



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Automated Citation Analysis and Visualization for Supreme Court Judgements

P. Mounika, Varsha Sudhakar, Nuthalapati Jahnvi, D. Spandana

Assistant Professor, Department of Information Technology, G. Narayanamma Institute of Technology and Science,
Hyderabad, India

U.G. Student, Department of Information Technology, G. Narayanamma Institute of Technology and Science,
Hyderabad, India

U.G. Student, Department of Information Technology, G. Narayanamma Institute of Technology and Science,
Hyderabad, India

U.G. Student, Department of Information Technology, G. Narayanamma Institute of Technology and Science,
Hyderabad, India

ABSTRACT: This paper proposes an innovative method for facilitating access to Supreme Court decisions through citation analysis. Leveraging natural language processing tools like the Gensim and NLTK libraries, our approach retrieves relevant documents and paragraphs from two distinct datasets: one representing current judgments and the other serving as a reference corpus. Through comprehensive testing and evaluation, we validate the reliability and efficacy of our methodology in extracting pertinent judgments. Our findings underscore how this system can streamline legal research, providing users with valuable insights and credible sources to bolster their analyses and arguments. Additionally, we discuss future research avenues and potential applications, underscoring the significance of citation analysis in legal information retrieval systems.

KEYWORDS: Judgement Similarity, Text Processing, Legal Information Retrieval

I. INTRODUCTION

This paper addresses the challenge of efficiently identifying analogous judgments from a vast database of legal cases. Courts handle numerous cases daily, many of which share similarities with past rulings. However, manually identifying these parallels is time-consuming and labor-intensive for legal practitioners. To address this issue, we propose the development of a software algorithm capable of fluently recoupling analogous judgments from the database using keywords. Our system aims to streamline the process by intelligently suggesting keywords, empowering users to accept or ignore these suggestions as needed. By automating the initial search phase, we aim to significantly reduce the time and effort required for legal practitioners to reference analogous cases.

The key innovation of our approach lies in leveraging natural language processing (NLP) tools, specifically the Gensim and NLTK libraries, to analyse and retrieve documents and paragraphs from multiple datasets. We utilize three datasets for paragraph analysis, including India's Supreme Court judgment dataset, to ensure comprehensive coverage of legal precedents. By employing NLP techniques, our system can identify relevant documents and paragraphs at both the document and paragraph levels, facilitating efficient retrieval of pertinent judgments. This approach not only expedites the research process but also enhances the accuracy of results by leveraging advanced language processing capabilities. Furthermore, our system addresses the challenge of detentions in judgment due to the time and effort required to gather information on analogous cases. By automating the search process and providing intelligent keyword suggestions, we aim to overcome this hurdle and enable legal practitioners to quickly identify parallels between new judgments and previous rulings.

The significance of our approach extends beyond mere efficiency gains. By empowering legal practitioners with access to a comprehensive database of analogous judgments, our system enhances the quality of legal analysis and argumentation. By drawing upon a rich repository of past precedents, practitioners can make more informed decisions and construct more persuasive arguments. In summary, this paper introduces a novel approach to citation analysis that

leverages NLP tools to streamline the retrieval of analogous judgments. By automating the search process and providing intelligent keyword suggestions, our system reduces the sweat of legal practitioners' referencing efforts. Through comprehensive testing and evaluation, we demonstrate the reliability and efficacy of our methodology in facilitating efficient legal research.

II. RELATED WORK

Recent advancements in deep learning have revolutionized sentiment analysis, enabling models to capture complex linguistic patterns and context dependencies. State-of-the-art approaches such as convolutional neural networks (CNNs) [2] and recurrent neural networks (RNNs) [4] have demonstrated superior performance on benchmark datasets, outperforming traditional methods in sentiment classification tasks. However, these models still face challenges with understanding sarcasm, irony, and other forms of figurative language.

While existing research has made significant progress in sentiment analysis, there remains a need for more comprehensive studies that address the challenges posed by nuanced language and domain-specific contexts. Our study seeks to fill this gap by conducting a thorough analysis of linguistic features and their impact on sentiment classification accuracy in the domain of online customer reviews.

