

TENSIONFLOW
PRESENTS



**ATTORNEY-CLIENT
MATCHER**

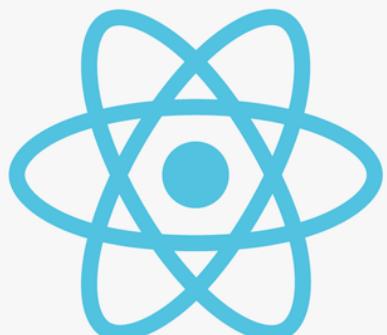
KASHISH GANDHI - ML, FRONTEND
HEMANT SINGH - ML, BACKEND
ABHAY MATHUR - ML, BACKEND
SHREYA SHAH - ML, FRONTEND

PROBLEM STATEMENT 2

To develop an MLOps-powered platform for precise client-attorney matchmaking. This platform will utilize advanced techniques to connect clients with attorneys based on their unique legal concerns. It should efficiently process user requests and ensure an equal gender ratio among recommended lawyers. The goal is to streamline legal expertise access, revolutionizing the legal industry by providing a user-friendly and efficient experience. Additionally, a vector database will be implemented to store and search word embeddings of lawyer data. Finally, a Client-Server Interface will be developed for seamless query input and recommendations.



TECH STACKS



React

django

Flask



VectorDB



Streamlit



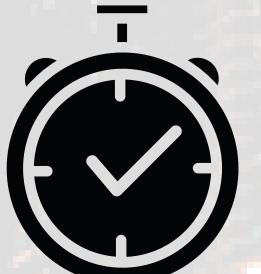
OpenAI

falcon a.i.

PROJECT TIMELINE

Data preparation and enhancement are necessary for accurate model training. Creating a robust algorithm for precise client-attorney matchmaking and designing an intuitive platform for seamless user interaction are both important steps.

PreProcessing



Models



User Interface



PREPROCESSING

Initial Dataset - Using LLM and Regex to extract keywords from sentences

Part 1

Anay Ahluwalia has **20** years of experience in **Corporate Law, Consumer Protection Law, and Labor Law**.

He has a Client Feedback of **5.0** out of 5.0.

Part 3

His Jurisdiction is **Supreme Court**.

Part 5

He charges **420.6850078605348** USD per hour.

Part 2

Part 4

He takes **117.69593880898871** Avg Days for Disposal.

He speaks: **Telugu, Hindi** and **Hindi**.

Part 6

Part 7

He practices at **Saini PLC**, and is based in **Hyderabad**.

She provides **pro bono** services to the community.

Part 8

Her Client Demographics is **Large Corporations**.

Part 9

PREPROCESSING

Gender Classification

Data Discrepancy

Vedika Vala has 8 years of experience in Environmental Law, Corporate Law, and Criminal Law. He has a Client Feedback of 4.0 out of 5.0.

~~Canva~~

~~Canva~~

✓. He has a Client Feedback of 5.0 out of 5.0. His Jurisdiction is Supreme Court. He charges 420.6850078605348 USD per hour.
~~Canva~~ ~~bad~~. She provides pro bono services to the community. Her Client Demographics is Large Corporations.

~~Canva~~

- **Data Discrepancy:** "Vedika" had inconsistent gender assignments, along with some other lawyers which included "He" and "She" both in the same information
- **Custom LSTM Model:** Developed an LSTM neural network for gender prediction on basis of name, trained on Indian names data from Kaggle.
- **Genderize.io API:** Integrated a third-party API to assist in gender predictions based on names.
- **Combined Approach:** Using both the LSTM model and API to ensure consistent and accurate gender assignments.

```
...
Epoch 49/50
1350/1350 [=====] - 15s 11ms/step - loss: 0.0156 - accuracy: 0.9955 - val_loss: 18.4643 - val_accuracy: 0.0449
Epoch 50/50
1350/1350 [=====] - 15s 11ms/step - loss: 0.0170 - accuracy: 0.9946 - val_loss: 20.0582 - val_accuracy: 0.0380
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
<keras.callbacks.History at 0x2247774d990>
```

Accuracy: 99.46%

PREPROCESSING

Gender Classification

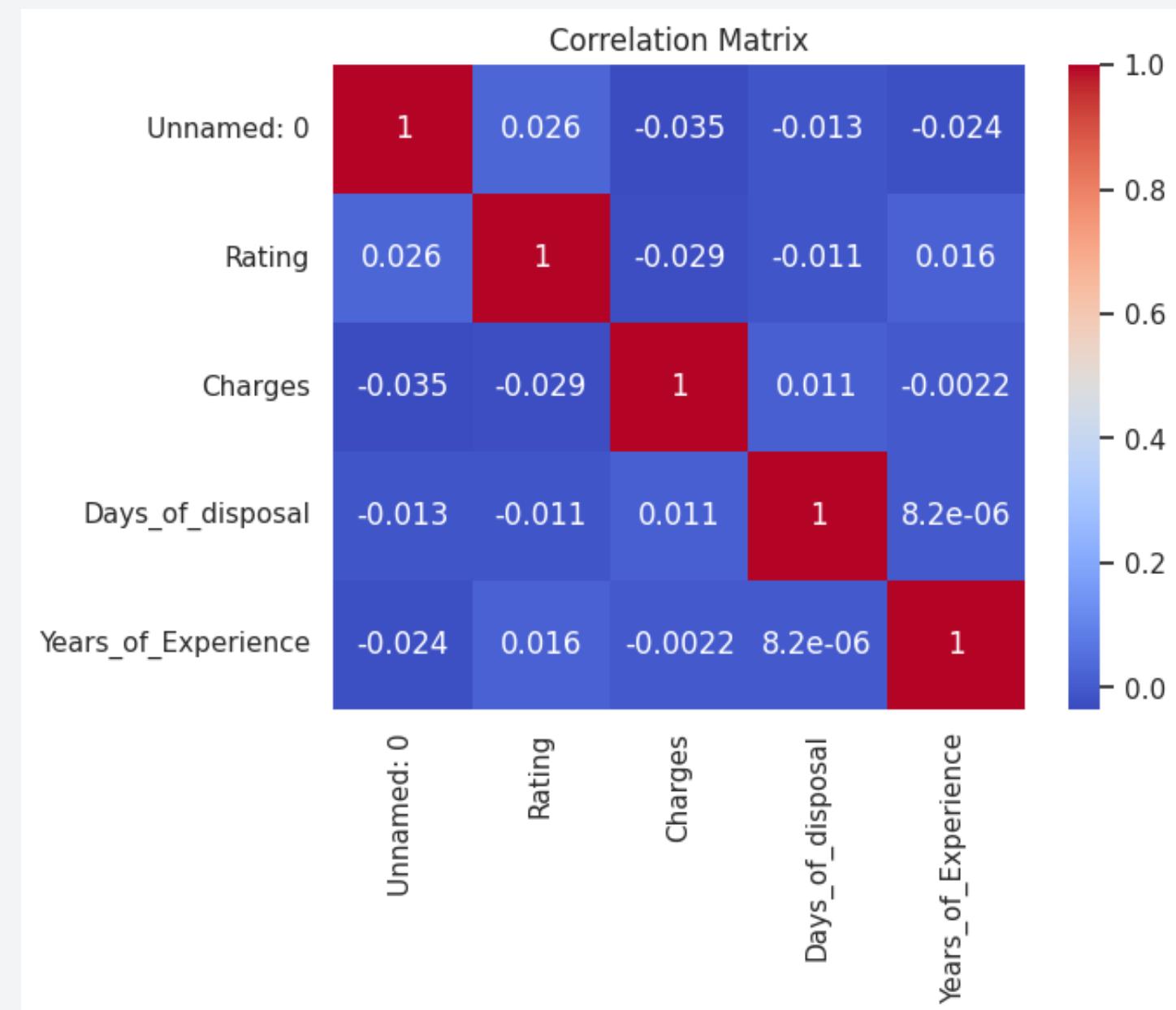
Addition of Non-Binary gender

Incorporating Non-binary Gender Data with Faker Library:

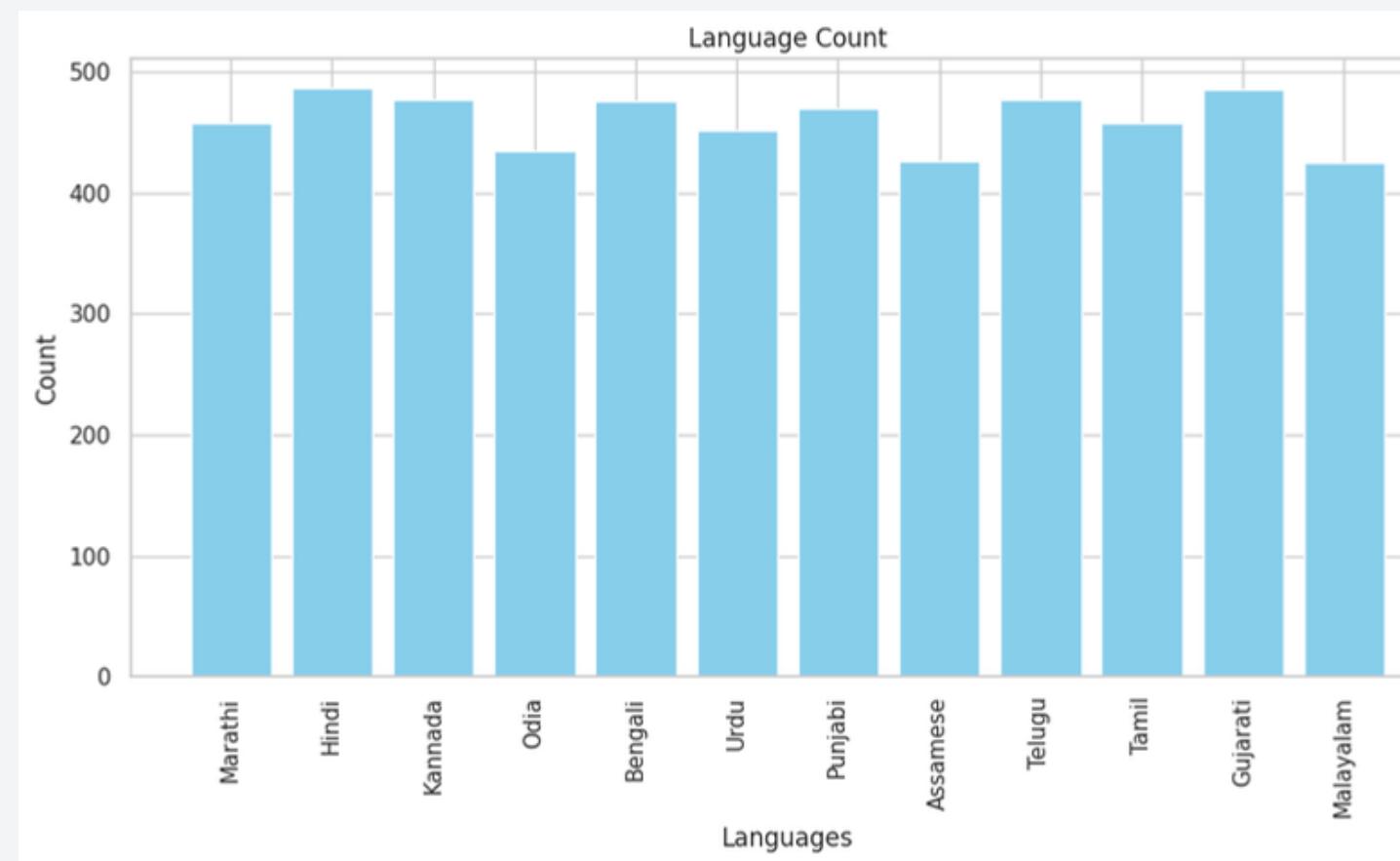
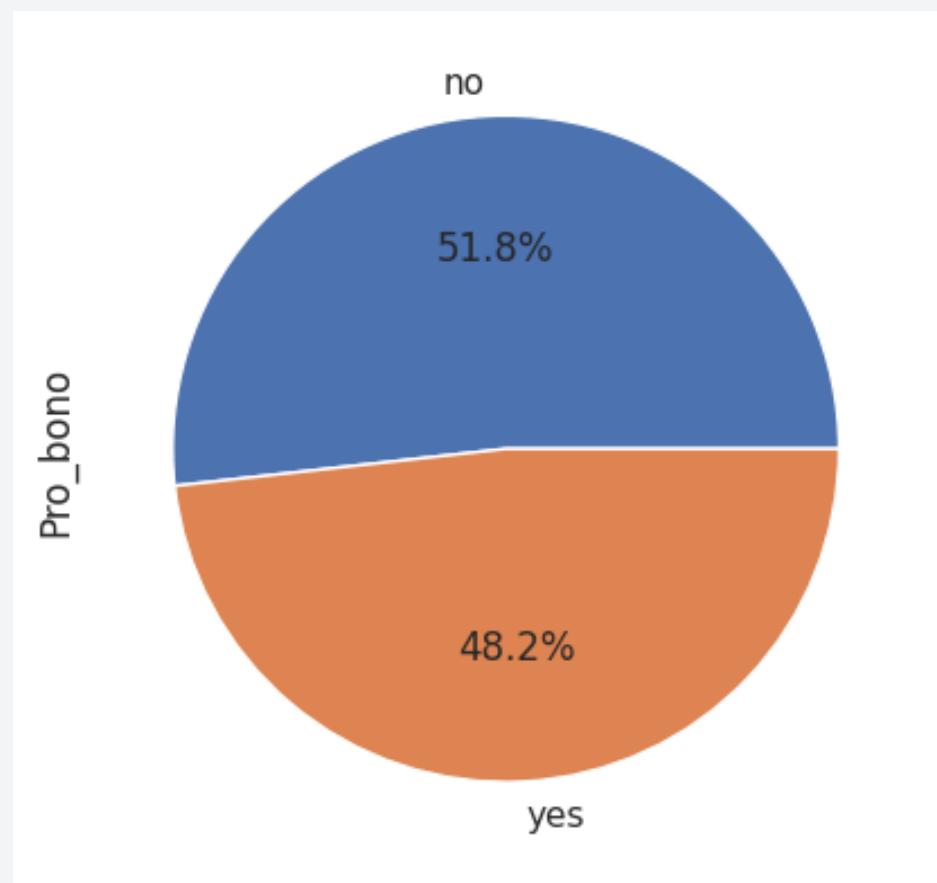
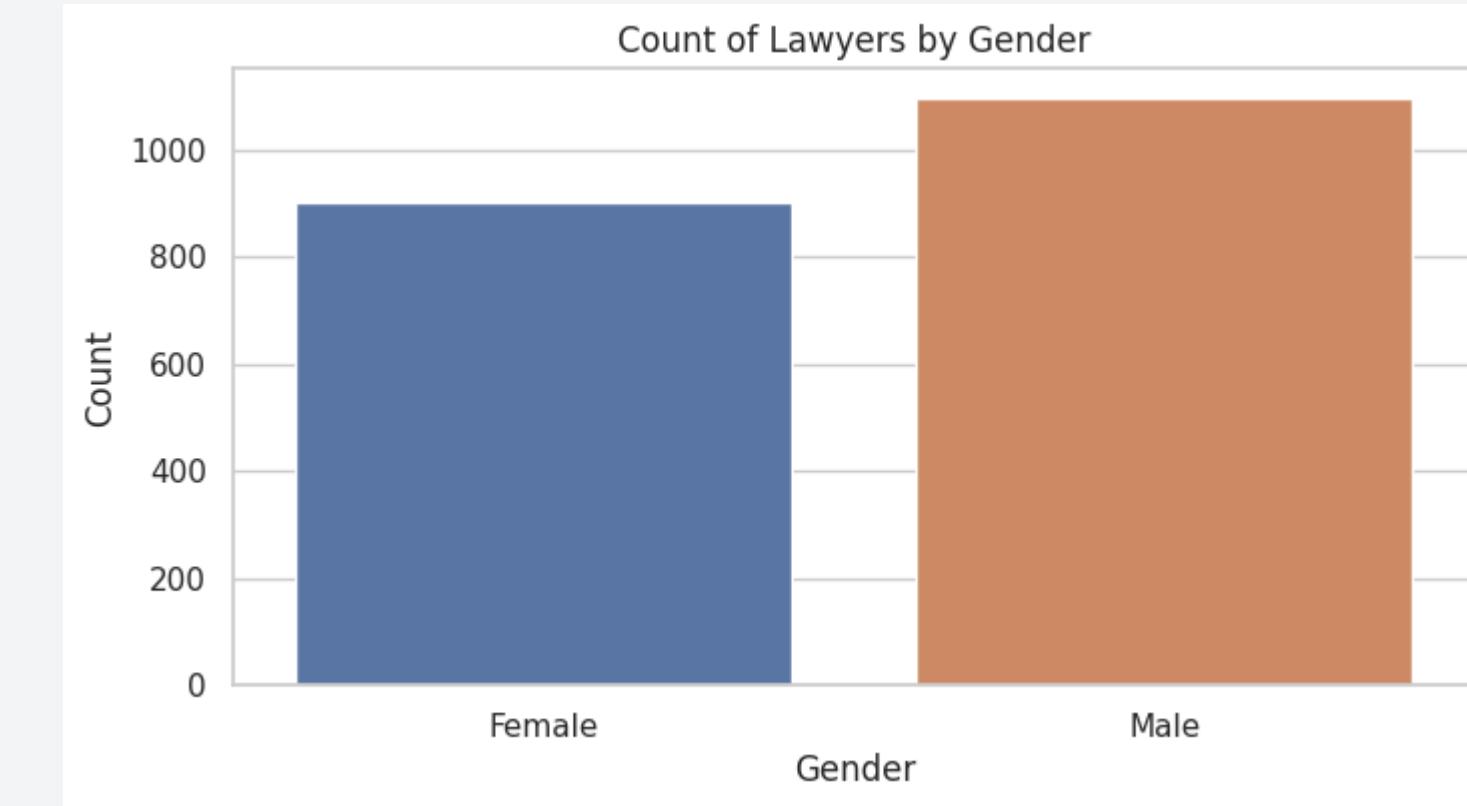
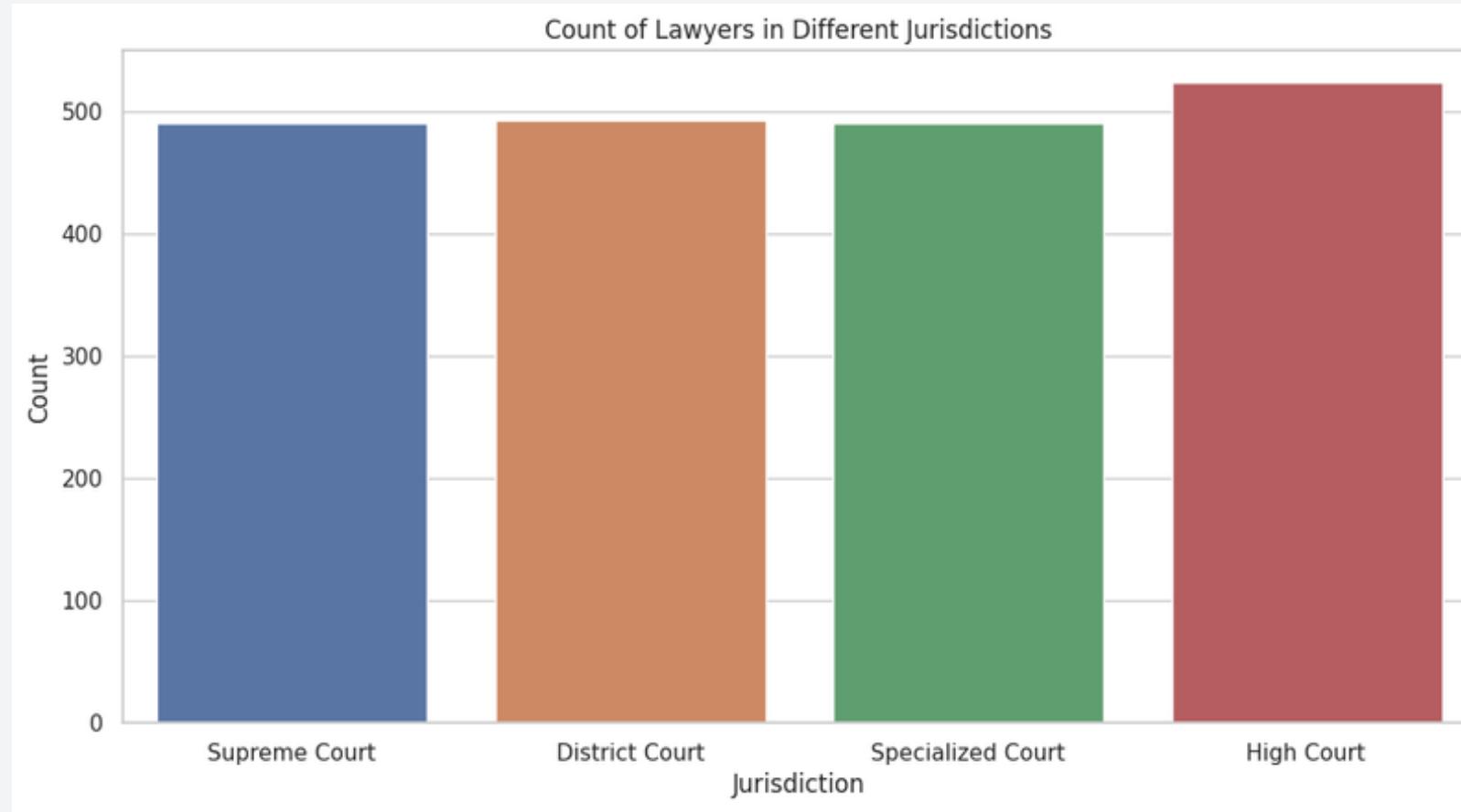
- We encountered challenges when trying to include non-binary gender data with the Faker library in Python.
- To maintain data integrity and preserve existing correlations, we opted not to directly inject non-binary data into the real-time dataset.

