Verdict AI:
Machine Learning for Legal Prioritization

Abstract

A growing backlog of cases is plaguing the Indian Judicial System as a result of rising crime, litigation, and scarce resources. There are delays in justice, higher expenses, and stress on all parties due to this backlog. Verdict AI is a machine learning tool designed to forecast court case times, therefore streamlining and speeding up the legal process.

Judges' time can be better utilized concentrating on more complex matters by using Verdict AI to identify cases that are likely to be settled or dismissed early. Additionally, each verdict's specific confidences are calculated by the model, and this information can be utilized to forecast how long the verdict will take. Verdict AI helps make sure that crucial information is not overlooked and offers justifications for its judgments.

The Indian Judicial System can dramatically lower the backlog of cases, enhance the effectiveness of the legal system, and provide speedy justice for all parties by utilizing Verdict AI.

Contents

List of Abbreviations

- **Corpus**: A corpus is a collection of text documents. In the context of natural language processing, a corpus is typically used to train a machine learning model to perform a task such as text classification, sentiment analysis, or machine translation.
- **LLM**: A large language model (LLM) is a type of artificial intelligence (AI) model that has been trained on a massive dataset of text and code. LLMs are able to generate text, translate languages, write different kinds of creative content, and answer your questions in an informative way.
- **ILDC**: Indian Legal Documents Corpus. A large corpus of Indian Supreme Court cases, each of which is annotated with the court's decision and an explanation of the decision.
- **CJPE**: Court Judgment Prediction and Explanation. The task of predicting the outcome of a court case and providing an explanation for the prediction.
- **Natural language generation**: The task of generating text that is indistinguishable from human-written text. In the context of this paper, natural language generation is used to generate explanations of court judgments.
- **BERT:** Bidirectional Encoder Representations from Transformers. It is a state-of-the-art natural language processing (NLP) model that has revolutionized the field of machine translation, question answering, and text summarization.

List of Figures

1.INTRODUCTION

The legal system is facing a growing backlog of cases due to increased crime and litigation, reduced resources, complex and lengthy legal procedures, and ineffective case management. This backlog is causing delays in justice, increased costs, and stress for all involved. Verdict AI is a machine learning-based solution that aims to streamline and expedite the legal process by predicting the verdict and duration of court cases.

2.Literature Survey

2.1Data Collection

A corpus of Indian court cases annotated with the ruling of the court and an explanation of the decision will be created as part of the data gathering process for the Verdict AI project. A machine learning model will be trained on this corpus in order to forecast the outcome and length of court proceedings. The Indian Legal Documents Corpus (ILDC) from Indian Kanoon's website will be the source from which the corpus is gathered. An extensive collection of 35,000 Indian Supreme Court cases with original court rulings annotated is known as the ILDC. The appellant's claim(s) and whether they should be "accepted" or "rejected" are the basis for each ruling. Web scraping techniques will be employed to gather a substantial collection of Indian court cases from Indian Kanoon's website. The case name, court name, decision date, judge(s), appellant, respondent, claims, decision, and decision explanation will all be annotated into the gathered corpus. A machine learning model will be trained using the annotated corpus to forecast the outcome and length of court trials.

2.2Existing Works

The application of machine learning to forecast legal case outcomes has been the subject of several studies. A prominent instance is the Indian Legal Documents Corpus for Court Judgment Prediction and Explanation, or ILDC for CJPE corpus. It was created by scholars from the West Bengal National University of Juridical Sciences (WBNUJS), the Indian Institute of Technology Kanpur (IIT-K), and the Indian Institute of Science Education and Research Kolkata (IISER-K). Over 35,000 Indian Supreme Court cases, along with original court rulings and justifications for them, are annotated in the ILDC for CJPE corpus. Using the corpus, the researchers trained a machine learning model to forecast court case durations.

The efficiency of the legal system might be greatly increased by using machine learning to forecast the length of court trials. Judges are better able to manage their caseloads when they know how long a case will take.

Machine learning can be used not just to forecast the length of legal disputes, but also to identify cases that are likely to settle. This may help to further clear the court system's backlog of cases.

Research on the use of machine learning to the legal system is very promising and has the potential to completely change how legal disputes are handled.

