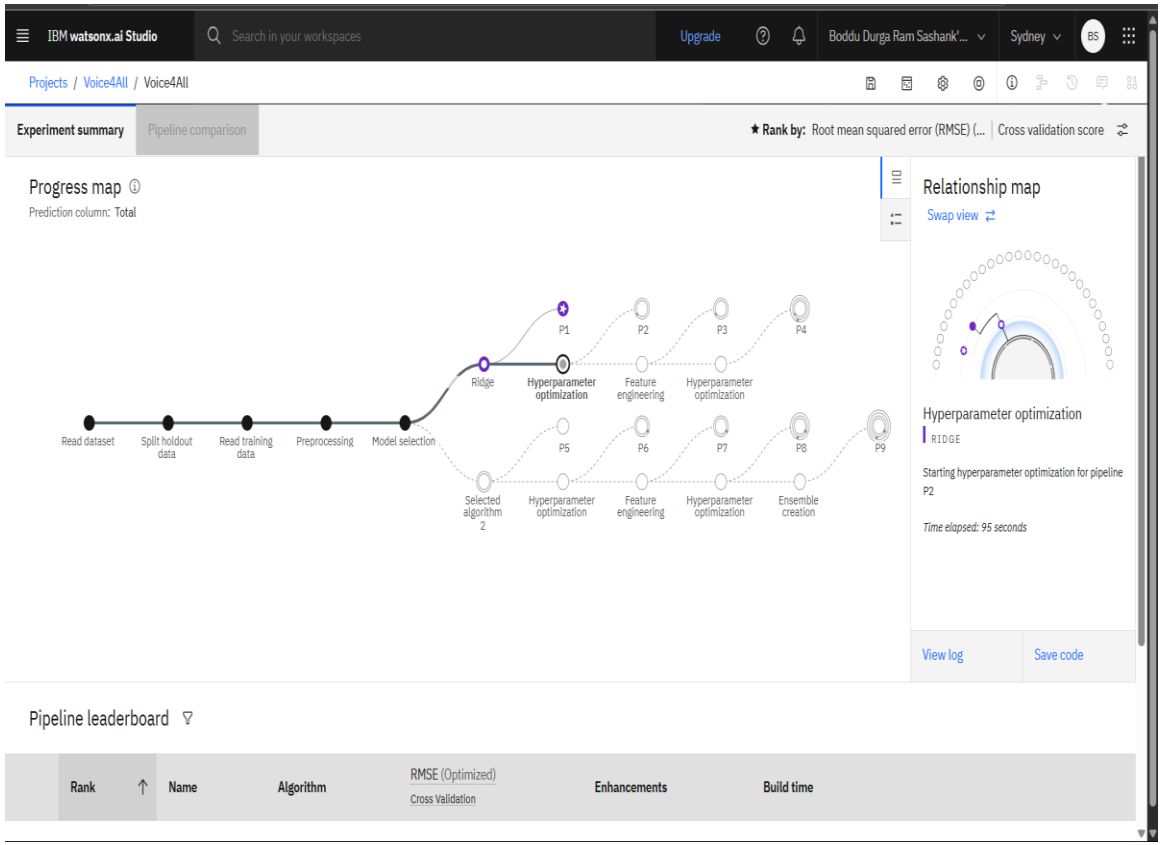
Model Deployment

- After identifying the best-performing pipeline, the model is **deployed to the cloud** using IBM Watsonx.ai's integrated tools:
- **Model Promotion**:
  - The selected model is **promoted to a deployment space** within Watsonx.ai.
- **Deployment Creation**:
  - A **new deployment** is created (real-time or batch) to make the model accessible via API.
  - The deployment is named and version-controlled.
- **Model Testing**:
  - The deployed model is tested using **new input values** directly from the cloud UI.
  - The predicted output (i.e., **total case registrations**) is generated in real time.

