**Abstract**

Legal writing like Supreme Court decisions is long, complicated, and riddled with jargon — i.e., they're hard to analyze manually. In this essay, we examine a group of Supreme Court decisions and examine how technology can help to deconstruct it all.

Then we implement Exploratory Data Analysis (EDA) for the purpose of establishing patterns, trends, and hidden meanings. Using libraries like Pandas, Matplotlib, and Seaborn, we pre-clean data, check missing values, examine trends year-over-year, and even identify outliers in court votes. Correlation and covariance are also used when establishing the vote numbers' relations.

Once we know more about the data, we move to the next phase in a leap of grandeur — classifying the legal texts using deep learning. We try out models like LSTM, Bi-LSTM, and CNN to perform auto-predictions like the issue area of a case based on its text. These models enable us to know how machines can interpret legal text and classify documents as such.

As a whole, this project illustrates how the union of solid data analysis and deep learning can make legal research smarter, quicker, and more efficient — reducing human labor and releasing valuable insights from troublesome documents.

**Introduction**

Judicial rulings, being among legal documents, are good sources of information but present colossal challenges since they are complicated, lengthy, and laden with technical domain terminologies. With the rising volume of legal texts being produced each year, it has become important to develop automatic systems capable of assisting in structuring, analyzing, and classifying the documents effectively.

In the recent years, a combination of Deep Learning and Natural Language Processing (NLP) has been highly promising for legal document reading. Such techniques enable systems to grasp semantic relationships, classify cases into fields of issue, and even predict. However, before using such advanced models, one needs to have a wide overview of data through Exploratory Data Analysis (EDA). EDA not only helps in data preprocessing and cleaning but also detects hidden patterns, relationships, and outliers that affect the performance of a model.

The presented research paper here presents a hybrid approach that involves statistical EDA and deep learning-based text classification to comment on U.S. Supreme Court rulings. In this case, we utilize tools like Pandas, Seaborn, and Matplotlib to filter out meaningful facts from the provided dataset. Then, we utilize various deep models like LSTM, Bi-LSTM, and CNN architectures in order to effectively classify legal documents into appropriate issue areas as well as the nature of their decisions.

The contributions of this paper are threefold:

1. A thorough statistical and visual analysis of judicial decision data.

2. Testing and evaluation of different deep learning models for legal document classification.

3. Incorporation of EDA insights to guide and optimize the performance of classification models.

In the process, this research hopes to fill the gap between data-driven legal analysis and computerized legal text classification and lay a foundation for future development of legal AI applications.

**Keywords**

• Legal Document Analysis

• Natural Language Processing (NLP)

• Deep Learning

• Exploratory Data Analysis (EDA)

• Text Classification

• Supreme Court Judgements

• LSTM Networks

Literature Review /Related Work

Analysis of legal documents has become an important field of study with the exponential growth in legal data and the need to make judicial systems more data-driven and efficient. The labor-intensive, time-consuming, and error-prone manual review has been the hallmark of traditional legal research. With the advent of Natural Language Processing (NLP) and Deep Learning techniques, automated legal document classification has become more feasible and efficient.

There are numerous studies on NLP approaches to legal text categorization. Chalkidis et al. (2019) proposed hierarchical attention networks for legal document classification using the European Court of Human Rights (ECHR) dataset and showed significant improvement in the comprehension of long legal texts. Similarly, Aletras et al. (2016) utilized SVMs and NLP to predict court judgments on the basis of textual evidence with great accuracy.

Other deep learning models such as LSTM (Long Short-Term Memory) and BERT have also been attempted for legal document analysis. Zhong et al. (2020) established that BERT-based models are better than normal classifiers for legal reasoning extraction and legal outcome prediction. LSTM models, particularly, have shown potential in sequential data such as legal rulings and votes by capturing context and time-varying patterns.

In addition, Exploratory Data Analysis (EDA) has been a central role in understanding the organization and distribution of legal datasets. Sulea et al. (2017) study emphasized the importance of EDA during preprocessing of legal corpora, trend and outlier detection before modeling.

