Volume 3, Issue VII, July 2024, **No. 43, pp. 556-572**
Submitted 29/06/2024; Final peer review 11/7/2024
Online Publication 17/7/2025
Available Online at http://www.ijortacs.com

DEVELOPMENT OF A HYBRID INTELLIGENT MODEL FOR CONTRACT LITIGATION

[1*]Nnaemeka C. Onyemelukwe, [2]Ogochukwu C Okeke
[1,2]Department of Computer Science, Chukwuemeka Odumegwu Ojukwu University
Anambra state
[1*]Corresponding Author's Email: Onyemelukwennaemeka@gamil.com

Abstract

The study addresses the challenges faced in contract litigation due to the complexity and volume of legal documents. Traditional legal text analysis methods are often labor-intensive and prone to errors, leading to inefficiencies and potential risks. To address these issues, the study explores the use of advanced technologies, such as natural language processing (NLP) and deep learning, to automate and improve the accuracy of legal document analysis. The study proposes a novel model that integrates BM25, a Convolutional Neural Network (CNN), and a Bidirectional Long Short-Term Memory (BiLSTM). This hybrid approach combines the strengths of BM25 for efficient information retrieval, CNN for capturing local text patterns, and BiLSTM for modeling sequential dependencies. The aim is to enhance the system's ability to analyze legal documents with greater accuracy and contextual understanding. The study seeks to develop and evaluate this integrated model to provide legal professionals with a powerful tool for streamlining document review, extracting key information, and making informed legal decisions in contract litigation. Additionally, the study explores the implications and challenges associated with implementing such technology in real-world legal settings, contributing to advancements in technology-assisted legal analysis and improving access to justice.

Keyword: Legal document analysis, Natural language processing (NLP), Deep learning, Hybrid model (BM25, CNN, BiLSTM), Technology-assisted legal analysis

1. INTRODUCTION

Legal content produced every day in court cases is increasing and is now accessible in electronic form. The analysis of legal documents using various machine learning and deep learning approaches has harvested much interest in recent years. (Giwa & Kodjovi, 2023). Legal principles are applied to situations that may result in the lodging of a case through the use of legislation. The process takes time, and it is crucial to apply the relevant statutes and find the precedent cases. There is a need for an automated system that can determine which statutes and cases from the past prior cases are appropriate in any contract litigation. This will make it easier for the attorneys to comprehend the case in its entirety and determine where the issue lies. (Castano et al., 2024)

Contract litigation, a cornerstone of the legal system, plays a pivotal role in resolving disputes and upholding the sanctity of agreements. In this domain, the accurate interpretation of contractual terms and conditions is paramount, as it influences the outcomes of legal proceedings. (Ahmad *et al.*, 2022) However, the processes governing contract litigation have long grappled with inherent challenges, comprising inefficiency, subjectivity, limited scalability, and the risk of errors. The manual analysis of precedent cases and interpretation of complex legal documents (statute cases)is not only labour-intensive but also prone to human error. This often results in delays in resolving disputes and escalates costs for all parties involved. The subjective nature of human analysis introduces inconsistencies in decision-making, posing a significant threat to the fairness and accuracy of contract litigation outcomes. As the volume of contracts continues to grow, legal professionals face the formidable challenge of managing and analyzing various precedent/statutes case documents within reasonable time frames (Hassan, 2022). This limitation not only impedes the timely resolution of contract disputes but also adds complexity and uncertainty to the legal landscape. The need for a transformative change in contract litigation has never been more apparent. Additionally, the majority of evaluations of case law retrieving have been carried out on limited collections and concentrate on related tasks like question-answering or recommendation systems, but not a rank system (Locke & Zuccon, 2022). The legal search providers are unwilling to divulge to the competition the trade secrets to their success in information retrieval.

The researcher has geared research efforts toward effective artificial intelligence (AI) and machine learning (ML) Technology which expanded the range of analytical techniques that can analyze legal documents, creating meaningful new study opportunities for both computer science and law experts. One such change is to represent legal regulations as executable code. In a legal document, the addition of each party's signature at the end records and confirms an agreement between two or more parties. There has been a resurgence of interest in finding a solution to the problems related to legal document review and how these documents can be categorized as privileged, responsive, and relevant using an AL and ML algorithm. Based on this, the research will concentrate on designing and implementing a hybrid model for the Artificial Intelligence for Legal Assistance (AILA) system which is a novel approach that integrates cutting-edge deep learning methods with contextual awareness to optimize the Technology-Assisted Review (TAR) process in legal contract litigation, which aims to understand the legal context, evolving case dynamics and user preferences to provide more accurate and efficient document review.