Consideration of Data Correlations:

- Recognizing the potential disruption to data correlations, especially in fields like criminal law, we decided to explore alternative approaches.
- Future Integration Possibility: It's important to note that while we chose not to add the data directly, the system is designed in a way that makes it easy to incorporate this feature in the future, allowing for a more robust and inclusive dataset.

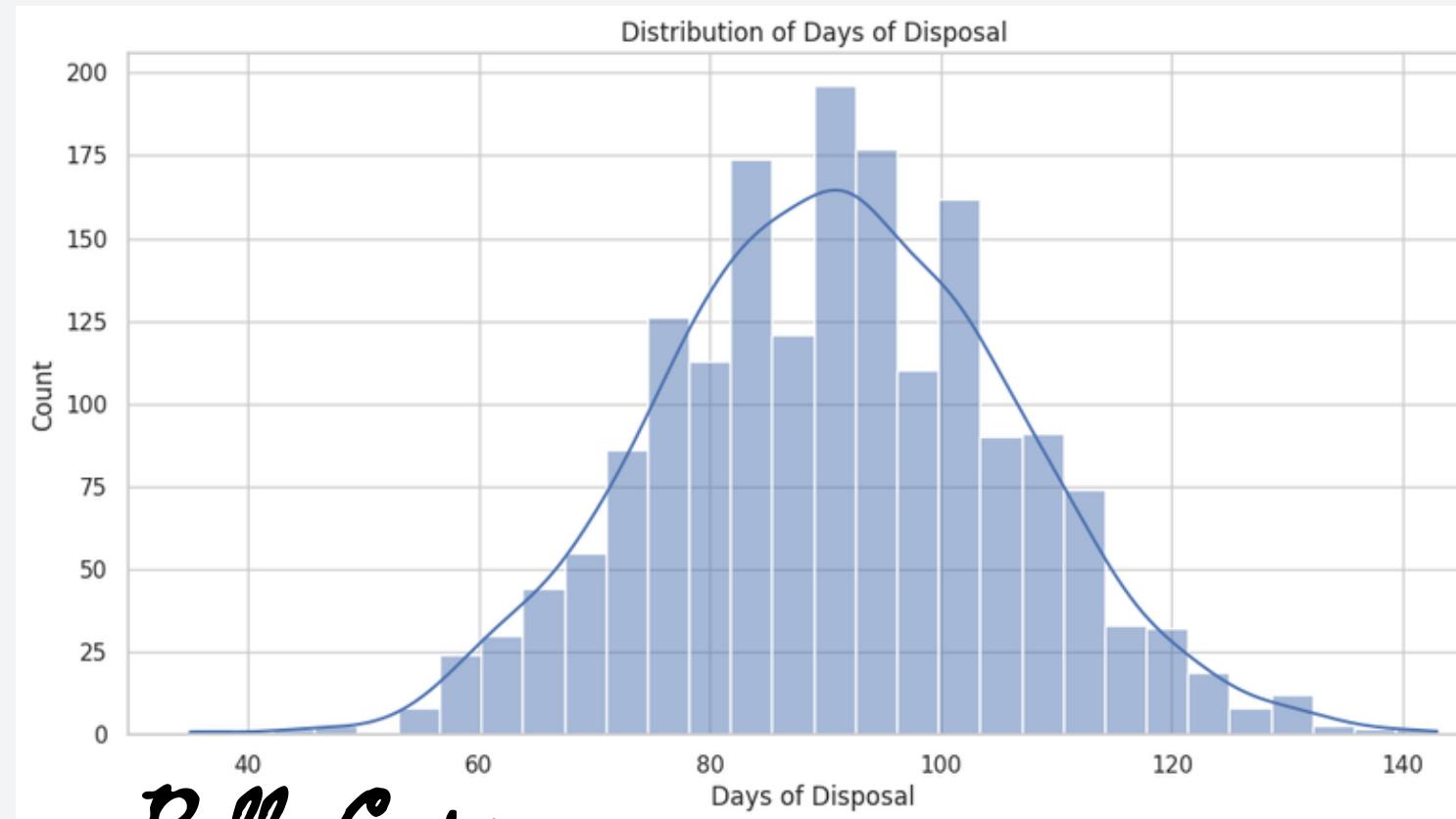


EXPLORATORY DATA ANALYSIS



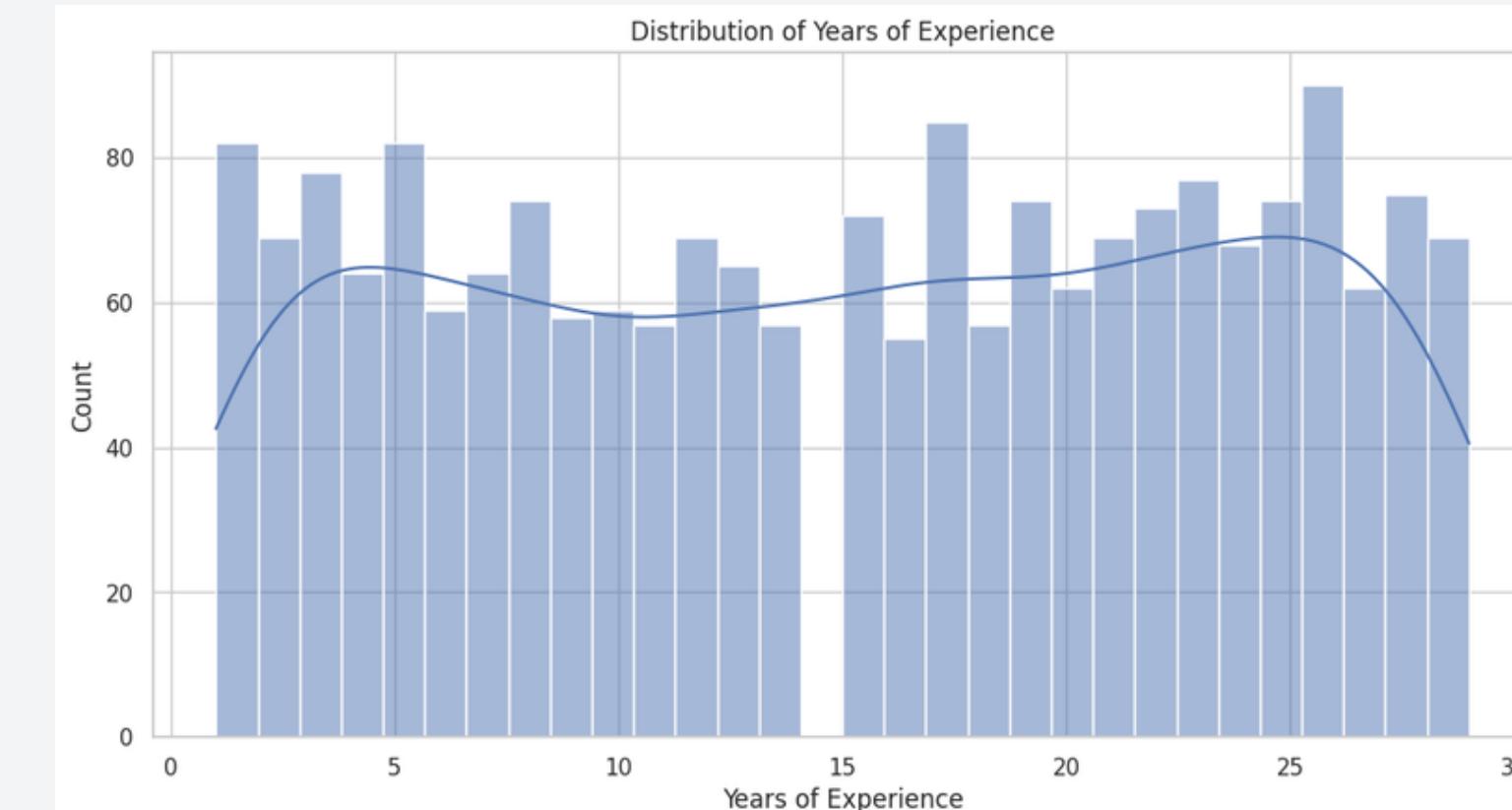
Data was quite consistent and other discrepancies were less