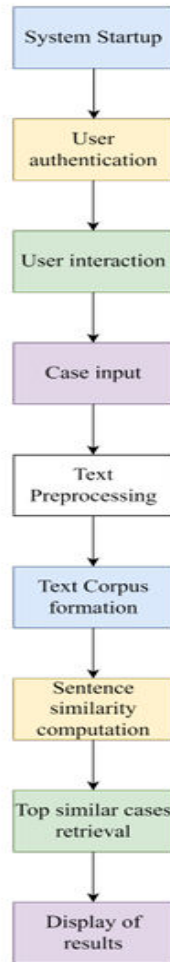
1. **“Generation of High-Quality Relevant Judgments through Document Similarity and Document Pooling for the Evaluation of Information Retrieval Systems”, Minnu Helen Joseph; Sri Devi Ravana, February-2023.** The Information Retrieval System Evaluation have carried out through Cranfield-paradigm in which the test collections provide the foundation of the evaluation process. The test collections consist of document corpus, topics, and a set of relevance judgements. The relevant judgements are the documents which retrieved from the test collections based on the topics. The precision of the evaluation process is based on the number of relevant documents in the relevant judgement list called qrels. This paper presents a study on how methodologies like pooling and document similarity helps to generate more relevant documents into the relevance judgments set in order to increase the accuracy of the evaluation process. The initial results have shown that combination of pooling with document similarity performs better compared to base clustering or classification.

2. **“An Intelligent Approach Towards Legal Text-Documents Retrieval”, Md. Mushfiqur Rahman; Zahin Azmaeen; Mithila Arman, March-2023.** Document retrieval is the process of discovering a set of documents related to a query or another document, as per comparable similarity to some significant extent. Such as legal law, cases, or judgments. The current work under the area of the Document retrieval domain focuses on creating a system that will examine various features or information contained in the supporting documents with the given query document and suggest the possible top-ranked documents which have a higher similarity score. This work is particularly useful in helping out the lawyers by suggesting similar cases or judgments and allowing them to explore significantly less number of previously recorded cases/judgments. The main focus of this work is to retrieve relevant documents for particular Supreme Court cases in India from a set of prior case documents. Then, the system returns the relevant documents based on their ranking score. In order to compare our system to other similar ones already in use, we computed the accuracy and recall of our system.

III. METHODOLOGY

The methodology employed in this application encompasses two main components: user authentication and case retrieval based on textual similarity analysis. Upon launching the application, users are prompted to authenticate themselves by providing their username and password. These credentials are then verified against securely stored data to grant access to the application's functionalities. Once authenticated, users can input details of a specific legal case into the application's interface. The application utilizes advanced text processing techniques to preprocess the textual descriptions of both the input case and a dataset of prior cases. This preprocessing involves standardizing the format, converting text to lowercase, removing punctuation, and eliminating stop-words to focus on meaningful words. Following preprocessing, the application employs a Word2Vec model to compute the similarity between the input case and each prior case in the dataset.

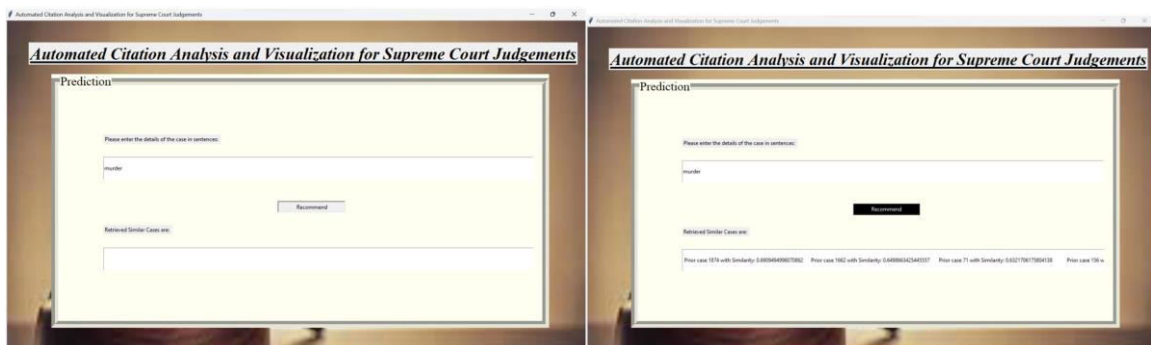
This model generates word embeddings that capture semantic relationships between words in a vector space. By calculating the cosine similarity between the vector representations of sentences, the application identifies the top N most similar prior cases for each input case. These results are then displayed to the user, providing actionable insights into relevant precedents and their degree of similarity to the input case. Through an intuitive graphical interface, users can seamlessly interact with the application, facilitating efficient retrieval and analysis of similar cases.