The proposed system lies in its integration of two complementary approaches for legal document retrieval: Best Match 25 (BM25) and a Convolutional Neural Network (CNN) with an aggregated Bidirectional Long Short-Term Memory (BiLSTM) model. BM25 is an information retrieval method that effectively ranks documents based on term frequencies and document length (Gomede, 2023). On the other hand, the CNN-BiLSTM model captures semantic relationships and contextual information within the text (Rhanoui et al., 2019). The integration of these methods offers a comprehensive approach that combines traditional keyword-based retrieval with deep learning techniques to improve the retrieval of legal documents. This hybrid system leverages the strengths of both methods to enhance the accuracy and relevance of document ranking, particularly in legal information retrieval.

Saxena,(2023) shows that aggregated Bidirectional Long Short-Term Memory (BiLSTM) is a crucial component of the proposed legal document retrieval system and plays a significant role in improving retrieval accuracy, particularly in comprehending the contextual information and semantic relationships included in the legal text. With the help of this project, legal professionals will locate relevant precedent case laws and legal status documents more quickly and accurately by developing a hybrid model for legal document retrieval. The intention behind this method is to bring together the benefits of deep learning models' semantic understanding with the advantages of conventional information retrieval methods BM25 (Gomede, 2023). Enhancing the ranking of relevant files, increasing retrieval accuracy, and eventually helping legal practitioners save time and effort are the main objectives. The system hopes to accomplish these objectives and offer a useful resource for legal research and decision-making. Using (IR LEGPRE, 2021), determine the most pertinent statutes for a given situation (Statute Retrieval) and relevant earlier decisions in the specific situation (Precedent Retrieval).

Millidge *et al*., 2021 (2021) gives the definition of predictive coding as machine learning that applies human-inputted data about document relevance to a considerably bigger document data collection. Using keyword search, filtering, summarizing, extracting, and sampling for e-discovery document review. Predictive coding technology, also known as tech-assisted review (TAR) or automated review (AR), is utilized in the legal case review process to locate digitally stored data (DSD) documents that are responsive. Artificial intelligence is used in predictive coding to create software that learns from experience and makes better decisions while greatly accelerating the review process to save time and money. The researcher found certain issues that show the method employed by other researchers, based on an assessment of the literature. Another concern is determining the best strategy for performing at your very best when reviewing legal documents with multilingual features. In contrast to the proposed design, most existing research that uses the Navia Bays bag of words (BOW) classification technique to address AI document review in litigation cases frequently results in zero probability issues because it assumes that all predictors are independent, which does not always occur when a human is reviewing legal documents, (Ifeanyi-Reuben *et al*., 2018). For illustration, most current systems will translate "school" into the native language Igbo using the BOW model for text representation, which fails to distinguish the compound word "ụlọ" or "akwụkwọ" as a separate term, which is incorrect. The goal of the suggested design is to resolve the issue.

Studies on support vector machine (SVM) show that, because of its analysis, confidence ratings rather than posterior probabilities are not suitable for the classification of enormous data sets. As they support most predictive coding on posterior probability, even though it performs well with small data sets, the suggested research tries to create a better one.

Since legal knowledge safeguards people's lives and is more valuable, it stimulates rivalry and delays legal proceedings. (Pisner *et al*., 2020). When two parties cannot agree on a dispute's fair resolution, litigation is the procedure through which the researcher finally resolves the issue in a court of law. There isn't much to brag about, despite the greater service rate and fee that a client pays the attorney for document evaluations. As a result, the proposed technology will significantly disrupt the legal industry, damaging various practitioners' financial prospects while benefiting from lower-cost input made possible by the new architecture. The common approaches implemented in supervised learning methods for lexical features and metadata features are linear models which include LR (logistic regression) and SVM (Support Vector Machine). Even the suggested study demonstrated that linear models such as LR and SVM outperformed BERT in legal discovery topics in Jeb Bush's email collection. (Yang *et al*., 2021)

According to Keeling (2019), attorney review is the process where lawyers must analyze a document and decide on its applicability, significance, and potential concessions. E-discovery is the forensic collection and production of all documents important to a legal case. Attorneys evaluate one document at a time but are frequently required to review hundreds of thousands, which only helps to compound the problem that has to be handled. The volume of documents that must be gathered, sorted, and annotated in a corporate investigation is frequently enormous. This process takes a long time, and it's difficult to separate responsive information from non-responsive content that should not be disclosed to the opposing side when reviewing.