FILLING NAN VALUES

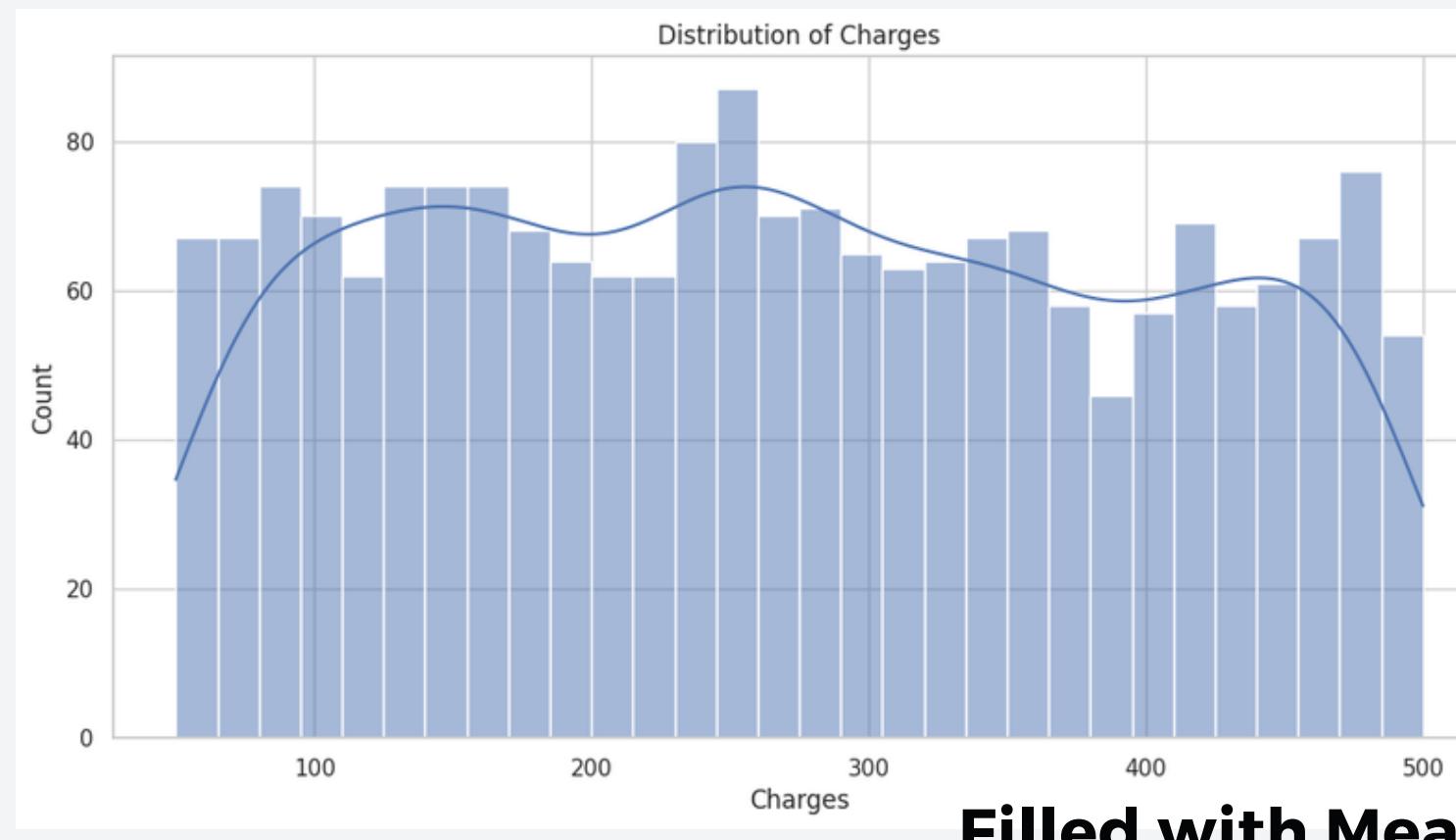


Bell-Curve

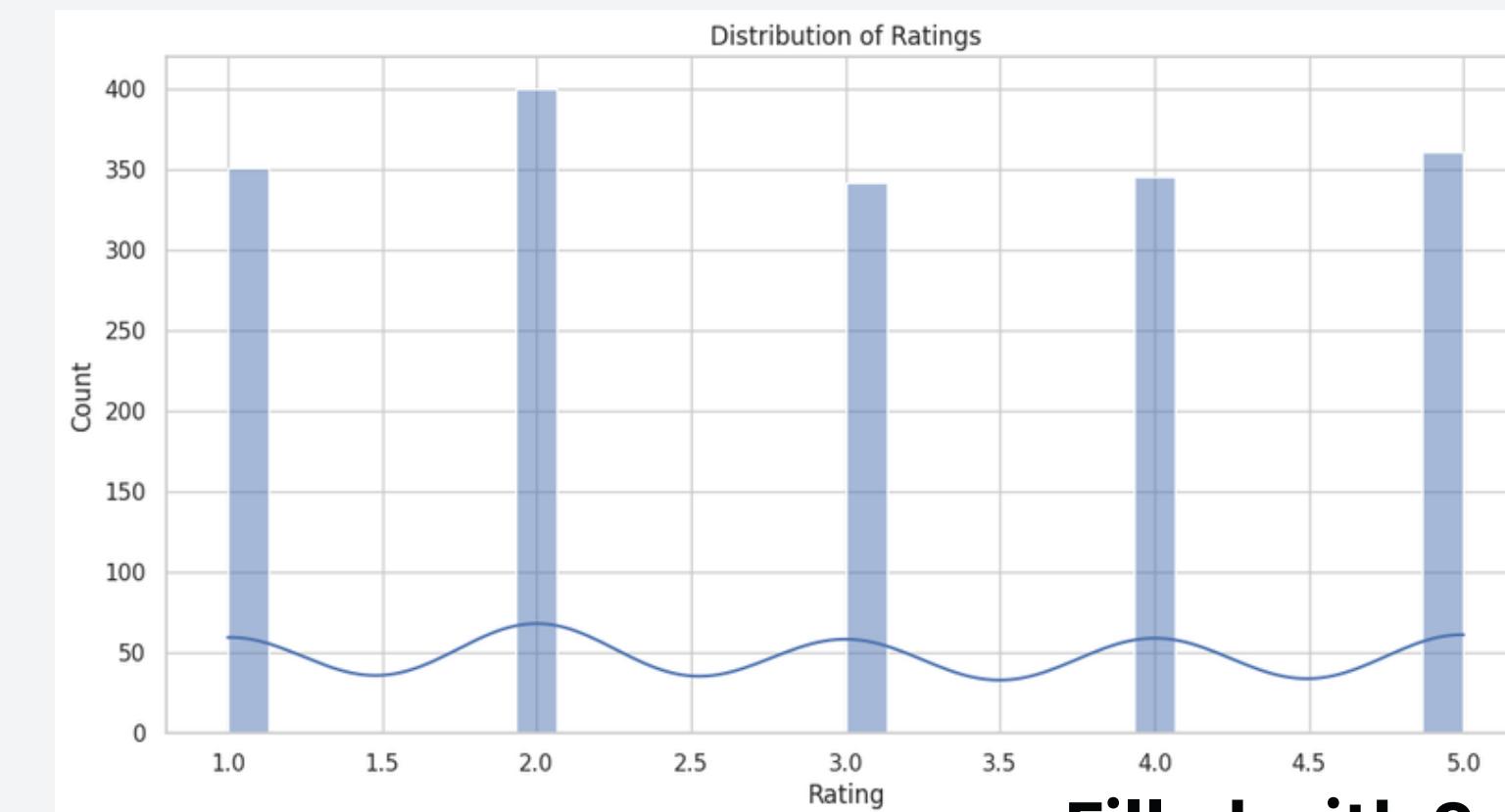
Filled with Mode



Filled with Mean



Filled with Mean



Filled with 0

FINAL DATASET

From this

	Lawyer Names	Information
0	Anay Ahluwalia	Anay Ahluwalia has 20 years of experience in C...
1	Amav Dubey	Amav Dubey has 29 years of experience in Inte...
2	Lavanya Lalla	Lavanya Lalla has 15 years of experience in Hu...
3	Armaan Khosla	Armaan Khosla has 11 years of experience in Co...
4	Advik Chaudry	Advik Chaudry has 8 years of experience in Hum...