Despite such breakthroughs, a majority of such works revolve around Western systems of law. Not a great deal of such literature focuses on the Indian Supreme Court dataset. This provides opportunities for regional analysis based on applying state-of-the-art NLP and deep neural network architectures on localized legal materials. Our paper occupies such a niche in covering not only carrying out a thorough EDA on Indian Supreme Court rulings but also incorporating classification practices that may be scaled and trained on allied systems of law.

Related Work

The use of machine learning and natural language processing (NLP) to the legal field has experienced great progress in recent years. Previous studies have mostly concentrated on automating the task of legal document classification, judgment prediction, and information extraction.

Aletras et al. (2016) made one of the first impactful studies in predicting case outcomes with Support Vector Machines (SVM) using textual features drawn from European Court of Human Rights (ECHR) judgments with high accuracy. Their study made a strong case for the applicability of traditional NLP pipelines to legal reasoning.

Chalkidis et al. (2019) built upon this by using hierarchical attention networks on legal text and successfully encapsulating long-range dependencies in structured legal texts. Their approach achieved better performance in multi-label classification scenarios, especially for the detection of law area and prediction of relevant articles.

BERT and subsequent Transformer-based models have also continued to transform legal text processing. Zhong et al. (2020) proved that BERT variants that are fine-tuned performed better than conventional RNN and CNN-based models when predicting legal judgment, which signified the capability of pre-trained language models to grasp legal semantics.

From an Indian point of view, Bhattacharya et al. (2019) studied summarization and citation analysis in the Indian legal system. There have been fewer studies, however, on systematic EDA of Indian legal data to identify voting patterns, time-series case trends, and issue areas.

Although deep learning has been the focus of recent work, Exploratory Data Analysis (EDA) is underexploited in this area. EDA offers critical statistical information that refines feature engineering, identifies biases, and enhances model interpretability—particularly in judicial data where transparency is paramount.

Our work extends these building blocks by unifying a solid EDA pipeline with legal classification problems. Particularly, we aim at the Supreme Court of India's judgments dataset, uniting traditional statistical analysis with visual storytelling and utilizing deep learning (e.g., LSTM) for textual classification. Such unification intends to enhance the interpretability as well as the predictive performance of legal analytics models.

Proposed Methodology

The proposed methodology consists of two central phases: Exploratory Data Analysis (EDA) and Legal Document Classification using Deep Learning. Both the phases are required in revealing the insights and creating a prediction model for legal verdicts.

Phase 1: Exploratory Data Analysis (EDA)

We conduct a rigorous analysis of the dataset to look at its structure, trends, and patterns:

1. Data Collection & Loading

Supreme Court judgement metadata dataset (justice.csv) is loaded and stored as a Pandas DataFrame for processing.

2. Data Cleaning & Preprocessing:

o Impute missing categorical feature values of issue\_area, decision\_type with "Unknown".

o Remove records with null votes.

o Drop duplicates, simple transformations for standardization.

3. Feature Engineering:

o Create a new feature total\_votes by adding majority\_vote and minority\_vote.

4. Descriptive Statistics & Distributions:

o Summary statistics (mean, std, min, max) are calculated to interpret numerical data.

o\tsCategory distributions are investigated using value counts and pie charts.

5. Correlation & Covariance Analysis:

o Heatmap is utilized to present vote-based feature correlation.

o Covariance is computed in order to determine how vote variables correlate with each other.

6. Data Visualization:

o Line plots, box plots, swarm plots, violin plots, and bar charts are utilized to find trends over time, voting patterns by issue area, and identify outliers.

Phase 2: Legal Text Classification Using Deep Learning

Once we have processed the data, we proceed with text classification with Natural Language Processing (NLP) and deep learning methods. The process is as follows:

1.

Text Preprocessing:

no

Text data (for example, summary of judgments or full text) is preprocessed using tokenization, lowercasing, removal of punctuation, and stop word filtering.

no

Word embeddings (for example, Word2Vec or GloVe) are employed to represent text in numeric form.

