

CAPSTONE PROJECT

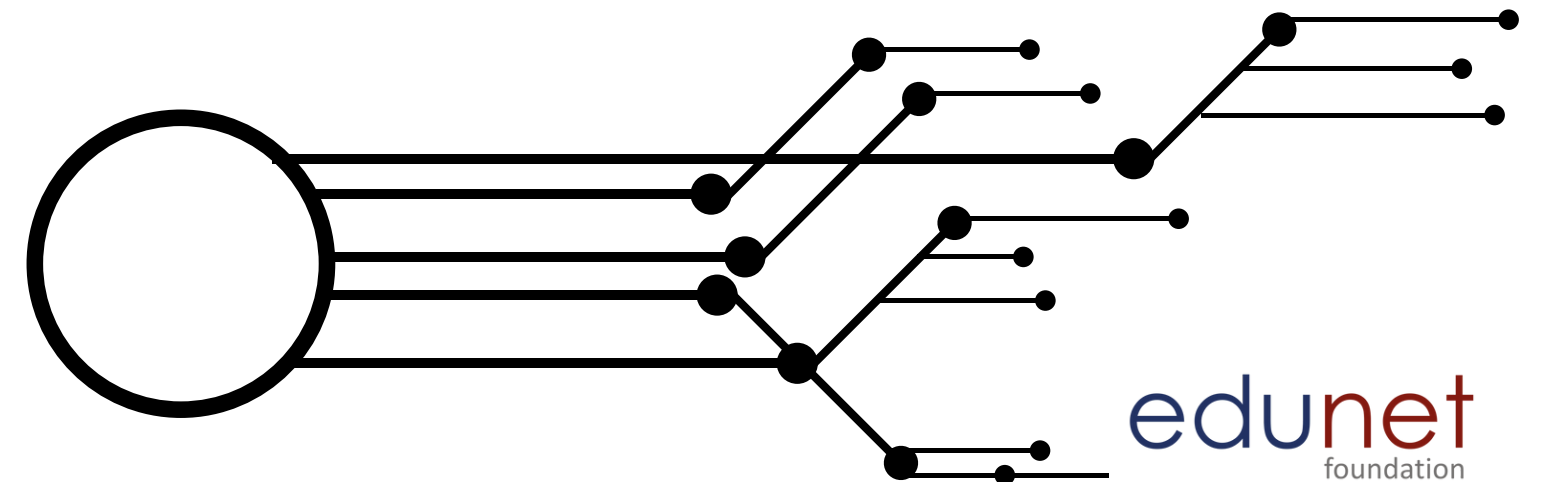
Analyzing Demographic and Regional Disparities in Tele-Law Cases

Presented By:
SHWETA NAUTIYAL
SUNDERDEEP GROUP OF INSTITUTIONS
B.TECH CSE(AIML)



OUTLINE






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



Problem Statement (37)

Despite the expansion of the Tele-Law initiative across states and districts, there is limited understanding of demographic utilization patterns and regional disparities in legal aid access. The challenge is to analyze Tele-Law case registration data to uncover gender-wise, caste-wise, and geographic disparities in service utilization across CSCs. Uneven representation among marginalized groups (SC, ST, OBC) and low outreach in certain districts raise concerns about equity and effectiveness. Moreover, the varying number of CSCs per region complicates direct comparisons. This problem demands a data-driven approach to evaluate inclusivity and optimize service delivery.

Proposed Solution

-  **Data Preprocessing:** Clean the dataset, create derived metrics (e.g., Cases per CSC, Female%, SC/ST/OBC%).
-  **Exploratory Analysis:** Visualize gender-wise, caste-wise, and regional usage patterns; highlight underserved districts.
-  **ML Component:** Use classification or clustering to identify low-access districts.
-  **Insights & Recommendations:** Suggest improvements to enhance Tele-Law outreach for marginalized groups.
-  **Deployment:** Implement using IBM Watson Studio and Cloud Object Storage.