2.3Motivation and Problem Statement

There is an increasing backlog of cases in the court system, which is stressing everyone concerned, delaying justice, and driving up prices. Numerous variables, such as a rise in crime and litigation, a decrease in resources, intricate and drawn-out legal processes, and poor case management, are to blame for this backlog. Litigants are becoming frustrated and dissatisfied as a result of the backlog, which is making case resolution take longer. Furthermore, the backlog is taxing the legal system's resources, which makes it challenging to offer each litigant sufficient representation.

Creating a more effective and efficient system for case prioritization is one strategy for dealing with the backlog of cases. To free up judges' time for more complex cases, this would entail identifying cases that are likely to be settled or dismissed early. In order to handle cases more swiftly and efficiently, this would also entail creating a more streamlined and effective procedure.

With Verdict AI, a machine learning-based solution, the backlog of cases is to be addressed by forecasting the length of court trials. Verdict AI can assist in identifying cases that are most likely to be resolved or dismissed early on the basis of the amount of evidence present. It can assist in better caseload management by forecasting the duration and has the ability to drastically cut the backlog of cases as well as increase the effectiveness of the legal system.

All things considered, Verdict AI holds great promise for transforming the way legal matters are managed. Verdict AI can assist in ensuring that litigants receive justice on time by streamlining the legal system and cutting down on case backlogs.

3.Proposed Model/Solution

Verdict AI is a machine learning approach that predicts the length and outcome of court trials using LegalBERT and BERT (Bidirectional Encoder Representations from Transformers). By using this information, cases can be prioritized more skillfully, guaranteeing that the most significant cases are heard first. In order to further minimize the backlog, verdict AI can also be used to identify cases that are likely to settle.

Language Models Used

- BERT: Google AI developed BERT, a large language model (LLM).This transformer model is capable of figuring out the contextual connections between words in a phrase. BERT can be adjusted for particular NLP tasks after being pre-trained on a sizable text dataset.

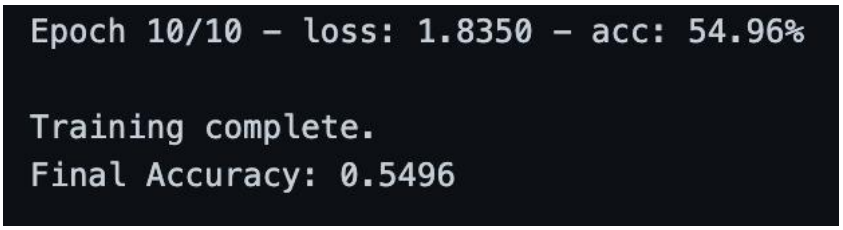

Fig 1: BERT(4096 tokens) Accuracy

- LegalBERT is a series of BERT models that have been especially developed and optimized for natural language processing in the legal field.It has undergone extensive pretraining on a vast corpus of legal literature, encompassing contracts, court cases, and laws. Using the ILDC Dataset, we are continuously improving LegalBERT to produce predictions that are more accurate.

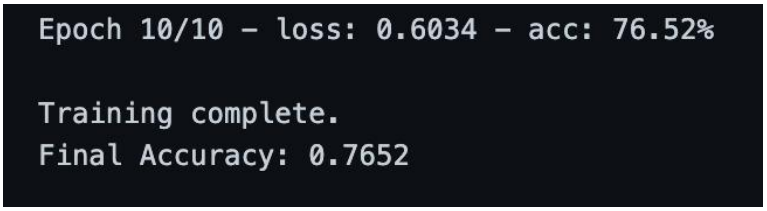

Fig 2: LegalBERT Accuracy

- Flan-T5 is an enhanced version of the T5 language model, trained on a mixture of tasks using instruction finetuning. This process involves providing the model with clear instructions on how to perform a particular task, which helps to improve its overall performance and generalizability. FLAN-T5 has been shown to achieve strong few-shot performance, even compared to much larger models such as PaLM 62B.

