My Lawyer : Digitalized Law System

Mr. Atharv Phadnis
Department Of Computer Science and Engineering
Savitribai Phule Pune University, Pune, India
atharv.phadnis.scoe@gmail.com

Mr. Shubham Raut
Department Of Computer Science and Engineering
Savitribai Phule Pune University, Pune, India
rautshubham28@gmail.com

Mr. Kunal Patil
Department Of Computer Science and Engineering
Savitribai Phule Pune University, Pune, India
kunalpatil0409@gmail.com

Mr. Sarvesh Katneshwarkar
Department Of Computer Science and Engineering
Savitribai Phule Pune University, Pune, India
sarveshkatneshwarkar17@gmail.com

Prof. Jyoti B. Kulkarni
Department Of Computer Science and Engineering
Savitribai Phule Pune University,
Pune, India
jbkulkarni.scoe@sinhgad.edu

***Abstract*—As we see Indian Law system is way too complex and almost every citizen gets confused about what legal procedure should be followed at certain situations and circumstances in social life. For e.g.: - A person starting a new business is confused and not aware about the no. of licenses he should pass, where to register his business, how to certify his business and other legal matters related to business. Also due to lack of literacy in laws, often many government officials, lawyers and any related stake holders misguide and thug common citizens. The system would document the entire "Constitution of India", our Law structure, various norms, etc. The user just needs to enter his/her problem in normal human language and the system would correctly provide the user with the entire law procedure it should follow along with essential government helplines, addresses and concerned stake holders. User can verify any law used to threaten him/her or tell any concerned person about the correct law. This would increase the social awareness among the society and help people through all legal matters from renting a house document to registration of new property, issuing legal notices to understanding legal notices, increasing legal contacts to studying legal law structure.**

I. INTRODUCTION

Law System is always consisting of complex and tedious activities that often pile up and result in delayed proceedings and deadlines of tasks. The current Indian Law system consists of lengthy paper work, multiple legal offices visits and delayed timelines are a daily routine. Digitalizing this entire system would fairly boost up its work efficiency and would ease the law process for the common citizens.

Today we have digitalized solutions for every sector of social structure like Finance, Banking, Education, Tourism, etc. but still there doesn't exist any such digitalized solution in the legal sector. This motivated us to create a system that would ease the efforts needed in legal matters and provide digitalized, in-place services to the users related to handle their legal matters.

The system uses supervised machine learning algorithms like KNN and SVM to classify the given user query into various domain in the general classification phase and further categorize it into domain specific case studies in the specialised classification phase.

The system consists of various modules like Online Legal assistance, Online Petition filling, Online License generation, Legal Template generator and Communication Modules. Each of these modules serve for specific requirements by the end users. Online legal assistance works by classifying the user query and accordingly identify related case studies to provide legal advice to the user. Online License generation module, generates digital licenses by verifying the documents uploaded by the user in the system and loading the same on government severs. Similarly, Legal Template generator, encrypts user data on legal templates and facilitate the user with the same. Online Petition module helps user to issue petition to various courts of the country easily and the communication module helps user connect with required law related stake-holders based on the legal matter of the user.

The end users of the system are the common people having less knowledge of law & order would be greatly benefited by this project as it will enable them to get legal assistance & knowledge about trivial to serious legal issues without any explicit help of a law practitioner. Issues like filling applications, filing petitions, etc. can be easily carried out with this system. Users can also contact law practitioners if needed, directly through the system. This system will also provide lawyers with *"virtual offices"*, where they can carry out their work, interact with normal users requiring their assistance/service through the system itself.

II. LITERATURE SURVEY

The main task of the design is to identify valid and most interesting features from the given query and to use them as parameters to classify, categorize and interpret into valid notes. And for this, research works on supervised learning algorithms like KNN, SVM, NLP, etc. were studied well.