To address the question of how a response variable depends on one or more predictors, some researchers have employed regression analysis and comparable techniques in some previous work. The mean, also known as a response variable outcome, is typically used to determine dependence, and independent variables typically exhibit a linear relationship between the slope and the intercept. Because of the discrete nature of the model, the approach is less suitable for handling the high-dimensional data that was made available for the current investigation. The regression model is unstable when data from legal documents is extracted. The purpose of this hybrid intelligent machine learning algorithms show more capabilityof handling high-dimensional data than conventional mathematical models by integrating BM25 with the CNN-aggregated BiLSTM model in legal text retrieval (Rhanoui *et al*., 2019) to enhance document ranking by combining ranking BM25 search algorithm with deep learning, and to improve semantic understanding in a legal context

The challenges and ambitions of human life have grown increasingly complex throughout time, demanding our laborious search for appropriate solutions. This attitude toward problem-solving has ultimately brought us to the current technological period when there is ongoing research and development aimed at making machines as intelligent as or even more intelligent than humans (Giwa & Kodjovi, 2023), which is inherited through natural evolution—into the artificial platform throughout this process so that computers, or other machines, might use it to solve complicated problems. It is therefore imperative in this dispensation of "Big legal Data" that the repository for the clause in the sentence boundary classification and tokenization used to identify the specific relevance in legal validity review are deciding factors to achieve a more efficient and effective predictive document review which the researcher seeks to achieve by using sentence segmentation method in the pre-processing step to provide simple sentences as input to normative rule generation in building the classifier for the algorithm, more detail will explain in the method section.

Artificial intelligence for legal help in court documents is now a reality thanks to the entry of new professionals into the legal field. These experts have influenced the discovery process and how attorneys obtain information on the examination of court documents, such as previous decisions and the status of cases pertaining to contract disputes. To gain more insight into these important questions related to modeling a privilege classifier, such as which set of annotations can be relied upon as a reliable basis for evaluation and which remaining annotations produce the greatest outcomes when utilized to train the model. In conclusion, the thesis aims to clarify how lawyers and robots might collaborate in a constructive way.

Our review aims to fill a significant gap in the literature by investigating the use of AI and ML-based systems in Nigeria's legal system for fairness, accountability, and transparency. However, there is little empirical research on how Nigeria's legal system incorporates AI and ML-based decision support and knowledge discovery systems into their actual work practices within the existing research on artificial intelligence and law (Grimm, 2021).The study will combine quantitative and qualitative research methodologies, using user studies, tests, and comparisons with other TAR systems and traditional manual review methods to evaluate the system's effectiveness and user satisfaction. The findings from this investigation will have significant implications for the field of legal document review and information retrieval. The study will contribute to the development of more efficient approaches to document review, which can help reduce the time and cost associated with legal cases and improve the accuracy and quality of document review outcomes. Ultimately, the study aims to promote the use of advanced natural programming language (NLP) techniques in TAR systems and help bridge the gap between human and machine intelligence in the legal field. This paper aims to create an AI-Driven Contract Analysis model by utilizing the combined strengths of machine learning and artificial intelligence technologies. This cutting-edge system is intended to quickly and accurately analyze contract terms and conditions, locate important clauses, and extract relevant data from legal documents with incomparable efficiency. It goes beyond simple automation, though. It involves the combination of AI capabilities and human expertise, resulting in an effective hybrid model.

2. LITERATURE REVIEW

Reviewing legal documents, including decrees, agreements, jurisdictional rulings, and law review articles is a crucial and time-consuming task for any legal scholar or practitioner (Dyevre, 2021). Several ML and NLP processing algorithms can be used to evaluate unstructured data. Additionally, they can efficiently ease their labour with the aid of information retrieval techniques, like classification modelling, word embedding, and transfer of instruction. Legal document review using machine learning: A systematic literature review by (Keeling, R., et al.,2020) provides an overview of the state-of-the-art in machine learning techniques for legal document review, including deep learning algorithms. It identifies key challenges and opportunities for future research in this area.