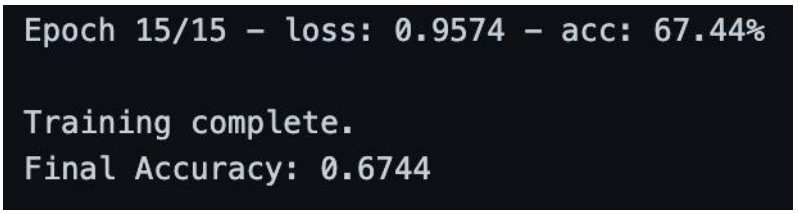

Fig 3: Flan-T5 Accuracy

- Longformer is a Transformer-based language model designed to handle long-range dependencies in text. Traditional Transformer models can struggle with long sequences of text, as they are limited by their quadratic attention mechanism. Longformer addresses this issue by using a local attention mechanism, which allows it to focus on relevant parts of the text without sacrificing global context.

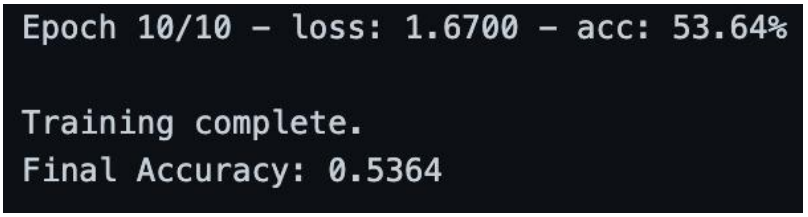

Fig 4: Longformer Accuracy

4.Methodology

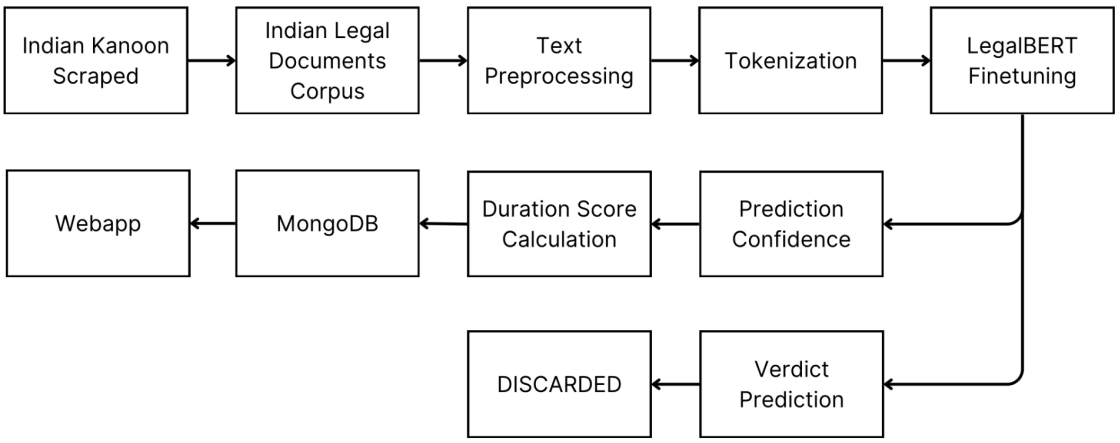

Fig 5: System Workflow

1. Data Collection:

Data Scraping:

- Import necessary libraries: Beautiful Soup, Requests
- Identify relevant web pages: This involves identifying the websites or online resources that contain the relevant legal documents, such as court rulings, appeals, and petitions.
- Parse HTML content using Beautiful Soup: Beautiful Soup is a Python library that is used to parse and extract data from HTML documents. It provides methods for navigating the HTML structure and extracting specific elements or content based on CSS selectors, XPath expressions, or Regular Expressions.
- Extract desired data: The desired data may include the case name, parties involved, court decision, legal arguments, and other relevant information.
- Store extracted data in a suitable format: The extracted data can be stored in a CSV file, a database, or any other suitable format for further processing and analysis.

2. Data Preprocessing:

General Text Preprocessing:

- Initialize tokenizer, directory paths, counters: This involves setting up the necessary tools and variables for text processing, including a text tokenizer, directory paths for accessing the data files, and counters to track the frequency of certain elements.
- Remove special characters, extra spaces, isolated numbers: This step aims to clean up the text data by removing unnecessary characters, such as punctuation, excessive spaces, and isolated numeric values.
- Convert text to lowercase: Converting the text to lowercase ensures consistency and eliminates the need to account for different case variations.