The Nishant Jain and GauravGoel work [1] , presents how the AI technologies can interact with the end-user using a chatbot, help them understand the situation and provide guidance by searching millions of solved cases. As Every case is unique, there are some similarities found in many solved cases. To find that similar case based on the input, the proposed system uses KNN based approach and then generate interrogative questions to gather more insight of the situation.

The Amin Sleimi, Marcello Ceci, MehrdadSabetzadeh, Lionel Briand and John Dannwork[2], presented an approach to automatically recommend templates for legal requirements

based on legal statements. To do this, they first defined a set of templates are grouped into three categories: statements with no counterpart, statements with a correlative statement, and statement with an implied statement. Which then devised, using NLP, automated rules for recommending suitable requirements templates.

Sebastian Blank and Florian Wilhelm's work[3], proposes a method to query unstructured databases using sentences written in natural language.It emphasizes on NoSQLdatabases.It makes use of Deep learning models to provide a solution for converting natural language sentences to database queries.

HananeBais, Mustapha Machkour and LahcenKoutti's work[4], puts forth a method for querying structured databasesusing natural language sentences.It is designed to work for SQL databases.A series of steps are performed on the input sentence query(in natural language) to produce relevant SQL query. The steps include – morphological analysis, syntactic analysis, semantic analysis, IXLQ (Intermediate XML Logical Query), DB query generation. It makes use of various tools and concepts like dictionary, knowledge base, semantic rules.

Riya Sil, Abhishek Roy, Bharat Bhushan and A.K. Mazumdar's work[5] , proposes an approach to use artificial intelligence to predict outcomes of legal matters based on legal data trained on the model. It initiates a discussion on how various artificial intelligence algorithms, let it be supervised, unsupervised, reinforcement machine learning algorithms can be used to achieve accurate predictability of legal outcomes to any legal matter. It predicts that using this approach could automate nearly 100000 jobs in the field of law.

Rishi Chhatwal, Peter Gronvall, Robert Keeling and Dr. Jianping Zhang' work [6], propose the concept of explainable predictive coding and simple snippets within responsive documents to summarize legal papers/documents.Research in Explainable AI has focused on two areas of explainable machine learning: Model based explanation (using decision trees and if – then rules) and prediction-based explanation (using SVM, Multilayer Neural Network, etc.). The purpose of this paper was to demonstrate the feasibility of explainable predictive coding in the context of professional services in the legal space.

K. Nikolskaia and V. Naumov in their work[7] stated that, Artificial intelligence technologies are increasingly penetrating into various spheres of human life. Modern artificial intelligence algorithms are not capable of fully imitating legal thinking. However, artificial intelligence algorithms are capable of automating some of the routine work of lawyers. This allows professionals to deal with really difficult cases without being distracted by routine complaints. The paper discusses the use of artificial intelligence in the right.

Guiying Yang's work[8] stated that, the current legal information retrieval methods, keywords are used for extraction and neural networks are used for search. However, the keywords of cases in legal information are different. Simply using keywords to extract the search results in insufficient accuracy. Therefore, an adaptive retrieval method of legal information based on artificial intelligence is proposed. First, establish a positional legal language model to determine the total number of spreads in the term. According to the definition of the term in the model, calculate the similarity of the semantic concept of the term. After confirming that the concept is similar to the term, calculate the co-occurring words in the term. Finally, an artificial intelligence algorithm is used to retrieve the co-occurring words. In order to verify the feasibility of the designed retrieval method and design methods were used to retrieve the contents of a certain legal information database. The experimental results show that the designed retrieval method has a higher retrieval accuracy and meets the design requirements.

In Xu Jing and Song Xiuhui's work[9] it was identified that, with the rapid development of artificial intelligence (AI), it has gradually integrated into people's lives, and uses its own advantages to help people do things that cannot be done. It seems that AI will continue to develop more comprehensively in the future with more functions and technologies to help people and take some occupations.