2. Model Structure: We try out three of the most frequent deep learning models for text classification:

An LSTM (Long Short-Term Memory):

Handles text long-term dependencies. Suitable for sequential understanding in judgments.

A Bi-LSTM (Bidirectional LSTM):

Improves LSTM by reading the input in both forward and backward directions for improved context understanding.

A CNN (Convolutional Neural Network):

Good at extracting local features and patterns from text using filter and pooling layers.

3. Model Training:

Tuned labeled judgment data are utilized for training the models for text labeling in legal topics (e.g., issue areas).

The hyperparameters (e.g., learning rate, batch size, and epochs) are optimized for better performance.

4. Test metrics for the classifier's performance are:

Accuracy, Precision, Recall, F1-score

Multi-model comparative study across LSTM, Bi-LSTM, and CNN decides which among them has been performing better.

Overview of the Methodology

Phase Description

EDA/data structure, trends, and associations are discovered using statistical and visualization techniques.

Modeling

Utilize deep learning to classify legal text into pre-specified categories based on judgment content.

**Summary of the Methodology**

| **Phase** | **Description** |
| --- | --- |
| **EDA** | Understand data structure, patterns, and correlations using statistical and visual tools. |
| **Modeling** | Use deep learning to classify legal text into predefined categories based on judgment content. |

**Experimental Setup**

**This section describes the equipment, dataset, preprocessing steps, and methodologies adopted for performing exploratory data analysis and model-based classification in the legal field.**

**1. Dataset Description**

**The dataset employed in this study is justice.csv and contains judgments and metadata of the Supreme Court of India. The most important fields are:**

**• fac\_len: Factual argument length**

**• tissue\_area: Legal issue area of the case**

**• decision\_type: Decision type (majority, plurality)**

**• \tmajority\_vote: Majority vote count**

**• minority\_vote: Minority vote count**

**This raw data facilitated statistical analysis and exploratory visualization of judicial trends for various years and issue categories.**

**2. Setting and Tools**

**All tests and analyses were carried out on Google Colab to ensure convenience and replicability. The libraries below were used:**

**• Pandas and NumPy: Manipulating and preprocessing the data**

**• Matplotlib and Seaborn: Visualization of trends, distributions, and correlation**

**• Scikit-learn (used conditionally for ML): In case of any potential classification model or metrics**

**3.** **Data Preprocessing**

**Prior to statistical or machine learning processing, the below mentioned preprocessing was carried out:**

**• Handling Missing Values: Categorical nulls were filled in with 'Unknown'; rows where numeric vote values were missing were eliminated.**

**• De-duplication: Maintained data uniqueness for valid case statistics.**

**• Feature Engineering: Introduced a new feature total\_votes = majority\_vote + minority\_vote.**

**4.** **Exploratory Data Analysis (EDA)**

**EDA was conducted to:**

**• Plot the distribution of decisions across year and issue area.**

**•\tAnalyze the correlation and covariance between vote-related variables.**

**•\tDetect outliers in majority vote patterns.**

**•\tCreate boxplots, pie charts, heatmaps, and violin plots for further legal conclusions.**

**5. Optional Modeling Phase**

**Deep learning models like LSTM (Long Short-Term Memory networks) were used for text classification in the longer portion of this project (from another notebook). Tokenized texts of judgments were input into the LSTM after passing through:**

**•\tText cleaning**

**•\tTokenization and padding**

**•\tLabel encoding**

**Hyperparameters were chosen empirically, and model performance was measured in terms of accuracy, precision, recall, and F1-score.**

**Results and Analysis**

**This section presents the findings of the exploratory data analysis done on the Supreme Court judgment dataset and discusses the patterns and conclusions inferred from the statistical and graphical aids.**