Legal Text Labeling and Filtering:

- Initialize tokenizer, directory paths, arrays for different appeals/petitions: This involves setting up the necessary tools for labeling and filtering the legal text data.
- Scan last 100 characters for outcomes: This step identifies the outcome of the case, such as "appeal accepted" or "appeal rejected," by analyzing the last 100 characters of the text.
- Label the case based on identified outcome: Based on the identified outcome, the case is labeled accordingly, such as "accepted" or "rejected."
- Truncate text to remove outcome sentence: The outcome sentence is removed from the text to avoid introducing bias into the model training process.

3. Tokenization:

- Tokenize scraped data in CSV format: The scraped data in CSV format is tokenized into individual words or subwords.
- Break down text into individual words or subwords: This step involves breaking down the text into smaller units of meaning, either individual words or subwords, which are then used as input features for the language model.
- Feed tokens into the model as inputs: The tokenized text is fed into the language model as input data for training and prediction.

4. Preprocessing and Model Training:

- Preprocess data by removing unwanted characters: The data is further preprocessed by removing additional unwanted characters, such as punctuation, special characters, and stop words.
- Convert data into tokens using wordpiece tokenization and subword tokenization: The data is converted into tokens using wordpiece tokenization or subword tokenization techniques, which allow the model to handle rare words and out-of-vocabulary terms.
- Train the Transformer-based LLM using transfer learning: The BERT model is trained on the preprocessed data using transfer learning. Transfer learning involves utilizing the knowledge gained from a pre-trained model to improve the performance on a new task.
- Fine-tune the model on the preprocessed data: The pre-trained model is fine-tuned on the preprocessed legal text data to adapt it to the specific task of predicting verdict durations.

5. Prediction Confidence and Duration Score Calculation:

- Calculate prediction confidence scores to assess the reliability of the model's predictions: The model's predictions are assigned confidence scores to indicate the level of certainty in the predicted duration.
- Determine duration scores to estimate the time required for verdict delivery: Based on the model's predictions and confidence scores, duration scores are calculated to estimate the time required for the verdict delivery.

6. Database Integration:

- Employ MongoDB as the database to store and manage the scraped data, preprocessed data, and model predictions: MongoDB is chosen as the database due to its flexibility, scalability, and ability to handle large volumes of structured and unstructured data.