Dipankar Chakrabarti, Neelam Patodia, Indranil Mitra and their colleagues in their work[10] cited that, Assessing risk for voluminous legal documents such as request for proposal, contracts is tedious and error prone. They have developed "risk-o-meter", a framework, based on machine learning and natural language processing to review and assess risks of any legal document. Their framework uses Paragraph Vector, an unsupervised model to generate vector representation of text. This enables the framework to learn contextual relations of legal terms and generate sensible context aware embedding. The framework then feeds the vector space into a supervised classification algorithm to predict whether a paragraph belongs to a pre-defined risk category or not. The framework thus extracts risk prone paragraphs. This technique efficiently overcomes the limitations of keyword-based search.

Theofanis Aravanis, Konstantinos Demiris, Pavlos Peppas's work [11] states that, Answer Set Programming is a declarative problem-solving approach, initially tailored to modelling problems in the area of Knowledge Representation and Reasoning. In this article, we provide a knowledge-based system, capable of representing and reasoning about legal knowledge in the context of Answer Set Programming - thus, modelling non-monotonicity that is inherent in legal arguments. The work, although limited to a specific indicative domain, namely, university regulations, has a variety of extensions. The overall approach constitutes a representative implementation of the Answer Set Programming's modelling methodology, as well as an enhancing of the bond between Artificial Intelligence and Legal Science, bringing us a step closer to a successful development of an automated legal reasoning system for real-world applications.

Ruining Zhan in his work[12] stated that, Information retrieval technology involves several aspects: such as database management, natural language processing, machine learning, and so on. Furthermore, basic technologies in information retrieval contain Web page downloading, document processing, and file indexing. The proposed intelligent economic law case retrieval algorithm is made up of four steps: 1) Feature extraction, 2) Feature integration, 3) Estimating weights for similarity computation, and 4) Economic law case matching. Particularly, particle swarm optimization is exploited to compute weights for similarity computation. Finally, experimental results demonstrate that the proposed system is able to achieve high quality economic law case retrieval results.

III. PROPOSED METHODOLOGY

A. System Architecture :

According to fig.1 as stated below, on the client side the user authentication system is used to authenticate the user with the

help of valid user-id and password. Each user has been assigned an independent account through which they can access the user interface. This user interface can be used to write the necessary queries to the system.

Further these queries are accepted on the server side. Over here using NLP, the most interesting features from the received query are identified and used as input to the supervised learning algorithm i.e. KNN. Then KNN, with the help of these features and available dataset, classifies the given query into various domains like Business, Corporate, Personal, etc. This Phase is known as 'General Classification Phase'.

Further, the same classification methodology is repeated but inside the respective domain. This is known as 'Specialised Classification Phase'. In this phase, the most relevant past case studies are identified using KNN and are used to generate legal summary/advice.