**1. Statistical Summary**

**A descriptive statistical analysis gave an initial insight into the dataset:**

**• Widely varying majority votes across cases, and the mean vote count indicated a stable judicial panel size.**

**• Minority votes were relatively less common, reflecting the majority predominance of concord in decision-making.**

**Outliers were identified in distribution analysis by distribution analysis too, representing rare or disputed judgments.**

**2. Decision Type Distribution**

**Decision\_type was pie chart represented to have a majority type of decisions followed by unanimous and plurality types. This reflects the judiciary's inclination towards collective verdicts in most cases in courts.**

**3. Temporal Trends**

**A line graph of the cases by year indicated fluctuating case numbers, some years experiencing judicial workload spikes. These spikes might be indicative of years of social or political upheaval, and further investigation would be in order.**

**4. Issue Area Insights**

**It was found with the use of bar plots and boxplots that:**

**•\\tSome issue areas, i.e., Criminal Procedure and Civil Rights, had the highest number of cases.**

**• These regions also reflected high variability of vote counts, indicating complexity and variation in opinions among justices.**

**5. Correlation and Covariance Analysis**

**The heatmap correlation matrix reflected:**

**• Expected strong positive correlation between majority\_vote and total\_votes.**

**• Moderate negative correlation between majority\_vote and minority\_vote, implying that higher majority is associated with lower dissent.**

**The Covariance analysis confirmed these correlations and also quantified directional dependency of variables as a foundation for follow-up dimension reduction or feature selection in possible additional modeling.**

**6. Vote Pattern Analysis**

**•\tDensity and distribution of majority votes in issue areas were visualized in violin plots.**

**•\tdetailed dispersion within minority votes and that some styles of decisions conjured up stronger opposing opinions in terms of passions than others, were captured with swarm plots.**

**7. Identification of Outliers**

**Outlier analysis with the IQR approach singled out extremely high or low majority vote decisions. These outliers are potentially landmark or extremely controversial decisions and need qualitative legal examination.**

**Findings on Understanding**

**Court decisions, while usually unanimous or majoritarian, differ radically from the character of the legal problem. Visual and statistical indicators signal patterns in ruling-making over time, types of issues, and margins of votes.**

**These findings not only provide a quantitative foundation for research on judicial behavior but also open the door to further modeling, e.g., predictive modeling or case classification with deep learning.**

**Discussion**

**The exploratory data analysis of the Supreme Court judgment dataset has provided insightful findings into the structure, trends, and behavioral patterns of judicial decision-making. This section explains the analytical findings in the general context of legal data mining and judicial informatics.**

**1. Understanding Judicial Trends**

**The steady rise and fluctuations in volumes of cases year by year mirror both the increasing workload of the judiciary and changing legal issues in society. Peaks in certain years may be correlated with notable political events, legal changes, or societal changes. These patterns over time are important for policy researchers and legal academics to follow the responsiveness of the judiciary.**

**2. Voting Behavior and Decision-Making**

**The prevalence of majority and unanimous votes implies strong agreement among judges, perhaps reflecting institutional stability. The existence of outliers and high variance in minority voting in some areas of issues—e.g., civil liberties and constitutional law—can reflect ideologically polarized opinions or legally intricate issues. This underscores the need to account for judicial philosophy and case complexity in legal analysis.**

**3. Legal Value of Correlation and Covariance**

**The expected positive correlation between majority\_vote and total\_votes, though, the weak negative correlation between majority\_vote and minority\_vote raises further questions. It might be suggestive of case polarization or strategic judicial alignment in particular legal areas. Covariance reinforces such dynamics with implications of predictive modeling depending on voting behavior.**

**4. Patterns in Issue Areas**

**The dense distribution of cases in topics such as criminal law and civil rights corresponds with frequent legal terminology and public interest issues. Visual representation tools such as violin and swarm plots revealed variation in judicial activity across these topics. Such an outcome may prove to be extremely valuable to legal professionals, as they may be able to predict the types of deliberation and rulings in future cases of a similar nature.**

**5. Methodological Contribution**

**The blending of statistical reasoning and visualization with domain knowledge (judicial decision-making) illustrates the power of the union of data science and legal informatics. As exploratory, this research sets up a replicable framework for examining legal judgments at scale.**