Wei, et al., (2018) in "The Role of Deep Learning in Legal Document Review" discuss the potential of deep learning algorithms for legal document review and identify several challenges and opportunities for their implementation. It also provides a detailed overview of the different deep learning architectures used for legal document review.

(Cormack & Grassman, 2014) pinpoints an overview of technology-assisted review techniques, including those that use deep learning algorithms, and compares their effectiveness and efficiency to manual review. It highlights the potential of TAR systems to improve the accuracy and consistency of legal document review (Hassan, 2022). Discussed the different deep learning techniques used for natural language processing in legal document review, including text classification and information extraction. It discusses the challenges and opportunities for implementing these techniques in the legal domain.

Castano et al., (2024) uses scientific research methods, particularly USML and SML techniques, to create a legal system artificially intelligently; in unsupervised machine learning, the data set is not labelled, or output is not defined in this situation; Word Embedding (grouping of words that are similar in meaning together in the vector space), Document Clustering with Word Embedding, Topic Modelling, and Latent Semantic Analysis are the key concepts used in legal search analysis. Word Cloud (visualization of the most prominent or frequent words displayed in a body of text), Latent Semantic Analysis, and Principal Component Analysis (Process of analyzing relationships among the set of documents and tokens within).

Hassan, (2022), used text mining and natural language processing to explain the technology-assisted review (TAR) of legal papers. Unsupervised machine learning techniques are a major topic of discussion in the paper. The grouping of papers based on legal text, he further suggested, should be the subject of a future effort.

Park *et al*., (2020), put forward a methodology for text categorization that utilizes cosine similarity to increase the precision of traditional classifiers. The improved classifiers perform exceptionally well in terms of accuracy and comparison to their traditional equivalents. This paper also decides by using a variety of experiments that represent strategies that are best for a given classifier based on total word count and TFIDF representing all datasets. Finally, they concluded by showing that TFIDF is preferable to the use of word count when considering cosine similarity.

According to Ifeanyi-Reuben & Ugwu (2018), the growth of information technology (IT) has promoted the usage of the Igbo language in text creation. In this study, the text structure of Igbo was examined, and a feature selection model for an intelligent Igbo text-based system was created. The most important features of Igbo text documents represented by two word-based ngram text representation models (unigram and bigram) were chosen to use the Mean TF-IDF measure. The Model is created using an object-oriented design methodology and is implemented using Python programming and NLTK tools. Based on language semantics, the outcome demonstrates that bigram-represented text provides more pertinent features which the proposed model aims to archive.

To accelerate their decisions and give legal counsel to attorneys based on historical data and the constitution, (Hassan, 2022) employed a data mining method to develop a technology-assisted review specifically for the "Supreme Court of Pakistan (SCP)" and the "Pakistan Bar Council (PBC)". He acquiesced by stating that unsupervised machine learning includes clustering, topic modelling, latent semantic analysis, and principal component analysis. Legal Sentiment Analysis and Opinion Mining (LSAOM) were also covered in the paper. Spoonmore, (2021) describe how the measurement of cosine similarity provides some insight into the alignment of text vectors between questions and replies. To forecast accurately if the question has been resolved, the context is still lacking.

A context-aware approach in his paper determine whether the current scenario and court records / previous cases are similar by using an efficient word representation that takes term dependency into account. Using the dataset made available by the (Castano et al., 2024) the task organizers, are assessed the approach. He then described the feature weighting strategy for the recommendation system to wrap up this paper. To recommend what they wanted in response to the questions they had to receive suggestions, they looked through statutes and cases of earlier rulings. He also went over the various approaches taken in assessing the questions. The researcher obtained score of 0.015 for task 1 and recall score of 0.05 for task 2 with Glove representation based on the aforementioned implementations. The system performance can be improved by applying different features and different similarity measures.

Sampath and Thenmozhi (2022) demonstrate in their article the precedent-setting method of retrieving a comparable previous case document in the legal sphere for the provided current case document. They identified that consulting previous cases is crucial to guarantee that the same circumstances are handled consistently throughout all situations. The utilization of a combination of statistical, semantic, and subtractive similarity features is suggested. Deep neural networks were sequenced using LSTM as the recurrent unit.