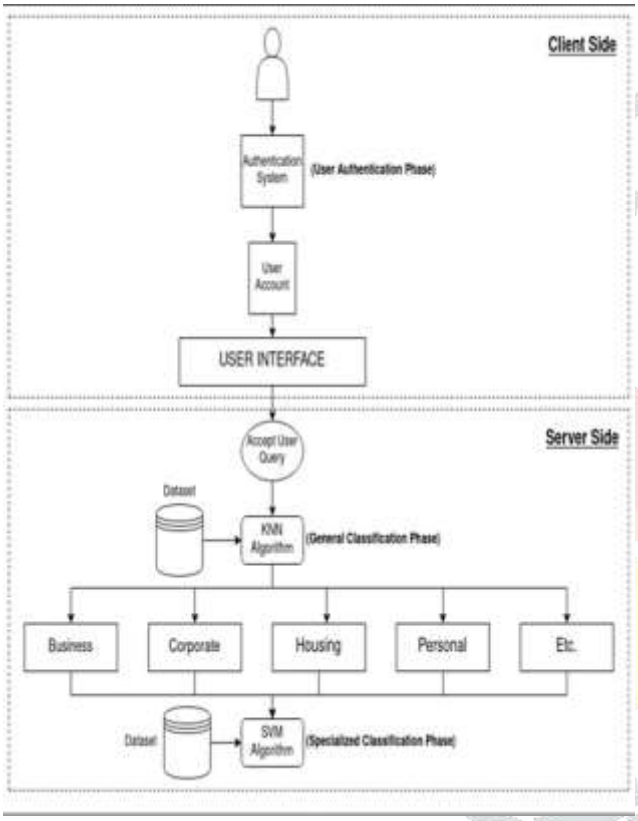


Fig. 1. General System Architecture

In fig.2, detailed structure of each intrinsic module has been specified. The various modules include the Online License Generation, Online Petition Filing, Online Legal Assistance, Communication and Template Generation Module.

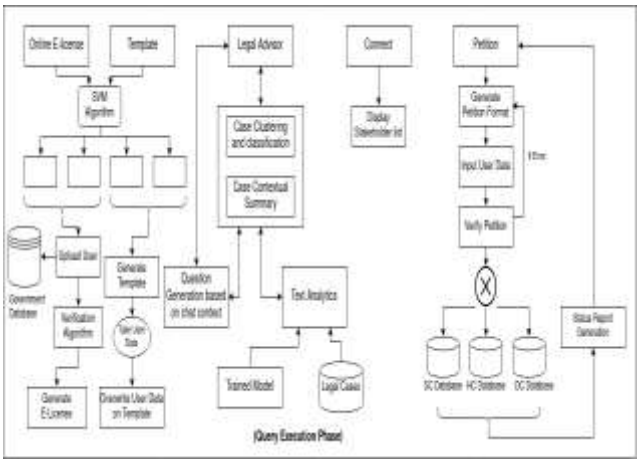


Fig.2. Detailed Modules Description

B. Mathematical Model :

a) Model for Sector Classification

$$SC = \{I, O, F, T\}$$

Where,

$$I = \{i1\}$$

I is the user query

$$O = \{Business, Corporate, Housing, Personal\}$$

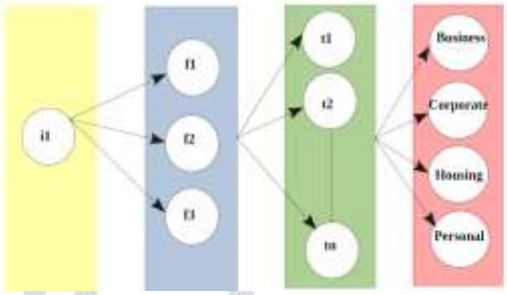
O is the set of classification outputs

$$F = \{f1, f2, f3\}$$

f1= function to remove stopwords f2= function to remove punctuation f3= function to generate word vectors.

$$T = \{t1, t2, t3, ...tn\}$$

T is the collection of labelled sectors to train the model for Sector Classification.



Input Functions Training Model Outputs

Fig.3. Model of sector classification

b) Model for Category Classification

$$CC = \{I, O, F, T\}$$

Where,

$$I = \{i1, i2, i3, ...in\}$$

I is the user query

$$O = \{o1 , o2 , o3 , o4 , o5\}$$

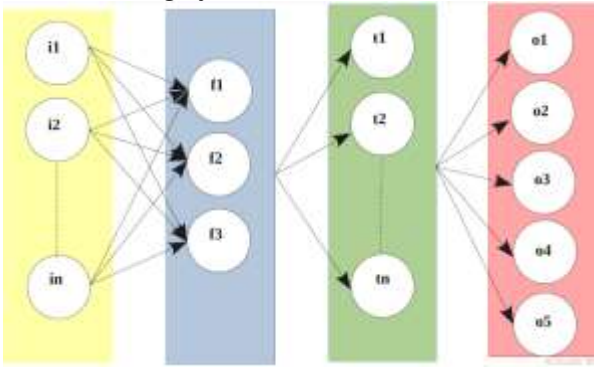
O1 = Online E-license
O2 = Template
O3 = Petition
O4 = Connection
O5 = Legal Advisor

$$F = \{f1, f2, f3\}$$

f1= function to remove stopwords
f2= function to remove punctuation
f3= function to generate word vectors.

$$T = \{t1, t2, t3, ...tn\}$$

T is the collection of labelled sectors to train the model for Category Classification.