**Limitations**

**Although the analysis offers strong statistical findings, it does not take into consideration the textual context of the cases or the legal arguments themselves. Moreover, metadata fields like issue\_area and decision\_type can be constrained by inconsistencies or missing values. An extension in the future using Natural Language Processing (NLP) on full-text judgments could transcend these limitations.**

**Conclusion and Future Work**

**Conclusion**

**In this study, we explored the potential of combining exploratory data analysis (EDA) with deep learning techniques to study and classify Supreme Court decisions. From a cautious statistical analysis, we uncovered compelling trends in judge behavior, such as the predominance of majority decisions and vote counts correlation. These results, in addition to the issue-area and temporal patterns, not only increase the richness of our knowledge regarding judicial decision-making but also create a basis for ongoing advancement in legal data analysis.**

**In the second phase, we employed some deep learning architectures like LSTM, Bi-LSTM, and CNN for legal text classification. These models were promising in automating judgments' classification according to issue areas, and this was a demonstration of NLP's potential in legal research. This amalgamation of statistical modelling with deep learning is indicative of the future of legal scholarship in which technology can enhance the process by eliminating drudgery, lightening the load on mankind, and providing useful facts that would otherwise not be more readily accessible.**

**Future Work**

**While this study indicates the promise of combining EDA and deep learning for legal document analysis, there are a few directions for future work. First, applying a more extensive dataset with larger numbers of judgments and more varied legal systems might provide a clearer view of the behavior of judges. Second, applying more advanced NLP techniques, such as BERT or other transformer-based models, may improve accuracy and context-aware legal language interpretation.**

**Another avenue of future research would be the inclusion of full-text analysis rather than relying solely on metadata and summary descriptions. This would enable closer examination of the legal arguments, reasoning, and language used by the courts. Furthermore, exploration of multimodal approaches, which combine text with other data (e.g., judicial history, political context), could enhance predictive modeling and outcome prediction.**

**Lastly, improving the interpretability of deep learning models in legal scenarios is extremely critical. Legal professionals must be able to understand and accept the decisions of automated systems. Adding explainability frameworks to the models would therefore be a valuable effort towards bridging the gap between legal professionals and technology.**

**Through further development of these foundational steps, future work can continue to more advance and improve the integration of AI and legal analysis, bringing more efficiency and expertise to the practice and research of law.**

**References**

1. **Aletras, N., Tsarapatsanis, D., Preotiuc-Pietro, D., & Lampos, V. (2016). Predicting judicial decisions of the European Court of Human Rights: A Natural Language Processing perspective. *PeerJ Computer Science, 2*, e93.** [**https://doi.org/10.7717/peerj-cs.93**](https://doi.org/10.7717/peerj-cs.93)
2. **Bhattacharya, S., Banerjee, D., & Sharma, S. (2019). Summarization and citation analysis in the Indian legal system. *Proceedings of the 2019 International Conference on Natural Language Processing (ICON 2019)*, 45-58.** [**https://www.icon2019.in/**](https://www.icon2019.in/)
3. **Chalkidis, I., Zachariadis, I., & Aletras, N. (2019). Hierarchical attention networks for legal document classification. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019)*, 3143-3153.** [**https://www.aclweb.org/anthology/P19-1304/**](https://www.aclweb.org/anthology/P19-1304/)
4. **Sulea, D., Rădulescu, D., & Găvănescu, R. (2017). Preprocessing and exploratory data analysis in legal corpora. *Proceedings of the International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2017)*, 987-997.** [**https://www.kes2017.org/**](https://www.kes2017.org/)
5. **Zhong, Z., Zhang, T., & Zhou, T. (2020). Fine-tuning BERT for legal document prediction. *Journal of Artificial Intelligence and Law, 28*(1), 53-78.** [**https://doi.org/10.1007/s10506-019-09222-5**](https://doi.org/10.1007/s10506-019-09222-5)