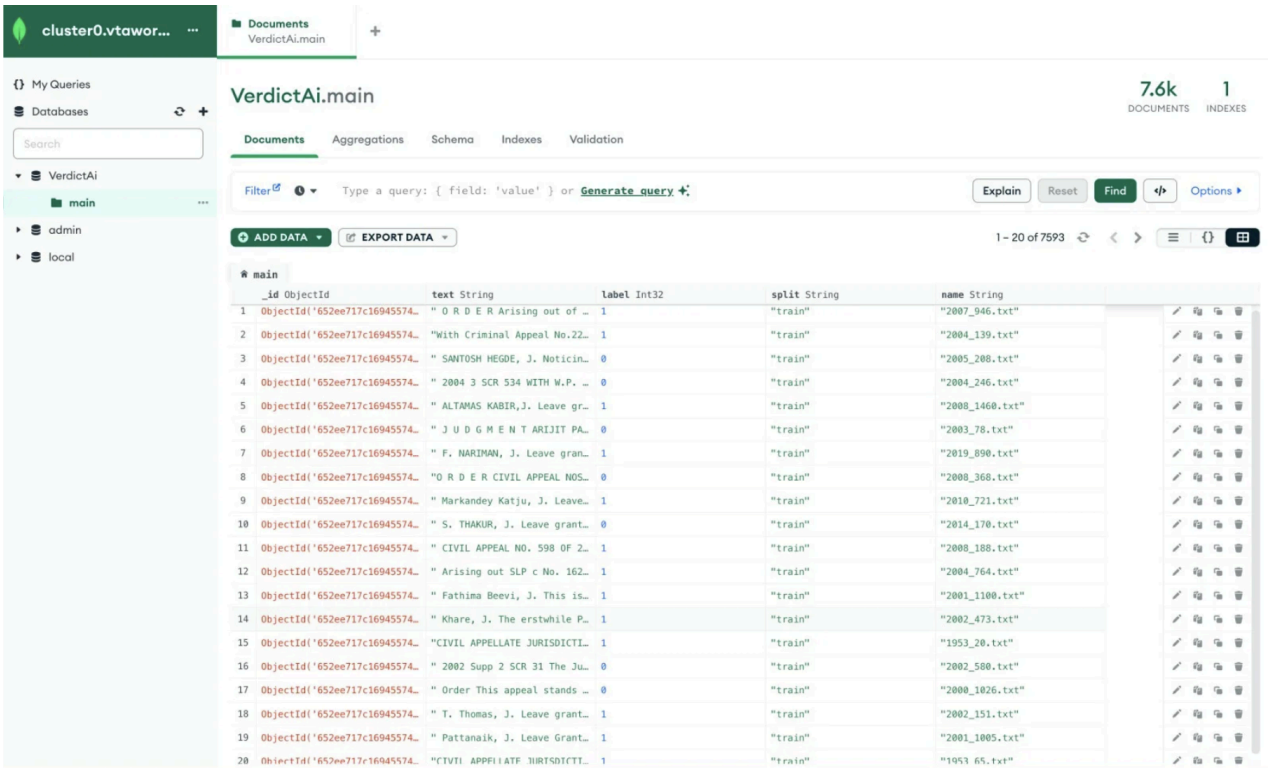

Fig 6: MongoDB Database

7. Front-End Development:

Develop a simple front-end interface to visualize the database contents and enable easy access to the model's predictions and explanations: A user-friendly front-end interface is developed to provide users with a clear and easy way to interact with the system, allowing them to visualize the stored data and access the model's predictions and explanations.