To this

Unnamed: 0	Rating	Jurisdiction	Charges	Days_of_Disposal	City	Years_of_Experience	Type_of_Lawyer	Languages	Pro_bono	Lawyer_Names	Law_firms	Demography	Gender	
0	0	5.0	Supreme Court	489	107.0	[Chennai]	21	['Constitutional', 'Immigration', 'Constitutio...']	['Marathi', 'Hindi']	no	Aaina Ganesan	[Tara Inc]	Small Businesses.	Female
1	1	4.0	District Court	70	91.0	[Mumbai]	15	['Human Rights', 'Labor', 'Corporate']	['Kannada', 'Hindi', 'Odia']	yes	Aaina Gopal	[Bhagat, 'Valdyas', 'Varma']	Small Businesses.	Female
2	2	4.0	District Court	282	106.0	[Mumbai]	22	['Medical', 'Immigration', 'Civil']	['Bengali', 'Urdu']	yes	Aaina Kanda	[Anne-Srinivasan]	Individuals.	Female
3	3	1.0	Specialized Court	490	103.0	[Mumbai]	6	['Consumer Protection', 'Environmental', 'Envi...']	['Punjabi', 'Bengali', 'Assamese']	yes	Aaina Kaul	[Amble Ltd]	Individuals.	Female
4	4	1.0	Specialized Court	441	85.0	[Hyderabad]	7	['Corporate', 'Criminal', 'Labor']	['Telugu', 'Urdu', 'Punjabi']	no	Aaina Kurian	[Mall-Tiwari]	Individuals.	Female

OUR APPLICATION

The image displays three screens of a mobile application interface, likely for a lawyer search and feedback platform, set against a background of a city skyline at dusk.

Top Left Screen (Large): Lawyer Finder

- Header:** Lawyer Finder
- Text:** Please provide details of your case*
- Form:** Give us a short description of your case (text input field)
- Text:** Location: *
- Form:** CHOOSE CITY (dropdown menu)
- Text:** Language Preference 1: * Language Preference 2: * Language Preference 3:
- Form:** CHOOSE LANGUAGE (dropdown menu) for all three fields
- Text:** Service Type: Case Type: Jurisdiction Level:
- Form:** CHOOSE SERVICE TYPE, CHOOSE CASE TYPE, CHOOSE JURISDICTION LEVEL (dropdown menus)
- Button:** Find Lawyer
- Result Card:** Mahika Agrawal, Real Estate Law, ★ 3.0, Experience: 13 years, Avg Charge: 216.06 USD/hour

Bottom Left Screen (Medium): Lawyer Profile

- Header:** Mahika Agrawal
- Text:** Real Estate Law, Rating: ★ 3.0, Experience: 13 years, Avg Charge: 216.06 USD/hour
- Text:** Upload PDF
- Text:** Please Upload the PDF of your last case here to receive an analysis
- Button:** Upload

Top Right Screen (Small): Feedback Form

- Header:** Mahika Agrawal
- Text:** Real Estate Law, Rating: ★ 3.0, Experience: 13 years, Avg Charge: 216.06 USD/hour
- Section:** Your Feedback
- Text:** Rating
- Form:** OUT OF 5 RATING (dropdown menu)
- Text:** Please provide detailed description of your experience (text input field)

MODELS USED

01

YAKE

Yet Another Key Extractor is used to extract the keywords from the user's query and then these queries are mapped to their respective categories

04

LSTM

To find the gender of a lawyer based on their first name (due to the discrepancy in the pronouns in the dataset)

02

FalconAI

With the help of the query as well as the keywords we got from the YAKE to classify the data into the 3 main buckets

05

Bert

We use Bert for Sentiment Analysis to make attorney ratings more accurate

03

Match-Making AI

After finding out where the query belongs with our sophisticated techniques like content collaborative filtering

06

OpenAI

We use OpenAI API to read the PDF page by page and draw important points about the case from its file



RECOMMENDATION SYSTEM

Query Preprocessing:

- User input query is received and tokenized.
- YAKE (Yet Another Keyword Extractor) algorithm is applied to extract relevant keywords from the query. YAKE identifies significant terms to represent the user's intent.

Legal Classification with FalconAI:

- Extracted keywords are utilized to query FalconAI, a (LLM).
- FalconAI classifies the query into specific law-based categories such as Criminal Law, Family Law, etc., using its pre-trained knowledge of legal domains.

RECOMMENDATION SYSTEM

Keyword Match Score Calculation:

- For each lawyer's description, a match score is computed by assessing the overlap between the extracted keywords and the keywords associated with the lawyer's expertise.
- The more common keywords, the higher the match score, reflecting relevance to the user's query.

Jaccard Similarity Analysis:

- Jaccard similarity is calculated between the set of keywords in the query and the set of keywords associated with each lawyer's expertise.
- Lawyers with higher Jaccard similarity are ranked more favorably, indicating a closer alignment with the user's query.

RECOMMENDATION SYSTEM

Collaborative Filtering for Lawyer Recommendation:

- Collaborative filtering techniques are applied to suggest the top 10 lawyers. This is based on collective user preferences and behavior, identifying lawyers who have been well-received or frequently chosen by users with similar queries.

Weighted Ranking:

- A weighted ranking system is employed to prioritize lawyers based on multiple factors. The weights are assigned as follows:
- Highest weight to the match score, reflecting the relevance of law-based categories.
 - Additional weights for user ratings, with higher-rated lawyers receiving preference.
 - Years of experience contribute to the ranking, favoring more experienced lawyers.
 - Charges for legal services are also considered, with reasonable fees being advantageous in the ranking.

RECOMMENDATION SYSTEM

Legal Chatbot and Lawyer Recommender

Enter your legal query:

Two brothers were tenant of a landlord in a commercial property. One brother had one son and a daughter (both minor) when he got divorced with his wife. The children went into the mother's custody at the time of divorce and after some years the husband (co tenant) also died. Now can the children of the deceased brother (co tenant) claim the right.

Get Recommendations

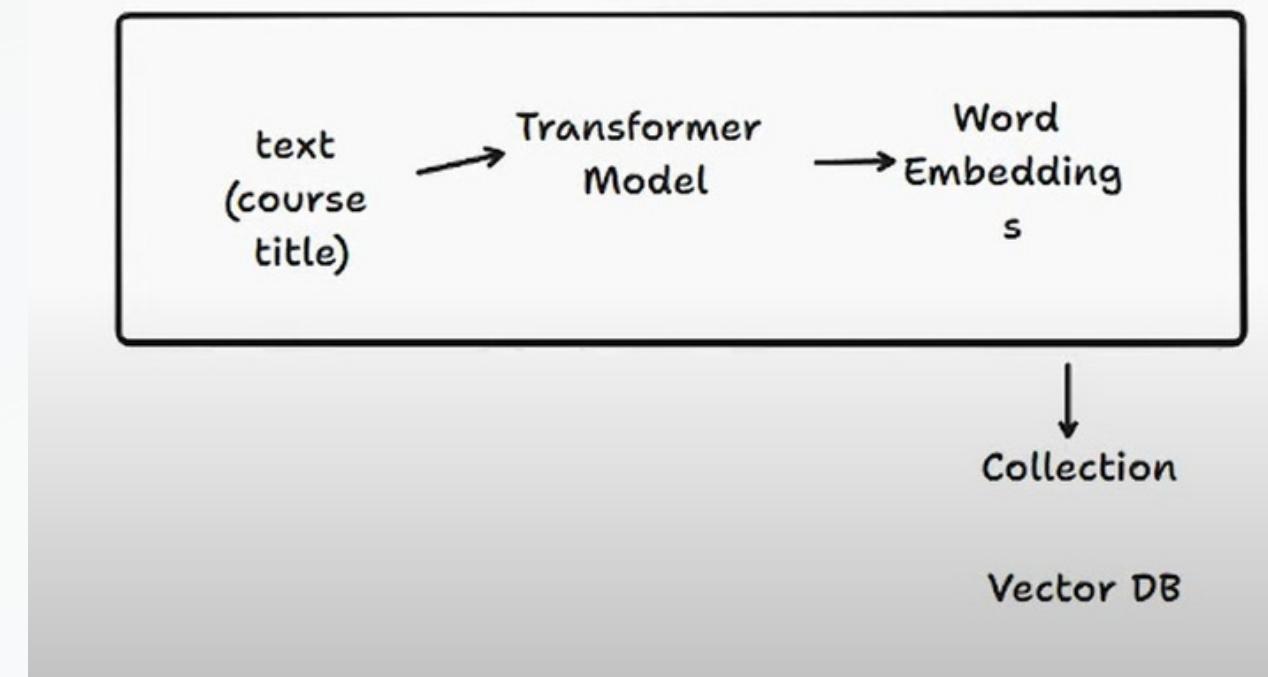
Top 10 Recommended Lawyers:

	Days_of_disposal	City	Years_of_Experience	Type_of_Lawyer
316	100	['Mumbai']	13	['Civil', 'Real Estate', 'Family']
68	100	['Delhi']	28	['Civil', 'Family', 'Real Estate']
1,363	91	['Hyderabad']	28	['Family', 'Civil', 'Consumer Protection']
819	95	['Delhi']	28	['Family', 'Civil', 'Banking and Finance']
684	83	['Hyderabad']	27	['Civil', 'Consumer Protection', 'Family']
608	97	['Chennai']	26	['Family', 'Immigration', 'Real Estate']
102	89	['Mumbai']	26	['Real Estate', 'Real Estate', 'Civil']
1,232	87	['Bangalore']	25	['Human Rights', 'Family', 'Real Estate']
339	88	['Bangalore']	24	['Family', 'Civil', 'Family']
600	86	['Hyderabad']	23	['Family', 'Civil', 'Medical']

"input": "Two brothers were tenant of a landlord in a commercial property. One brother had one son and a daughter (both minor) when he got divorced with his wife. The children went into the mother's custody at the time of divorce and after some years the husband (co tenant) also died. Now can the children of the deceased brother (co tenant) claim the right."

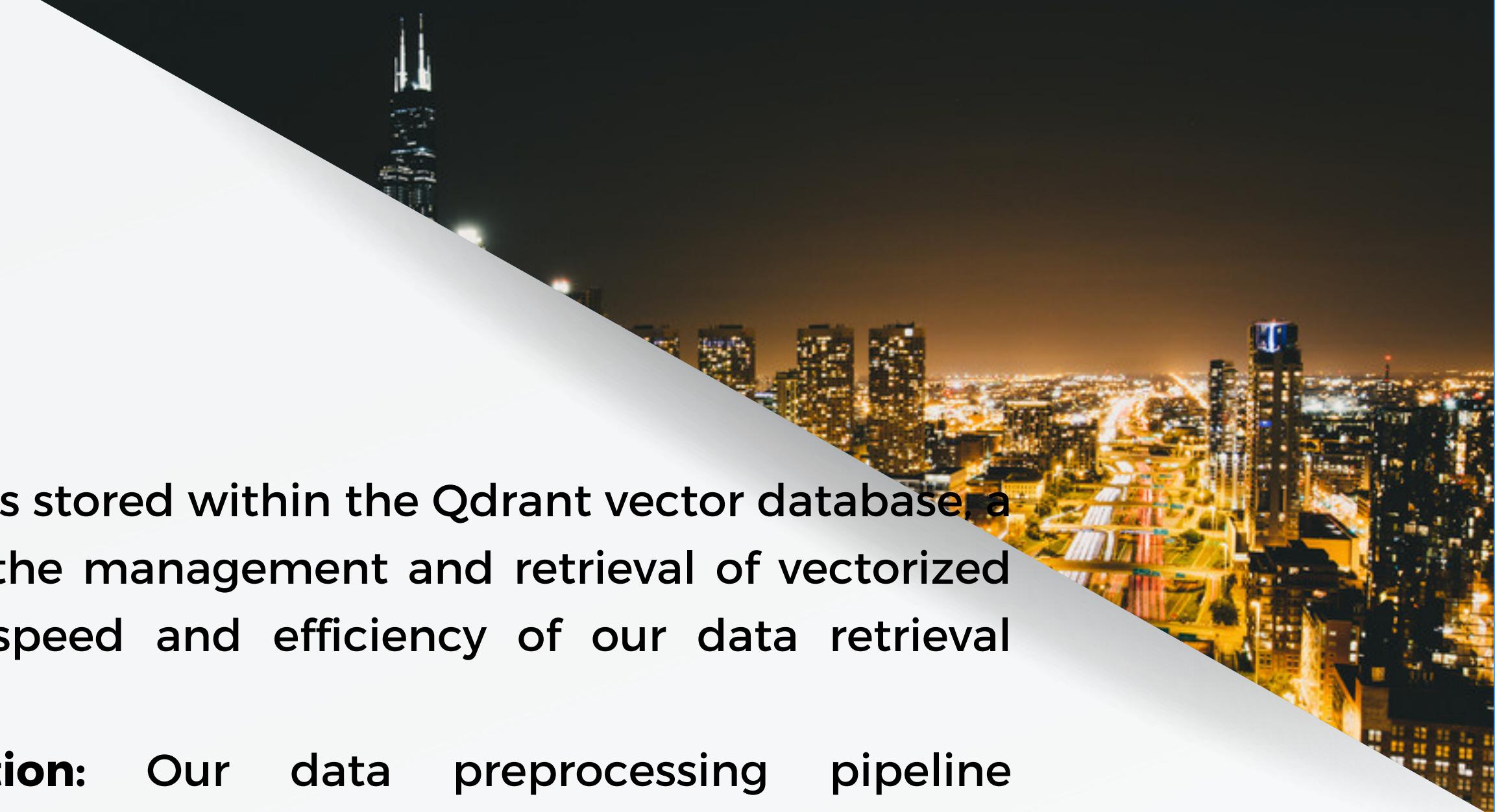
VECTORIZED DATABASE

- **Space Complexity Optimization:** We've devised a method to create embeddings for sentences that optimizes space usage, resulting in a highly efficient data storage and retrieval system.
- **Cosine Similarity for Classification:** We utilize cosine similarity, a widely recognized technique for comparing vector representations, to classify and discern semantic relationships between sentences with remarkable accuracy.



VECTORIZED DATABASE

- **Qdrant Vector Database:** Our data is stored within the Qdrant vector database, a specialized platform designed for the management and retrieval of vectorized data. This choice enhances the speed and efficiency of our data retrieval operations.
- **Sentence Transformer Integration:** Our data preprocessing pipeline incorporates the cutting-edge Sentence Transformer model, specifically the all-MiniLM-L6-v2 variant. This transformer model excels at converting sentences into dense vector representations, capturing their semantic content with precision.
- **Vectorized Data Integration:** We have seamlessly integrated our vectorized data into the Qdrant vector database, leveraging advanced vector storage and retrieval techniques. This integration ensures that our system operates at the intersection of efficient storage and retrieval.



VECTORIZED DATABASE

- **Data-Driven NLP:** With this sophisticated technological stack, we have established a data-driven approach to sentence embedding and classification. Our system sets a new standard for efficiency and accuracy in natural language processing tasks, marking a significant advancement in the field.

The screenshot shows a database management interface with a sidebar on the left displaying a file tree and a main table view on the right.