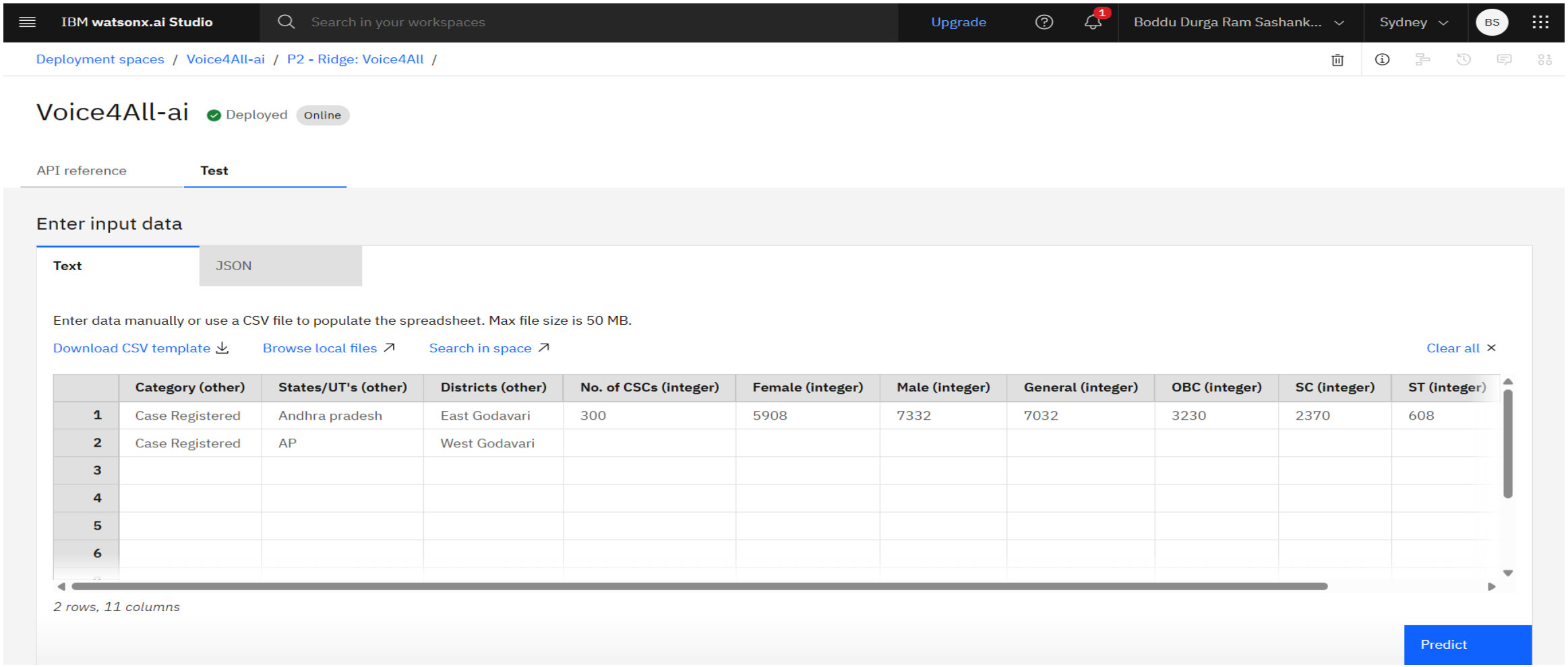
RELATIONSHIP MAP



PROGRESS MAP



RESULT

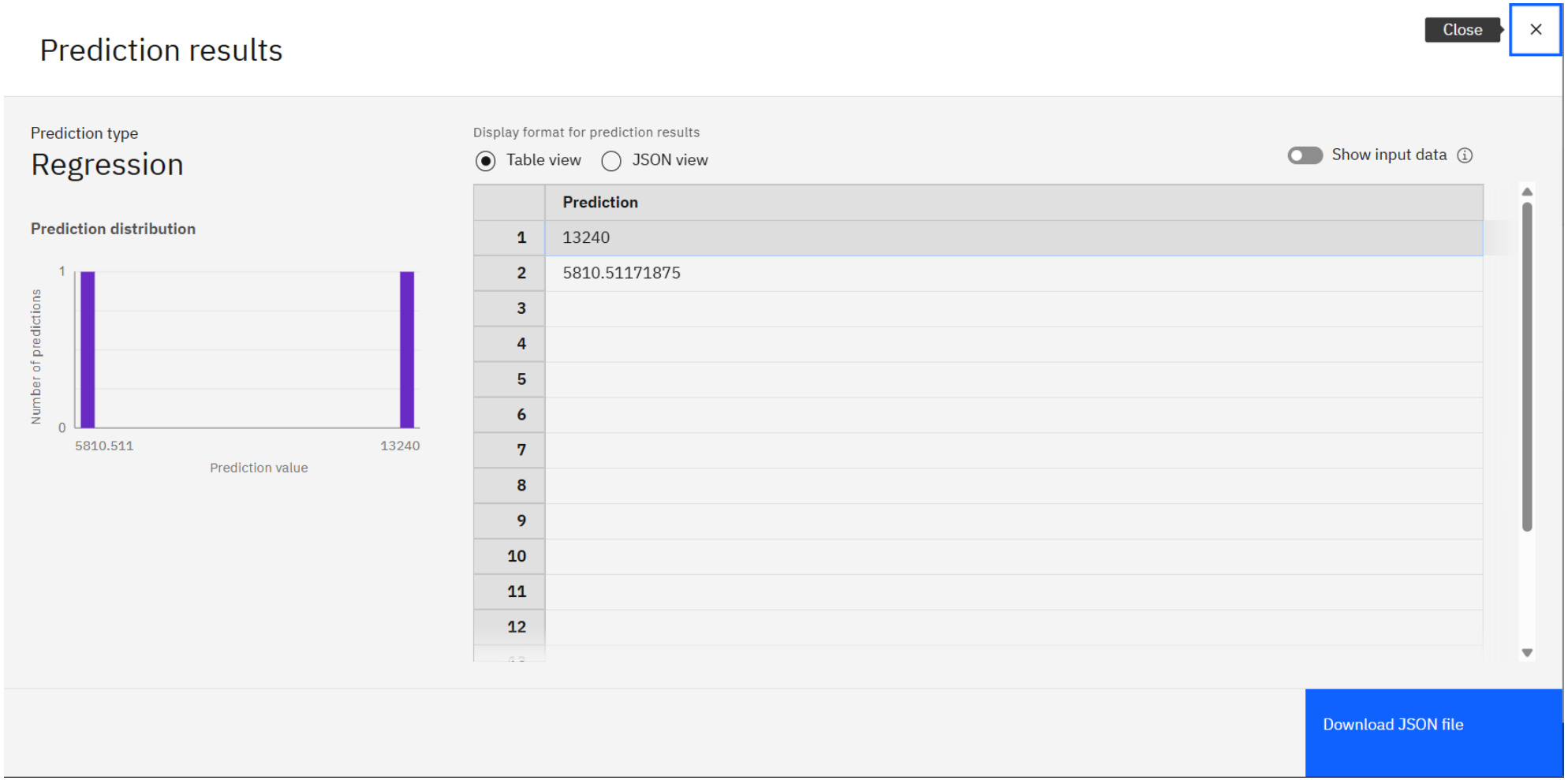


	Category (other)	States/UT's (other)	Districts (other)	No. of CSCs (integer)	Female (integer)	Male (integer)	General (integer)	OBC (integer)	SC (integer)	ST (integer)
1	Case Registered	Andhra pradesh	East Godavari	300	5908	7332	7032	3230	2370	608
2	Case Registered	AP	West Godavari							
3										
4										
5										
6										

INPUT DATA

edunet foundation

RESULT



PREDICTION RESULT

RESULT

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1567	Advice Ena	West Beng	North 24 P	179	1424	1703	3127	636	251	2059	181	3127							
1568	Advice Ena	West Beng	Paschim B	62	28	75	103	61	28	11	3	103							
1569	Advice Ena	West Beng	PurbaBard	215	643	1303	1946	1117	249	483	97	1946							
1570	Advice Ena	West Beng	Purulia	170	10154	11648	21802	3625	10253	6567	1357	21802							
1571	Advice Ena	West Beng	South 24 P	292	3259	7435	10694	5703	1653	3154	184	10694							
1572	Advice Ena	West Beng	West Medi	211	2152	3919	6071	1384	719	1714	2254	6071							
1573	Total Advic	Total (36 St	Tota (785 D	250000	3963499	6183286	10146785	2352749	3213067	3214657	1366312	10146785							
1574																			
1575	Case Regis	Andhra Pra	East Godav	300	5908	7332	13240	7032	3230	2370	608	13240							
1576	(Ctrl)																		
1577																			
1578																			
1579																			
1580																			
1581																			
1582																			
1583																			
1584																			
1585																			
1586																			
1587																			
1588																			
1589																			
1590																			
1591																			
1592																			
1593																			
1594																			

MODEL PREDICTION ACCURACY: COMPARISON WITH ACTUAL DATA

CONCLUSION