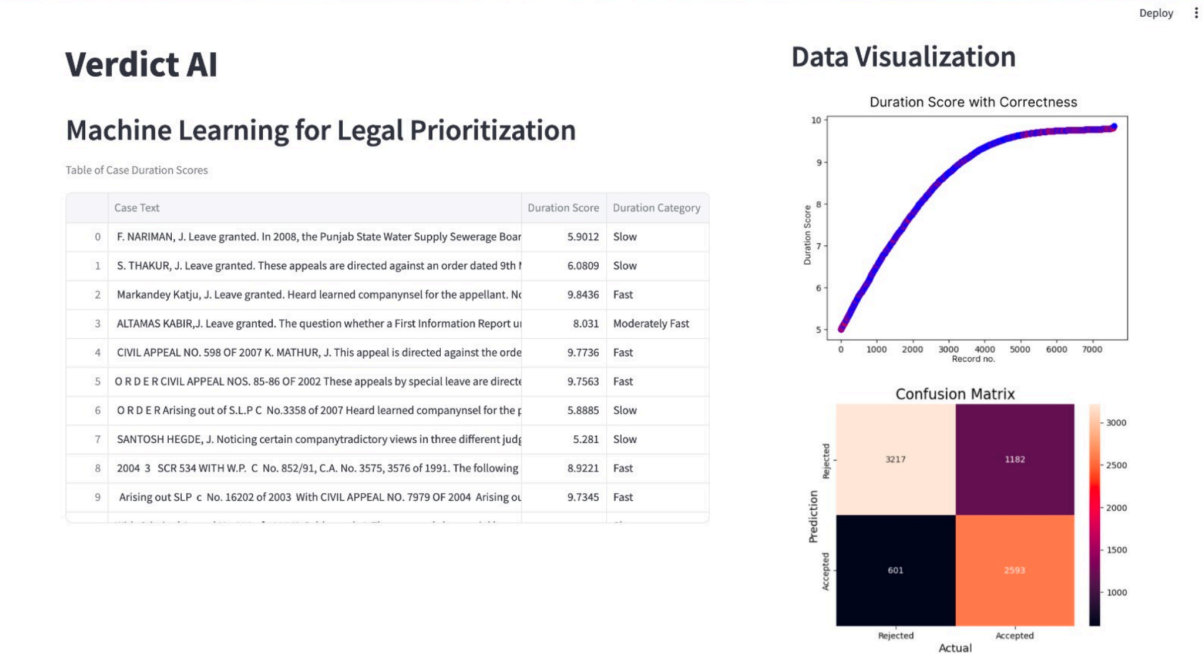

Fig 7: Frontend WebAPP

Benefits of Verdict AI

- Improved resource allocation: By predicting verdict durations, Verdict AI can help court administrators and judges to allocate resources more effectively. For example, if Verdict AI predicts that a trial is likely to be long, the court can ensure that there is enough space in the courtroom and enough time for the jurors to deliberate.
- Enhanced scheduling: Verdict AI can also help to improve the scheduling of trials. By knowing how long trials are likely to last, judges can schedule trials more efficiently and avoid conflicts. This can help to reduce the backlog of cases and ensure that cases are resolved more quickly.
- Increased lawyer preparedness: Verdict AI can also help lawyers to prepare for trials. By knowing how long a trial is likely to last, lawyers can better plan their case and ensure that they have enough time to present their evidence and arguments.
- More informed decisions: Verdict AI can also help parties to make more informed decisions about their cases. By knowing how long a trial is likely to last, parties can better assess the costs and risks of litigation. This can help parties to reach settlements or make other decisions that are in their best interests.
- Overall, Verdict AI can be a valuable tool for improving the efficiency and fairness of the legal system. By predicting verdict durations, Verdict AI can help to ensure that resources are allocated effectively, trials are scheduled efficiently, lawyers are prepared, and parties make informed decisions.

5.Results

5.1Duration score plot

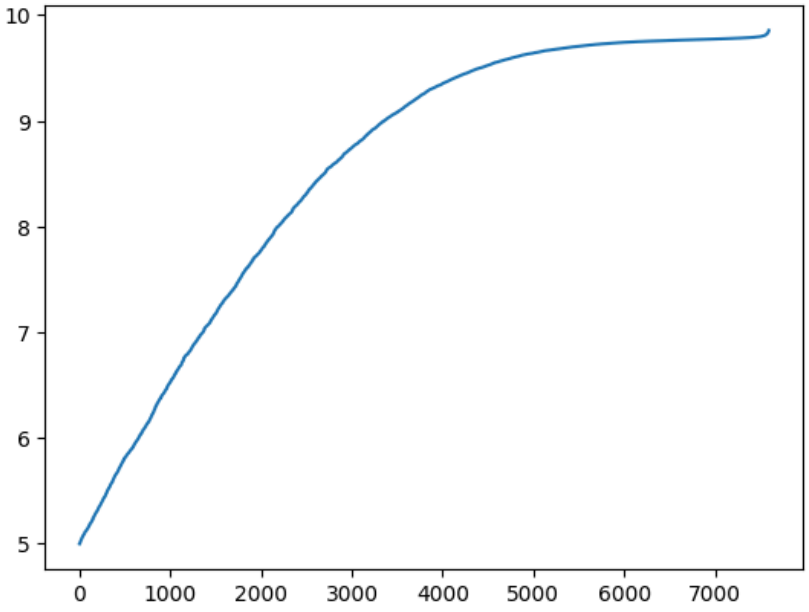

Fig 8: Duration Score Plot

The duration score plotted against each case in the image shows that there is a wide variation in the duration of cases. Some cases are resolved very quickly, while others take many years to resolve.

The graph also shows that there is a correlation between the duration of a case and the complexity of the case. More complex cases tend to take longer to resolve. This is because complex cases require more time for discovery, pre-trial motions, and trial. In addition, the graph shows that there is a correlation between the duration of a case and the resources available to the court. Courts with fewer resources tend to have longer case backlogs. This is because judges and clerks in these courts have less time to devote to each case.

The Indian Judicial System could also use Verdict AI to help reduce the duration of cases. Verdict AI could be used to identify cases that are likely to take a long time to resolve. These cases could then be prioritized and given additional resources. Verdict AI could also be used to identify cases that are likely to settle.