System Approach

- **Data Collection:**
- Used district-wise Tele-Law dataset with gender, caste, and CSC data.
- **Data Preparation (IBM Data Refinery):**
 - **Removed irrelevant columns**
 - **Created new features like Percent_Female, Percent_SC, Cases_per_CSC**
 - **Added Low_Equity label based on equity conditions**

Technologies Used:

- **IBM Watson Studio – project environment**
- **IBM Data Refinery – data cleaning, feature engineering**
- **IBM AutoAI – automatic model selection and training**
- **IBM Cloud Object Storage and IBM Cloud Lite – storing datasets and outputs**
- **IBM Watson Machine Learning – model deployment & API testing**

Algorithm & Deployment

Model Type:

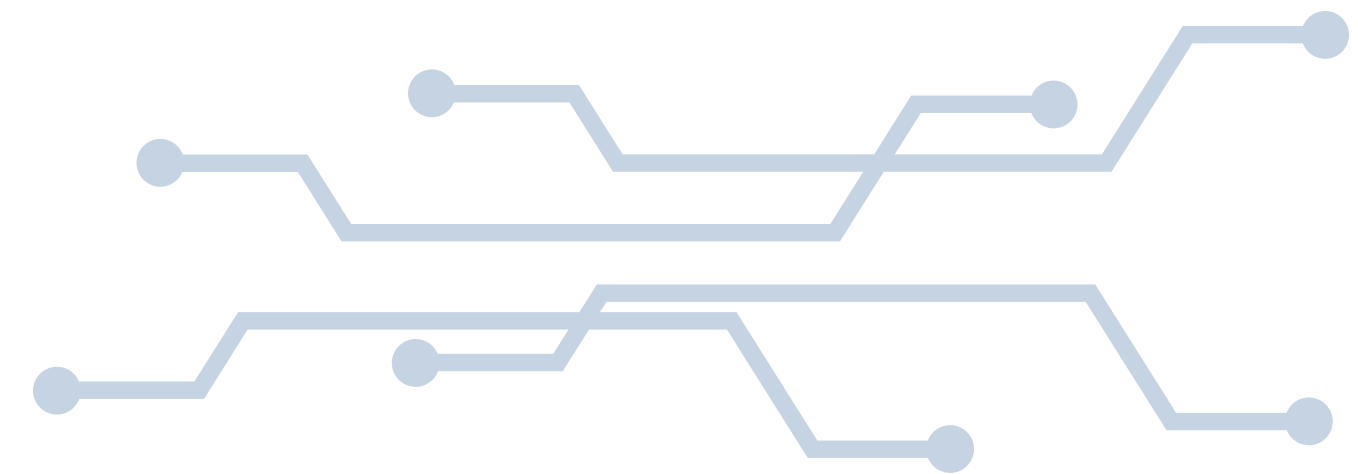
- **Binary classification** (predicting Low Equity: 0 or 1)

Algorithm:

- **AutoAI** tested multiple ML models
- Best-performing algorithm (**Decision Tree Classifier**) selected automatically based on accuracy

Deployment:

- Deployed via **Watson Machine Learning**
- Real-time predictions using test form
- **Input:** caste & gender percentages, service load
- **Output:** prediction + probability



Prediction column: low_Equity

```
graph LR; A[Read dataset] --> B[Split holdout data]; B --> C[Read training data]; C --> D[Preprocessing]; D --> E[Model selection]; E --> F[Snap Decision Tree Classifier]; E --> G[Decision Tree Classifier]; F --> H[Hyperparameter optimization]; G --> I[Hyperparameter optimization]; H --> J[Feature engineering]; I --> K[Feature engineering]; J --> L[Hyperparameter optimization]; K --> L; L --> P1; L --> P5; P1 --> P2; P2 --> P3; P3 --> P4; P5 --> P6; P6 --> P7; P7 --> P8; style P8 stroke-width:4px
```

Swap view

Experiment completed

8 PIPELINES GENERATED

8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 3 minutes

View log

Save code

Pipeline leaderboard

	Rank	Name	Algorithm	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 8	Decision Tree Classifier	1	HPO-1 FE HPO-2	00:00:20

Result

IBM watsonx.ai Studio

Search in your workspaces

Upgrade

Shweta Nautiyal's Account

Sydney

SN

Deployment spaces / Tele_Law_Project / P8 - Decision Tree Classifier: Auto_AI_Tele_Law_3 /