- This project effectively used **IBM Watsonx.ai's AutoAI** to develop a predictive model for estimating **Tele-Law case registrations** across Indian districts. By leveraging demographic inputs like **gender, caste**, and **CSC availability**, the model uncovers key utilization patterns and highlights regions with possible disparities in legal aid access.
- AutoAI streamlined the entire process, from **data preprocessing** to **model selection and evaluation**, enabling the creation of a high-performing regression model with minimal manual effort. The insights gained from this model can help **optimize CSC distribution**, identify **underrepresented communities**, and support **inclusive planning** for legal service delivery—ultimately contributing to a more equitable justice system.
- Moving forward, this predictive system can be enhanced by integrating **real-time data updates**, incorporating **socio-economic indicators**, and expanding the feature set to include factors like literacy rate or digital accessibility.

FUTURE SCOPE

- **Incorporate More Features:** Enhance the model by adding socio-economic indicators such as literacy rate, income level, and digital connectivity to improve prediction accuracy.
- **Real-Time Data Integration:** Enable the system to work with live Tele-Law data feeds for continuous model updates and timely insights.
- **District-Level Targeting:** Use the model outputs to identify and prioritize low-performing districts for focused outreach and CSC deployment.
- **Scalable Deployment:** Extend the model across other government schemes to assess service accessibility and inclusion at a national scale.
- **Dashboard Integration:** Link predictions with an interactive dashboard to help policymakers visualize disparities and act accordingly.
- **Predictive Alert System:** Develop a system that flags regions at risk of low utilization in advance, enabling preventive action.