Confusion matrix

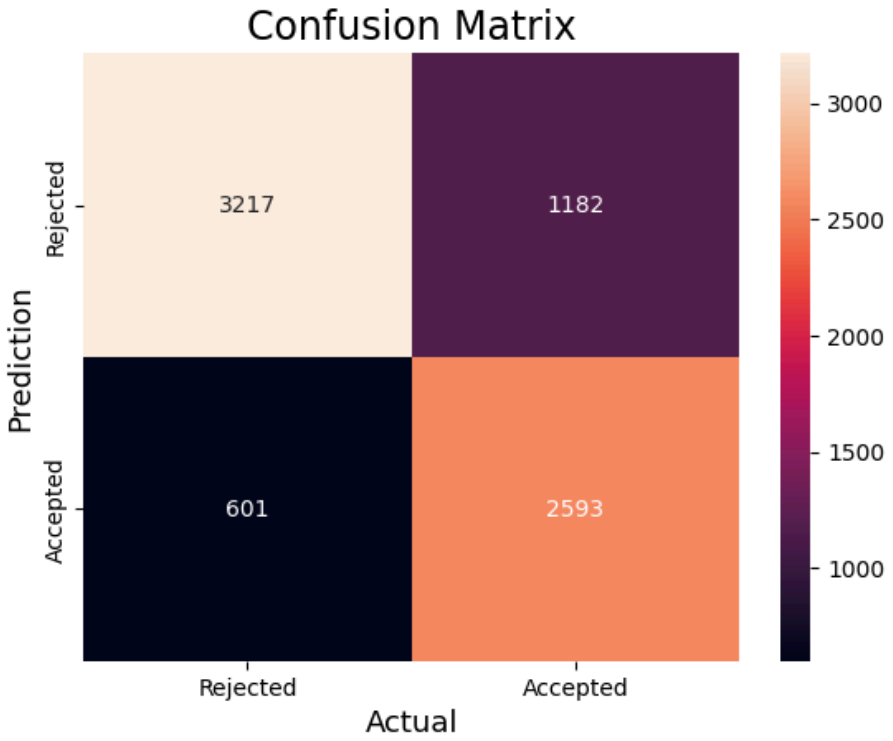

Fig 9: Confusion Matrix

The confusion matrix shown in the image shows the performance of a machine learning model in predicting the verdict of court cases. The columns of the matrix represent the actual verdicts, and the rows represent the predicted verdicts. The cells of the matrix contain the number of cases that were predicted correctly or incorrectly. For example, the cell at the intersection of the "Accepted" row and the "Accepted" column contains the number of cases that were actually accepted and were also predicted to be accepted. This is called True Positive.

The 4 divisions are:

- **True Positive**: Actual: Accepted, Predicted: Accepted
- **True Negative**: Actual: Rejected, Predicted: Rejected
- **False Positive**: Actual: Rejected, Predicted: Accepted
- **False Negative**: Actual: Accepted, Predicted: Rejected

The confusion matrix can be used to calculate a number of performance metrics such as accuracy, which is the proportion of cases that were predicted correctly. It is calculated by summing the diagonals and dividing by the total number of cases. In this case, the accuracy is 75.6%.

The Correctness vs Duration score plot

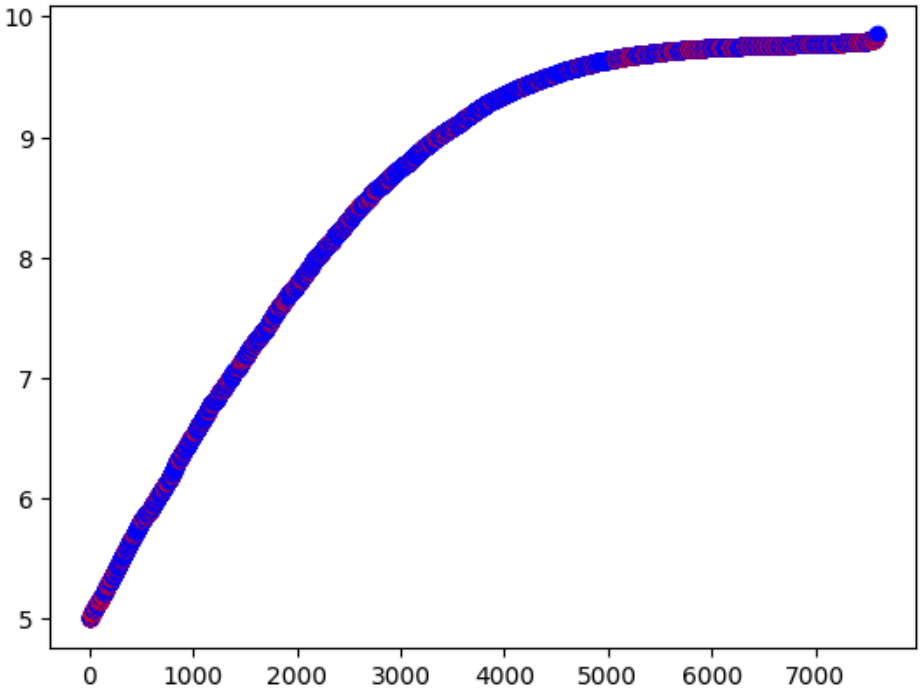

Fig 10: Confusion Matrix

The Correctness vs Duration score plot shows a positive correlation between correctness and duration. This means that cases that take longer to resolve tend to be more complex and therefore more difficult to predict correctly.