Input Functions Training Model Output

Fig.4. Model for Classification category

c) Model for Template Recommendation

TR ={I, O, F, T}

Where,

I={i1, i2}

i1 = Sector classified in 1st model

I2 = user query

O={o1}

O = Recommended template

F={f1}

f1= function to overwrite user data on recommended template

T={t1, t2, t3, ...tn}

T is the collection of labelled templates to train the model for template recommendation.

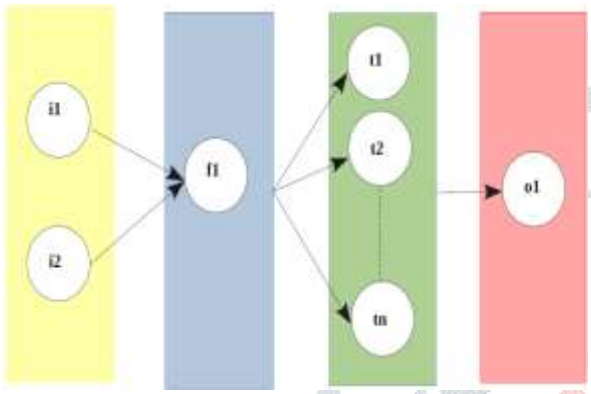


Fig.5. Model for Template Recommendation

d) Model for Online E-license

OL ={I, O, F, T}

Where,

I={i1, i2, i3}

i1 = Sector classified in 1st model

I2 = user query

I3 = user personal data

O={o1}

O = Generated E-license

F={f1, f2, f3}

f1 = function to verify data entered by the user

f2 = function to check for what type of license the user wants

f3 = function to generate E-license by overwriting the user data

T={t1, t2, t3, ...tn}

T is the collection of labelled templates to train the model for Online E-license

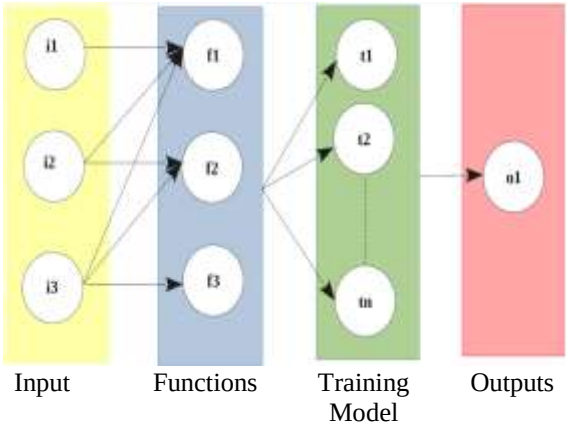


Fig.6. Model for Online E-License

e) Model for Filing Petition

FP ={I, O, F}

Where,

I={i1, i2}

i1 = Sector classified in 1st model

I2 = user query

O={o1}

O = Filed Petition

F={1,f2}

f1= function to generate Petition Format

f2= function to recommend similar Petition Filed

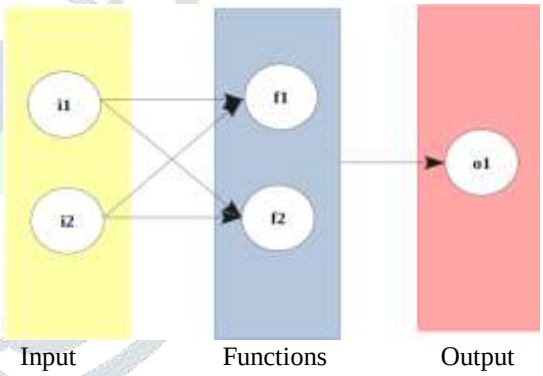


Fig.7. Model for Filing Petition

C. Algorithms:

a) NLP :

Natural Language Processing or NLP is a field of Artificial Intelligence that gives the machines the ability to read, understand and derive meaning from human languages.

