



# Artificial Intelligence in The Judicial System



# Data

Collecting relevant data and analyzing it is an important aspect of research and machine learning.

There are various sources on the internet for legal data, paid and unpaid. The following slides show what all data I collected and from where.



# Indiankanoon.org

Kanoon is a free legal search engine providing access to Indian laws, court judgments and tribunal judgments. Kanoon has been developed by computer scientist Sushant Sinha. The site contains over 1.4 million laws along with judgements from the Supreme Court of India, 24 high courts and 17 law tribunals. Also provided are Constituent Assembly debates and Law Commission reports. The aim of Kanoon is to make Indian law accessible to ordinary people by integrating legislation and case law and by breaking legal documents down into clauses so that relevant information can be more easily located.

<http://www.indiankanoon.org/>



# Ways used to extract data from indiankanoon

- Manual Search - Judgements can be filtered on the basis of
  - Court - Supreme court, High Court, District Court
  - Year of case
  - Month of case
  - Author of the article
- API - There are four kind of API requests that can be accessed are available at the following end points:
  - Search query - <https://api.indiankanoon.org/search/?formInput=<query>&pagenum=<pagenum>>
  - Document - <https://api.indiankanoon.org/doc/<docid>/>
  - Document fragments - <https://api.indiankanoon.org/docfragment/<docid>/?formInput=<query>>
  - Document Metainfo - <https://api.indiankanoon.org/docmeta/<docid>/>



# SCC Online

SCC Online includes the Supreme Court Cases report series (SCC) from 1969 - present, up to date unreported judgments from the Supreme Court and all High Courts, Tribunals and Commissions, and many High Court historic and recent report series such as the Indian Law Reports . SCC also includes useful case digests and articles.

SCC online also lets users download true copy of the documents.

SCC online is a paid feature and I used a free trial. The free trial lasts for about a week or 4 hours of browsing on the website, whichever comes first.

There are 2 ways to browse the website -

- ☐ Keyword searching
- ☐ Using dropdowns and filters

According to me, using keywords is handy when one knows what exactly they are looking for. The second feature is useful if one has a general idea of what is required but is not sure specifically what they are looking for.

<https://www.sconline.com>



## Bombay High Court Website

- ❖ Cases can be searched by filtering a period of one year or giving the exact case number
- ❖ Listing dates for each case are given and a document of what orders or judgement was given on a particular date can be downloaded
- ❖ Lower court details for the same case can be obtained in a short table type summary
- ❖ Recent judgements given by the Bombay HC are on the website
- ❖ <https://bombayhighcourt.nic.in/index.php>



## Other sources

### National Judicial Data Grid:

NJDG is a national repository of data relating to cases pending and disposed of in all district and taluka courts of the country. The portal has been developed around the concept of elastic search technology enabling efficient case management and monitoring of cases leading to effective disposal of cases.

### Indian Legal Data Corpus:

ILDC is a large corpus of 35k Indian Supreme Court cases annotated with original court decisions. A portion of the corpus (a separate test set) is annotated with gold standard explanations by legal experts. The dataset is prepared by scraping data from indiankanoon.org

It is a project which was undertaken by students from IITK, IISER-K and WBNUJS. The dataset can be obtained by requesting them access by filling this form - <https://forms.gle/psRcqPqbTaynRu6u5>



# Models tested

- I. Sent2Vec model from the paper - ILDC for CJPE
- II. Summarization models
  - A. Supervised
  - B. Unsupervised





# Sent2Vec

- Sent2Vec presents a simple but efficient unsupervised objective to train distributed representations of sentences.
- 3 models - Logistic Regression, SVM and Random Forest were applied to the mapped sentence vectors. The models were designed to predict an explainable outcome of a case.
- Dataset used was the ILDC.
- The highest accuracy achieved was by Logistic Regression at 55.44%.



# Summarization Models

- ★ The first goal of the research and this project is to supplement the judges and the judicial system, not supplant it by developing a highly accuracy case prediction model
- ★ Hence, research was shifted to implementing summarization models to help lawyers.
- ★ Precedents help a lawyer understand how the Court has dealt with similar scenarios in the past, and prepare the legal reasoning accordingly. Hence lawyers have to go through hundreds of prior cases. These cases are available as case judgments which are essentially long and free-flowing with dense legal text. Hence summarization of these documents helps lawyers skim through a large amount of documents in a relatively shorter time.
- ★ The goal of the model is to summarize the document in such a way that important and essential information is retained.
- ★ Two models were implemented by me - Supervised and Unsupervised using the paper 'A Comparative Study of Summarization Algorithms Applied to Legal Case Judgments' as reference
- ★ The Supervised model gives better ROUGE scores in comparison to the unsupervised model.

# An Example

## Input

1. Heard the learned Counsel for the appellants and the respondents at a considerable length.

2. Perused the record.

3. The Appeals from Order can be disposed of in following terms in view of the peculiar circumstances of the case.

5-92265-2020-AOst-order=.doc

4. The operative part of the impugned order dated 20th February, 2020 passed by the learned Joint Civil Judge, Senior Division, Pune is defective in the sense that at the interlocutory stage, the relief as has been granted is as good as a final relief in the suit. As such, the said order is set aside and substituted as follows :-

### ORDER

(a) The appellants – original defendants shall in view of its standing committee’s Resolution No.62 dated 22nd June, 2020 which came to be passed after the impugned order, accept the supply of uniforms of the students for the Academic Year 2020-21.

(b) The learned Civil Judge, Senior Division, Pune shall expedite the suit.

(c) Parties shall co-operative with the expeditious disposal of the suit without seeking unnecessary adjournments.

5. The Appeals stand disposed of in the aforesaid terms. No order as to costs.

6. In view of disposal of the appeals, pending Interim Applications are also disposed of.

(PRITHVIRAJ K. CHAVAN, J.)

## Output

1.2.3.The Appeals from Order can be disposed of in following terms in view of the peculiar circumstances of the case.doc

4.As such, the said order is set aside and substituted as follows :-

ORDER

(a) The appellants – original defendants shall in view of its standing committee’s Resolution No.

(b) The learned Civil Judge, Senior Division, Pune shall expedite the suit.

5.No order as to costs.6.



# Obstacles

I experienced a few obstacles such as:

1. Non availability of entire case documents  
- Even after looking up at a lot of sources to find legal data, documents which contained details about the entire case were not found. I could find only judgements for the most part.
2. Manupatra and westlaw are another paid source for legal data which offer free trial. I signed up for a free trial request with westlaw but they never reverted back to me. As for Manupatra while signing up on their portal, I would never receive a One time password (OTP) from their end.



# Bibliography

India's Court Hierarchy - <https://unimelb.libguides.com/c.php?g=928009&p=6704295>

NJDG - <https://njdg.ecourts.gov.in/njdgnew/index.php>

Indiankanoon API documentation - <https://api.indiankanoon.org/documentation/>

ILDC for CJPE - <https://github.com/Exploration-Lab/CJPE/tree/main/Models/Classical%20Models>

Case summarization -

[https://www.researchgate.net/publication/332256965 A Comparative Study of Summarization Algorithms Applied to Legal Case Judgments](https://www.researchgate.net/publication/332256965_A_Comparative_Study_of_Summarization_Algorithms_Applied_to_Legal_Case_Judgments)

<https://aclanthology.org/C16-2054.pdf>