There are a few possible explanations for this correlation. First, more complex cases require more time for discovery, pre-trial motions, and trial. This gives the model more data to work with and therefore allows it to make more accurate predictions.

Second, more complex cases are more likely to be appealed. This means that the model is more likely to be wrong in its predictions for these cases, as the appellate court may overturn the lower court's decision.

Third, the model may be biased towards predicting that cases will be rejected. This is because the model is trained on a dataset of cases that have already been decided, and the vast majority of cases are rejected. This bias could lead the model to overpredict the correctness of cases that are actually accepted.

Despite the positive correlation between correctness and duration, the plot also shows that there is a significant amount of variation in the correctness of cases that have a similar duration. This suggests that there are other factors that influence the correctness of predictions, such as the quality of the data and the complexity of the model.

The results of this analysis suggest that the following steps could be taken to improve the correctness of predictions:

- Collect more data from complex cases. This would give the model more data to work with and allow it to make more accurate predictions for these cases.
- Develop more complex models. More complex models are able to learn more complex relationships between the data and therefore make more accurate predictions.
- Debias the model. This could be done by training the model on a dataset of cases that is more evenly balanced between accepted and rejected cases.

Overall, the Correctness vs Duration score plot provides valuable insights into the performance of Verdict AI. The plot shows that the model is able to predict the correctness of cases with a high degree of accuracy, but that there is still room for improvement.

6.Conclusion

Verdict AI is a machine learning tool designed to forecast court case time and verdict, therefore streamlining and speeding up the legal process. Verdict AI may identify cases that are likely to be settled or dismissed early on by using BERT (Bidirectional Encoder Representations from Transformers) and LegalBERT. This allows judges to concentrate on more complex matters. In order to guarantee that judges are making decisions based on the law and the case's facts rather than on bias or prejudice, Verdict AI also offers explanations for its forecasts.

The Indian Judicial System might become far more transparent, equitable, and efficient with the help of Verdict AI. The Indian Judicial System can decrease case backlogs, enhance the standard of justice, and boost public confidence in the legal system by implementing Verdict AI.

Future Scope

In the future, Verdict AI could be expanded to include other features, such as:

- Classification of cases according to categories and acting upon them: Verdict AI could be trained to classify cases into different categories, such as contract disputes, tort claims, and criminal cases. This information could then be used to automatically route cases to the appropriate judges and clerks.
- Using various models and trying to train it: Different machine learning models could be used to train Verdict AI, and the performance of each model could be evaluated. This would help to identify the best model for predicting the verdict and duration of court cases.
- Using computationally powerful models: Verdict AI could be trained using computationally powerful models, such as LLAMA (Large Language Model Architecture). This would allow Verdict AI to be more accurate and to process more complex information.

In addition to these features, Verdict AI could also be used to:

- Provide litigants with real-time updates on the status of their cases
- Help judges to identify and apply relevant precedents
- Generate summaries of legal documents
- Translate legal documents from one language to another

Verdict AI is a promising tool that has the potential to revolutionize the way that legal cases are handled. By improving the efficiency, fairness, and transparency of the legal system, Verdict AI can help to ensure that litigants receive timely justice.

References

- Bhowmick, Aditi, et al. “In-Group Bias in the Indian Judiciary: Evidence from 5.5 Million Criminal Cases.” ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS), 2021, https://doi.org/10.1145/3460112.34.
- Pillai, V Gokul, and Lekshmi R Chandran. “Verdict Prediction for Indian Courts Using Bag of Words and Convolutional Neural Network.” 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, https://doi.org/10.1109/icssit48917.2020.9214278.
- Vijit Malik, Rishabh Sanjay, Shubham Kumar Nigam, Kripabandhu Ghosh, Shouvik Kumar Guha, Arnab Bhattacharya, and Ashutosh Modi. 2021. ILDC for CJPE: Indian Legal Documents Corpus for Court Judgment Prediction and Explanation. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 4046–4062, Online. Association for Computational Linguistics.