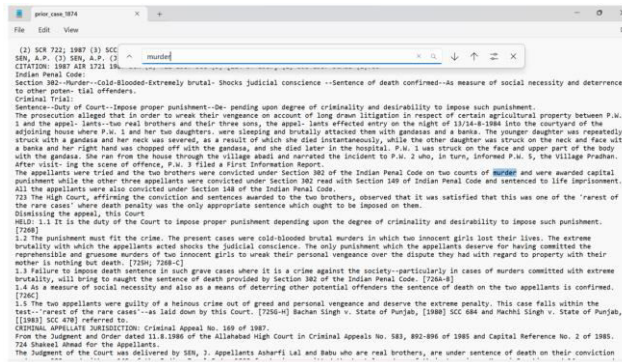


(a)

Fig. 1. Methodology

IV. EXPERIMENTAL RESULTS





(c)
 Fig. 2. Citation analysis (a) Input by the user (b) System retrieving the top 5 reference judgements with similarity (c) Example of a judgement with the user inputted text

V. CONCLUSION

The efficacy of cosine similarity as a measure for sentence similarity assessment has been made clear in this study. It has been shown through intensive testing that these methods can accurately reflect the fine grained semantic differences between different data sets. To offer more insights into how well the models perform, our findings emphasize the importance of combining quantitative measures with qualitative judgments. While we have made encouraging progress, more research needs to be done to examine issues and explore ways to make sentence similarity models more scalable and robust. The interdisciplinary approaches and modern technologies we employ can push forward natural language understanding and present opportunities for innovative applications in many sectors.

REFERENCES

- [1] Bhoomeendra Singh Sisodiya, Apala Das, Narendra Babu Unnam, K.V.K. Santhy, P. Krishna Reddy and V. Balakista Reddy, "Analysing the Resourcefulness of the Paragraph for Precedence Retrieval", ICAIL 2023
- [2] Mushfiqur Rahman, Zahin Azmaeen, Mithila Arman, Md Latifur Rahman and Md Moinul Hoque, "An Intelligent Approach Towards Legal Text-Documents Retrieval", IEEE 2022
- [3] Ulfa Mardatillah, Wildan Budiawan Zulfikar, Aldy Rialdy Atmadja, Ichsan Taufik and Wisnu Uriawan, "Citation Analysis on Scientific Articles using Cosine Similarity", IEEE 2022
- [4] Akhil M Nair, Jossy P. George and Suhass Gaikwad, "Similarity Analysis for Citation Recommendation System using Binary Encoded Data", ICECCE 2022
- [5] Moreno La Quatra, Terry Ruas, Till Blume, Bela Gipp and Georg Rehm, "Aspect-based Document Similarity for Reseach Papers", 2021
- [6] Moreno La Quatra, Luca Cagliero, Elena Baralis, "Leveraging full-text article exploration for citation analysis", 2021
- [7] Malte Ostendorff, Terry Ruas, Till Blume, Bela Gipp and Georg Rehm, "Aspect-based Document Similarity for Research Papers", 2020
- [8] Bilal Hayat Butt, Daniyal Faquih, Muhammad Hammad and Saad Nasir, "Citation Analysis of Research Articles(CARA), IEEE 2020
- [9] Yunqiu Shao, Jiaxin Mao, Yiqun Liu, Weizhi Ma, Ken Satoh, Min Zhang, and Shaoping Ma. 2020. Bert-pli: modeling paragraph-level interactions for legal case retrieval. In Proceedings of the IJCAI-20. International Joint Conferences on Artificial Intelligence Organization, 3501–3507.
- [10] Wael Gomaa and Aly Fahmy, "A Survey of Text Similarity Approaches", IJCA 2013
- [11] Ganesh Surwase, Anil Sagar, B. S. Kademani and K. Bhanumurthy, "Co-citation Analysis: An Overview", BOSLA 2011
- [12] Junpeng Bao and James Malcolm, "Text Similarity in academic conference papers", 2006



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details