Te

Prediction results

Close

X

Display format for prediction results

☒ Table view

☐ JSON view

Show input data

	prediction	probability
1	0	[1,0]
2	0	[1,0]
3	1	[0,1]
4	0	[1,0]
5	1	[0,1]
6	1	[0,1]
7	0	[1,0]
8	1	[0,1]
9	0	[1,0]
10		

Download JSON file

Charts

Search in your workspaces

Upgrade



Shweta Nautiyal's Account

Sydney

Chart

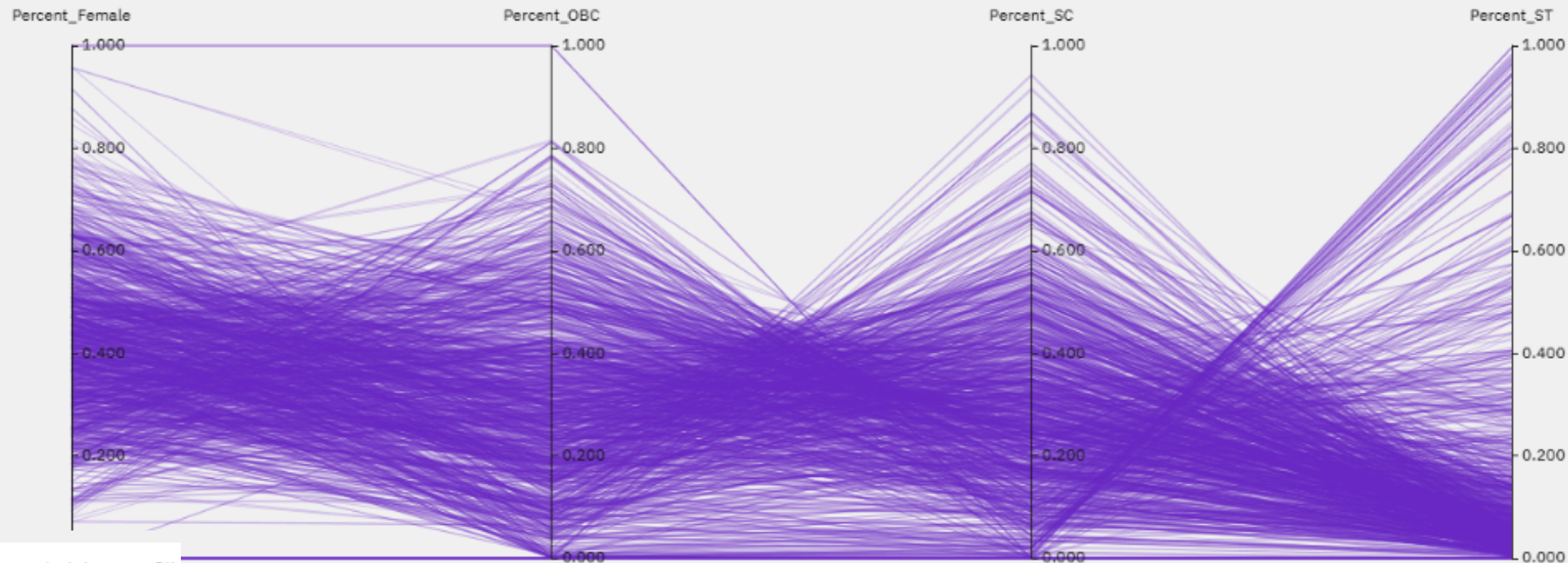


Output

Edit

parallel_2025080455948fs0pn0rxh9

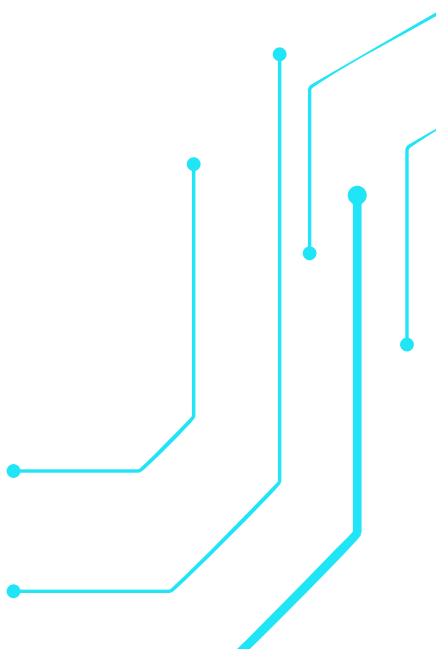
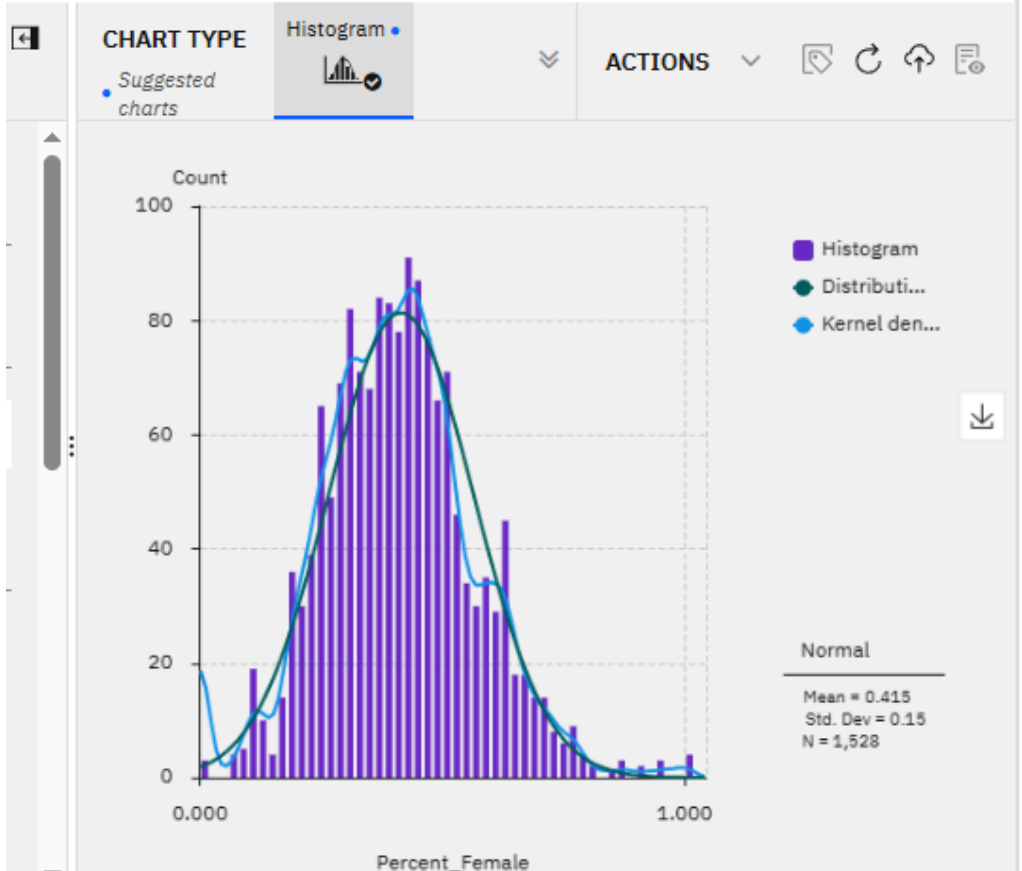
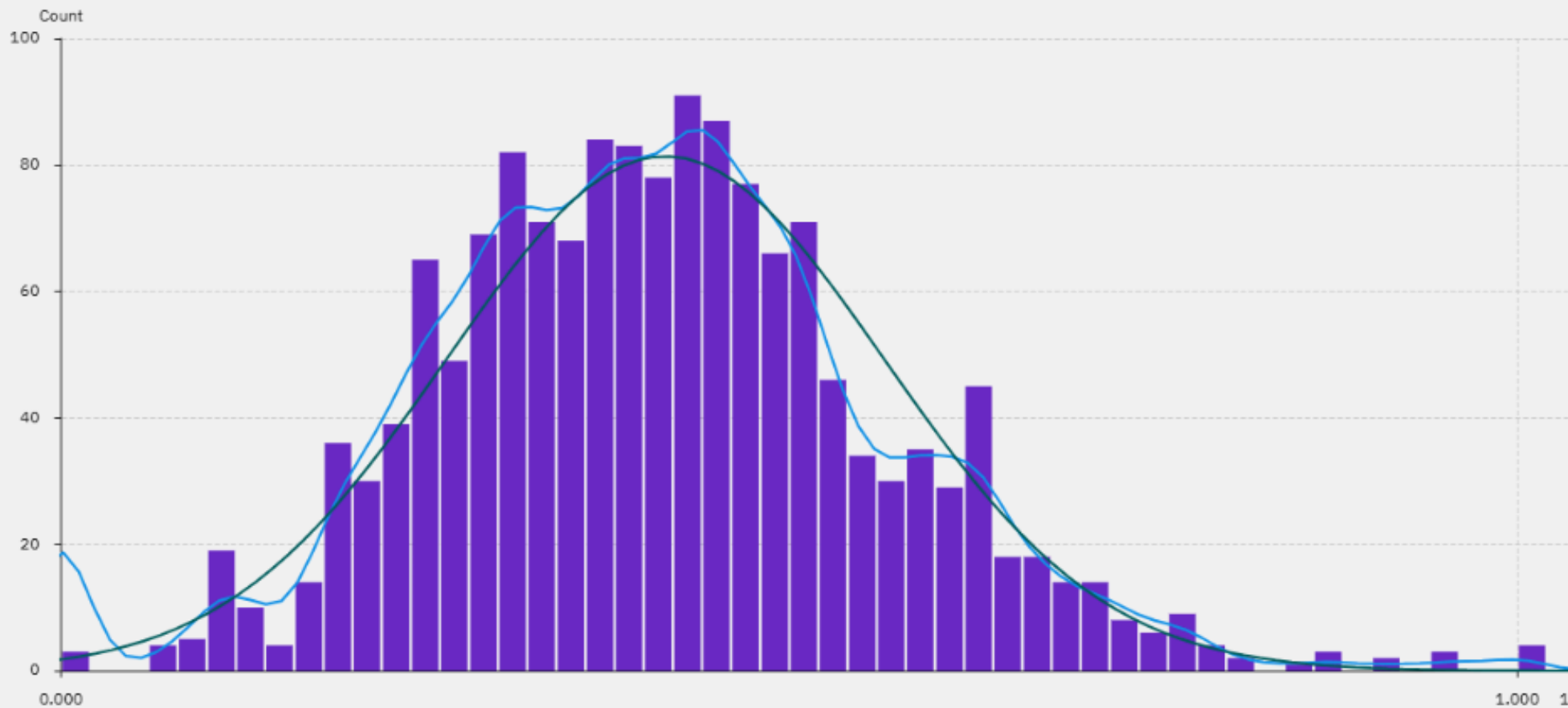
Chart is based on sampled dataset. Click edit