An existing methodology involves utilizing a hybrid deep learning (DL)-based decision support system (DSS), namely CNN with bidirectional long/short-term memory (BiLSTM), to effectively forecast court outcomes from past legal data (Ahmad, 2022). Through the process of feature selection, only the most pertinent features were chosen by giving priority to those that received high scores in the provided legal data set. Next, a hybrid model consisting of CNN and BiLSTM was used to predict the outcomes of court cases. With 91.52% accuracy, 91.74% precision, 89.54% recall, and an F1-score of 90.44%, the hybrid model's experimental results were promising when compared to other similar studies.

Jelali & Messina (2015) proposed online dispute resolution (ODR), an electronic online system intended to settle out-of-court disputes. Court decisions are formal and verbose, and disputant case descriptions are brief and informal. The ODR scheme, eMediation, matches these characteristics better. It uses methods from natural language processing and machine learning to achieve this. The experimental results substantiate the proposed approach's ability to improve court decision retrieval, consequently promoting an informed eMediation process. They advocated the application of more advanced machine-learning techniques.

Numerous studies (Katz *et al*., 2017) have examined the use of computational techniques such as machine learning (ML) to predict court outcomes using historical legal data. However, the sole focus of these investigations was the early prediction of court case decisions. Furthermore, they were limited by (i) an incorrect selection of predictor factors applied to court rulings and (ii) conventional encoders that failed to take into account the correlation between the predictor variables in the legal data. Further analyses of the most current research on experiments utilizing various deep learning, machine learning, and natural language processing methods to forecast administrative and judicial decisions (Francia *et al*., 2022). Among the researcher's most noteworthy findings are the following: The most widely used data mining techniques are Random Forest (RF), K Nearest Neighbors (K-NN), and Support Vector Machine (SVM), while Long-Term Memory (LSTM) and transformers like BERT are the most widely used deep learning techniques. The apparent superiority of machine learning techniques over deep learning techniques was one important finding in the reviewed papers.

2.3 Research Gap

The research gap in Nigerian law highlights the unexplored domain of AI in legal procedures. The proposed study addresses this by integrating BM25 and a Convolutional Neural Network (CNN) with an aggregated Bidirectional Long Short-Term Memory (BiLSTM) model for legal document retrieval. This hybrid system combines traditional keyword-based retrieval with deep learning techniques, enhancing the accuracy and relevance of document ranking. This approach is crucial for fostering fairness, accountability, dynamic contract interpretation, and integrating standards in digital modalities.

3. METHODOLOGY

The researcher used Object-Oriented Analysis and Design (OOAD) methodologies to analyze the present system as well as to design the proposed system with the primary aim of developing a novel model that integrates BM25, Convolutional Neural Network (CNN), and Bidirectional Long Short-Term Memory (BiLSTM) to enhance the accuracy and contextual understanding of legal document analysis in contract litigation. The process of object modeling involved identifying objects and grouping them into classes, identifying the relationships among classes, creating a user object model diagram, defining user object attributes, and defining the operations that should be performed on the classes. By following this OOAD-based methodology, the researcher aimed to develop a comprehensive and well-structured model that integrates advanced natural language processing and deep learning techniques to address the challenges in contract litigation document analysis.

4. SYSTEM ANALYSIS

The proposed approach combines BM25 with two neural networks namely CNN and BiLSTM,
In the system analysis of the model's hybridization, the research proposed the following architectural stages as shown in figure 2:

Stage 1 : Data processing

The data Pre-processing methods indulge in the process of preparing the data included with the filtering process of the dataset. The various stages used are;

i. pre-processing the data; text lower casing; punctuation and special character removal
ii. stop word removal; stemming / lemmatization

Stage 2: Tokenization:

Legal document precedents, statutes, cases, and queries are tokenized and then represented numerically.

Stage 3: CNN features Extraction using a filter to the input

Convolution: $Z = h * W$

Max Pooling: $h_{xy} = max_{i,j} h(x+i)(y+j)$ (1)

Connected Layer = $Z = W * h$ (2)

Where:

- h represents the input feature map, which could be an image or feature map from a previous layer.
- W represents the convolutional kernel or filter. These are small matrices used for feature extraction. During training, these kernels are learned to detect various patterns or features within the input data.
- Z represents the output feature map obtained by convolving the input feature map with the kernel.
- hxy represents the value at position (x,y)(x,y) in the output feature map after applying max pooling.
- h(x+i)(y+j)h(x+i)(y+j) represents the value at position (x+i,y+j)(x+i,y+j) in the input feature map.
- The notation max(i,j) denotes taking the maximum value over a specific region of the input feature map. Typically, this region is defined by a small window or kernel.