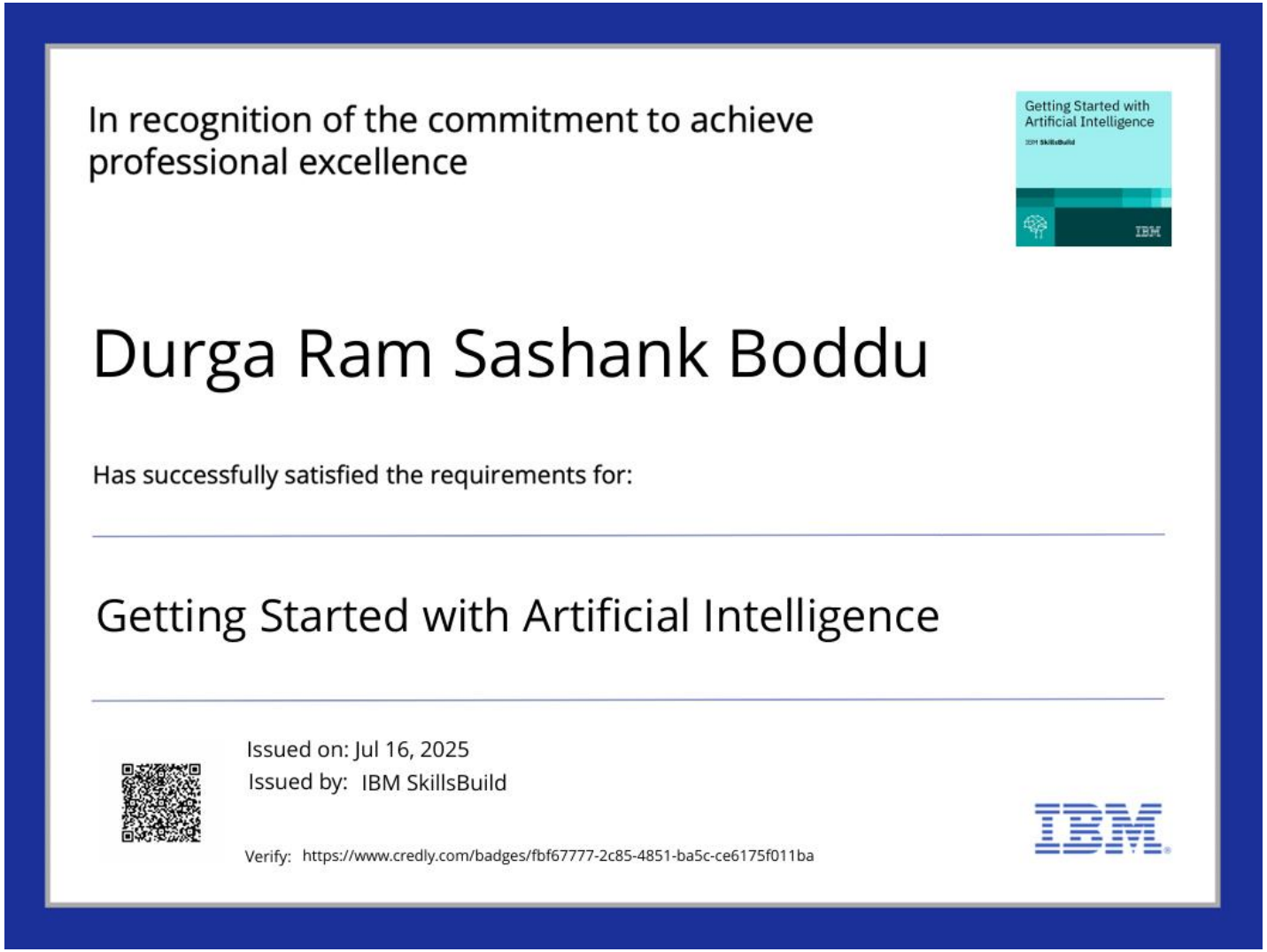
REFERENCES

- Official Dataset Source:
  District-wise Tele-Law Case Registration and Advice Enabled Data (FY 2021–22 to 2024–25) – data.gov.in
- IBM Cloud Platform: https://cloud.ibm.com/
- IBM Cloud: *Watsonx.ai Studio (AutoAI Tool Documentation)*

GITHUB LINK

Sashank206/IBM_ML-Project: IBM_Edunet Project

IBM CERTIFICATIONS



IBM CERTIFICATIONS

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution

IBM SkillsBuild

IBM

Boddu Durga Ram Sashank

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/f5e661b4-1446-42ae-a894-c2b178c79765

IBM

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Durga Ram Sashank Boddu

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