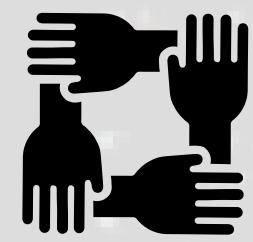
File Tree (Left):

- DATAHACK_2_TENSIONFLOW
- Vector Database
- recproject
- recommender
- __pycache__
- management
- migrations
- templates
- vector_database.db
- collection\lawyers_collection
- storage.sqlite
- .lock
- meta.json
- __init__.py
- admin.py
- apps.py
- filters.py
- forms.py
- models.py
- services.py
- tests.py
- urls.py
- utils.py
- vectorized_lawyers.pickle

Table View (Right):

	id	point	
1	gARLAC4=	3.93 KB	
2	gARLAS4=	3.94 KB	
3	gARLAi4=	3.91 KB	
4	gARLY4=	3.95 KB	
5	gARLB4=	3.92 KB	
6	gARLB54=	3.95 KB	
7	gARLBi4=	3.92 KB	
8	gARLBy4=	3.94 KB	
9	gARLC4=	3.93 KB	
10	gARLC54=	3.95 KB	
11	gARLCi4=	3.94 KB	
12	gARLCy4=	3.94 KB	
13	gARLD4=	3.95 KB	
14	gARLD54=	3.94 KB	
15	gARLDi4=	3.95 KB	
16	gARLDy4=	3.93 KB	

UNIQUE SELLING POINTS



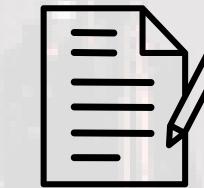
Sentimental Analysis

In addition to the rating that the user provides out of 5 stars, we used sentimental analysis on the comment they leave as feedback to make the rating more accurate



PDF Summariser

We allow the lawyer to upload the Case file PDF and get required information about the case like the judgement, location, etc.



Ask AI Chatbot

Ask AI Chatbot, an advanced AI tool that simplifies complex legal terminology into plain language,

SENTIMENT ANALYSIS

The image displays a web application interface for sentiment analysis, specifically using the Bert model. The application shows two distinct feedback entries, each with a rating and a detailed description.

Feedback 1: Mahika Agrawal, Real Estate Law, Rating: ★ 3.0, Experience: 13 years, Avg Charge: 216.06 USD/hour. The user rating is 4, and the API response shows a sentiment analysis rating of 5. The user's detailed description is "I loved the services".

Feedback 2: Mahika Agrawal, Real Estate Law, Rating: ★ 3.0, Experience: 13 years, Avg Charge: 216.06 USD/hour. The user rating is 2, and the API response shows a sentiment analysis rating of 1. The user's detailed description is "The services were really bad, I didn't like it at all".

The application includes a sidebar with navigation links: Home, Find Lawyer, History, Settings, and Help. A top bar features a "Ask AI" button and a logo for "LEGALNINJA". The browser's developer tools are open, showing the "Console" tab with several JavaScript errors and logs related to the application's functionality.

Using Bert

PDF SUMMARIZER

Page 1 Summary:

In the case of Anagnostakis v. Greece, the European Court of Human Rights found that there had been a violation of Article 8 (right to respect for priv

Page 2 Summary:

The case discussed in the text is about a woman named Karen who was fired from her job after she took time off to care for her sick daughter. Karen fil

Page 3 Summary:

The case of Anagnostakis v. Greece concerned the non-enforcement of court decisions setting the contact schedule between the applicant, a Greek nationa

Waiting for 1 minute...

Page 4 Summary:

The case of Anagnostakis v. Greece involves a father, born in 1983, who voluntarily recognised his son born out of wedlock with the consent of the moth

Page 5 Summary:

The case concerns a father's appeal against the decision of the Athens Court of Appeal which dismissed his appeal for shared custody of his son with th

Waiting for 1 minute...

Output is truncated. View as a [scrollable element](#) or open in a [text editor](#). Adjust cell output settings...

Pdf Plumber

```
response2 = openai.Completion.create(  
    model="text-davinci-003",  
    prompt=prompt2,  
    temperature=0.75,  
    max_tokens=500,  
    top_p=0.75,  
    frequency_penalty=0.6,  
    presence_penalty=0.8  
)  
  
for s in str(response2.choices[0].text).split('\n'):|  
    print(s.strip())  
[4]  
...  
  
Winning Party: Anagnostakis  
Losing Party: Greece  
Time Taken from start of closing of case: 2018-2021  
Type of Law: Human Rights  
Which Court: European Court of Human Rights
```

Open Ai

ASK AI CHATBOT

localhost:8501

Relaunch to update :

Deploy :

LegalNinja Chatbot

A Chatbot Powered by OpenAI LLM for overall education on Legal Knowledge

How can I help you?

What is pro bono in legal terms?

In legal terms, "pro bono" refers to providing legal services without charge or at a reduced fee to clients who cannot afford to pay for such services. It is derived from the Latin phrase "pro bono publico," which means "for the public good." Pro bono work is typically done to promote access to justice and help individuals or organizations in need. Lawyers and legal professionals often provide pro bono services to disadvantaged clients, charitable organizations, or public interest causes.

Your message ➤

LangChain-Chat with search

Hi, I'm a chatbot who can search the web. How can I help you?

What law does a divorce come under?

✓ Search: What law governs divorce?
I should search for the law that governs divorce.
Action: Search
Action Input: "What law governs divorce?"
Divorce law covers a wide range of legal issues that begin when a couple decides to divorce and end when the judge hands them the divorce decree and judgment of dissolution. Divorce is a long, never-pleasant procedure that requires patience, understanding, and often a good family lawyer. yourselves. Even then, having an attorney review your ... Family law in the United States encompasses a wide range of legal issues that pertain to relationships and domestic matters. One of the most prevalent and significant areas of family law is divorce. Divorce law governs the dissolution of marriages and the resolution of related issues such as child custody, support, and property division. Laws around divorce and custody matters can vary widely. When and how you may file for divorce can differ from one state to the next. Understanding your state's divorce laws and parental rights is important before going ahead. Learn About Your Divorce and Child Custody State Laws Contemplating a divorce is a major decision. Divorce: Divorce is the legal termination of a marriage. Family law addresses issues related to property division, alimony or spousal support, child custody, visitation rights, and child support. Child Custody and Support: When parents separate or divorce, family law helps determine the custody arrangement for their children. Divorce Law Center It is a reality of modern life that not every marriage lasts forever. The divorce rate has been steadily rising throughout the United States for decades. If you are considering ending your marriage, you should learn something about the process in advance. You may file for divorce only in a state where you are a resident.

✓ Complete!
I now know the final answer.
Final Answer: Divorce law governs the dissolution of marriages and the resolution of related issues such as child custody, support, and property division.

Divorce law governs the dissolution of marriages and the resolution of related issues such as child custody, support, and property division.

What is pro bono? ➤

Business Model

Key Partners:

- **Legal Databases:** Collaborate with legal databases and organizations to access accurate and up-to-date information.
- **Law Firms:** Establish partnerships with law firms to offer them promotional opportunities and facilitate connections with potential clients.



Subscription Model: Charge lawyers a recurring fee of Rs. 399/month for premium access to advanced features, such as detailed sentiment analysis and unlimited PDF summarization.



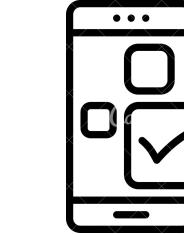
Transaction Fees: Collect 5% of fees when users hire recommended lawyers through the platform



Advertising and Promotion: Offer lawyers and law firms paid promotional options for increased visibility on the platform.

FUTURE SCOPE

Make an app and also include non binary gender so they can also get a platform to explore more job opportunities



Incorporating multilingual capabilities to serve a global audience and break down language barriers, allowing users to access legal information and explanations in their preferred language.



Docker enables seamless scalability by effortlessly replicating and deploying containerized applications across server clusters - Scalability





COST STRUCTURE



- 1. Development Costs:** Invest in the initial development and ongoing improvement of the platform and algorithms.
- 2. Data Maintenance:** Incur costs for continuously updating and validating the lawyer database.
- 3. Marketing and Promotion:** Allocate resources to promote the platform and attract users.
- 4. Legal Expertise:** Employ legal experts to ensure the quality and accuracy of recommendations.





MARKETING CHANNELS



- 1. Online Platform:** Deliver the service through a web application and mobile app.
- 2. Social Media and Marketing:** Use digital marketing, SEO, and social media to reach and engage with potential users.
- 3. Legal Directories:** Collaborate with legal directories and organizations to extend the reach of the platform.



TENSIONFLOW