It is broadly defined as the automatic manipulation of natural language, like speech and text, by software.

In our system helps to identify most interesting features that would help us as input to classification algorithms like KNN.

b) KNN :

K Nearest Neighbors, is one of the simplest Supervised Machine learning Algorithm.

It Classifies a data point based on how its neighbors are classified.

K in KNN is a parameter that refers to the number of nearest neighbors to include in the majority voting process.

A data point is classified by majority votes from its K nearest neighbors.

There are various applications of KNN in data mining like economic forecasting, search for semantically similar documents, predict whether a person will have diabetes or not etc.

The main purpose of KNN in our application is classifying the user query into appropriate domains and further identifying K-Nearest case studies related to given query.

IV. CONCLUSION

As we see Indian Law System is way too complex and almost every citizen gets confused about what legal procedure should be followed at certain situations and circumstances in social life. In this situation, *My Lawyer* would be a pioneer step to digitalize entire law system and help common citizen of India to understand and take legal steps confidently. It would also help increase Indian Law literacy among Indian citizens.

Such systems would improve the existing efficiency of Indian Law System and would further motivate the society to implement such solutions in various other domains.

REFERENCES

[1] *Nishant Jain and GauravGoel, "An Approach to Get Legal Assistance Using Artificial Intelligence" 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) , June 2020*

[2] Amin Sleimi, Marcello Ceci, MehrdadSabetzadeh, Lionel Briand, John Dann, "Automated Recommendation of Templates for Legal Requirements", IEEE 28th International Requirements Engineering Conference (RE), September 2020

[3] Sebastian Blank and Florian Wilhelm's work, "Querying NoSQL with Deep Learning to Answer Natural Language Questions", Thirty-Third AAAI Conference on Artificial Intelligence At: Honolulu, Hawaii, USAVolume: Vol 33 (2019)

[4] HananeBais, Mustapha Machkour and LahcenKoutti's work "Querying Database using a universal Natural Language Interface Based on Machine Learning", International Conference on Information Technology for Organizations Development (IT4OD), 2016.

[5] RiyaSil, Abhishek Roy, BharatBhushan, A.K. Mazumdar, "Artificial Intelligence and Machine Learning based Legal Application: The State-of-the-Art and Future Research Trends", International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2019.

[6] Rishi Chattwal, Peter Gronvall, Robert Keeling, Dr.Jianping Zhang, Dr.Haozhen Zhao, "Explainable Text Classification in Legal Document Review A Case Study of Explainable Predictive Coding", IEEE International Conference on Big Data (Big Data), 2018

[7] K. Nikolskaia and V. Naumov,"Artificial Intelligence in Law", International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), 2020

[8] Guiying Yang, "Adaptive Retrieval Method of Legal Information based on Artificial Intelligence", IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI), 2020

[9] Xu Jing; Song Xiuhui, "Artificial Intelligence Legal Consulting Based on Categorization", International Conference on Artificial Intelligence and Education (ICAIE), 2020

[10] Dipankar Chakrabarti; Neelam Patodia; Udayan Bhattacharya, "Use of Artificial Intelligence to Analyse Risk in Legal Documents for a Better Decision Support, TENCON 2018 - 2018 IEEE Region 10 Conference

[11] Theofanis Aravanis; Konstantinos Demiris, "Legal Reasoning in Answer Set Programming", IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI), 2018

[12] Ruining Zhan, "Design and Development of Intelligent Economic Law Case Retrieval System", International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), 2018