Edit

histogram_202508042453k1hdqx0hab

Chart is based on sampled dataset. Click



Conclusion

- The model successfully predicts whether a district is experiencing **low equity in Tele-Law access**.
- It uses **gender and caste distribution** to assess inclusiveness.
- The system helps identify underrepresented regions needing more outreach.
- Using **IBM's AutoAI and deployment tools** simplified the ML workflow.



Future scope

- **Add more features:** education level, rural/urban split, internet access
- Use geospatial data to **visualize** regional disparities
- Deploy as a **mobile/web app** for decision-makers
- **Expand** to other government schemes to assess inclusivity



References

- IBM Cloud Documentation
- AutoAI official docs
- Tele-Law data source (<https://www.data.gov.in/resource/district-wise-tele-law-case-registration-and-advice-enabled-data-fy-2021-22-2024-25>)
- IBM Watson Studio Tutorials

GitHub Repository Link

<https://github.com/Shweta141203/telelaw-equity-prediction>

IBM Certifications

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to

Shweta Nautiyal

for the completion of

Introduction to Artificial Intelligence

(MDL-211)

According to the Moodle system of record

Completion date: 20 Jul 2025 (GMT)

Learning hours: 1 hr 15 mins

In recognition of the commitment to achieve
professional excellence



Shweta Nautiyal

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence



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

Verify: <https://www.credly.com/badges/7fa35f8f-2554-4fdc-a942-a0d987ea1921>



IBM Certifications

IBM **SkillsBuild**

Completion Certificate



This certificate is presented to

Shweta Nautiyal

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