Stage 4: BiLSTM integration

In this stage, the extracted characteristics features (output) from the CNN are the input to the BiLSTM layer forward and backward network simultaneously which involves multiple gates and operations.

$$InputGate\ (i): [\ i_t\ = sigma(W\_\{ix\}\, x_t\ +\ W\_\{ih\}h\_\{t-1\}\ +\ b_i)\ ]$$

$$ForgetGate\ (f): [\ f_t\ =\ sigma(W\_\{fx\}x_t\ +\ W\_\{fh\}h\_\{t-1\}\ +\ b_f)\ ]$$

$$OutputGate\ (o): [\ o_t\ =\ sigma(W\_\{ox\}x_t\ +\ W\_\{oh\}h\_\{t-1\}\ +\ b_o)\ ]$$

$$CandidateGate\ (g): [\ g_t\ =\ tanh(W\_\{gx\}x_t\ +\ W\_\{gh\}h\_\{t-1\}\ +\ b_g)\ ]$$

$$HiddenState\ (h): [\ h_T\ =\ o_t odot tanh(c_t)];\ CellState: Ct = ft \odot Ct - 1 + it \odot gt$$

Where σ denotes the sigmoid activation function, which squashes the input values to be between 0 and 1. tanh is the hyperbolic tangent function, which outputs values between-1and1. ⊙ represents element-wise multiplication. X_t is the input at time step t, H_t is the hidden state at time step t, Sigma is the sigmoid activation function, Represents element-wise multiplication, W and b are the weight matrix and bias vector parameters.

Stage 5: **Combining BM25 with CNN-based BiLSTM (hybrid model)**

Multiply the BM25 weights with the word embeddings obtained from the CNN-BiLSTM mode

$$BM25(d,t)\ =\ IDF(t)\ *\ (tf(t,d)\ *\ (k1\ +\ 1))\ /\ (tf(t,d)\ +\ k1\ *\ (1\ - b\ +\ b * |d|\ /\ avgdl)) \quad (3)$$

$$C_t = \ f_t \circ C_{t-1} +\ i_t \circ \tanh(W_{xc} * X_t +\ W_{hc} * H_{t-1} + b_c) +\ BM25(d,t) \quad (4)$$

$$o_t\ =\ \sigma(W_{xo} * X_t +\ W_o * H_{t-1}\ + W_o \circ C_t +\ \ b_o) +\ BM25(d,t) \quad (5)$$

Where

$$IDF(t)\ =\ log((N - n\ +\ 0.5)\ /\ (n\ +\ 0.5)) \quad (6)$$

- N: $Totalnumberofdocuments$.
- n: Number of documents containing term t.
- BM25(d,t): BM25 score of document d for a given term t
- tf(t,d): Term Frequency of term tt in document d, representing how many times term t occurs in document d.
- k1: A parameter that controls the scaling of term frequencies. It's typically set empirically.
- b: Another parameter that controls the scaling of document length normalization. Also, set empirically.
- |d|: Length of document dd, often measured as the total number of terms.
- avgdl: Average document length in the document collection

Stage 6: Train the hybrid model based on Performance and Evaluation

It will evaluate the system against the other deep learning models. The system will calculate the relation of text in legal records and locate names by rank categories to display comprehension of the combined model and show a successful result on legal contract document analysis.

Program Model Specification

i. **Programming Languages:** Python program
ii. **Machine Learning Libraries**: TensorFlow, PyTorch, sci-kit-learn, or Keras for building and training machine learning models.
iii. **Data Processing Tools**: Pandas, NumPy, Seaborn, Matplotlib. Etc
iv. **Development Environments**: Jupyter Notebook, PyCharm, or Visual Studio

Hardware Requirement

i. High-performance CPU of Intel Core i5 or AMD Ryzen 5)
ii. **Random Access Memory (RAM) of** 8GB to 16GB of RAM and above
iii. 256GB SSD is typically sufficient, but larger capacities may be needed depending on the size of the dataset. The system flow chart is in figure 1.

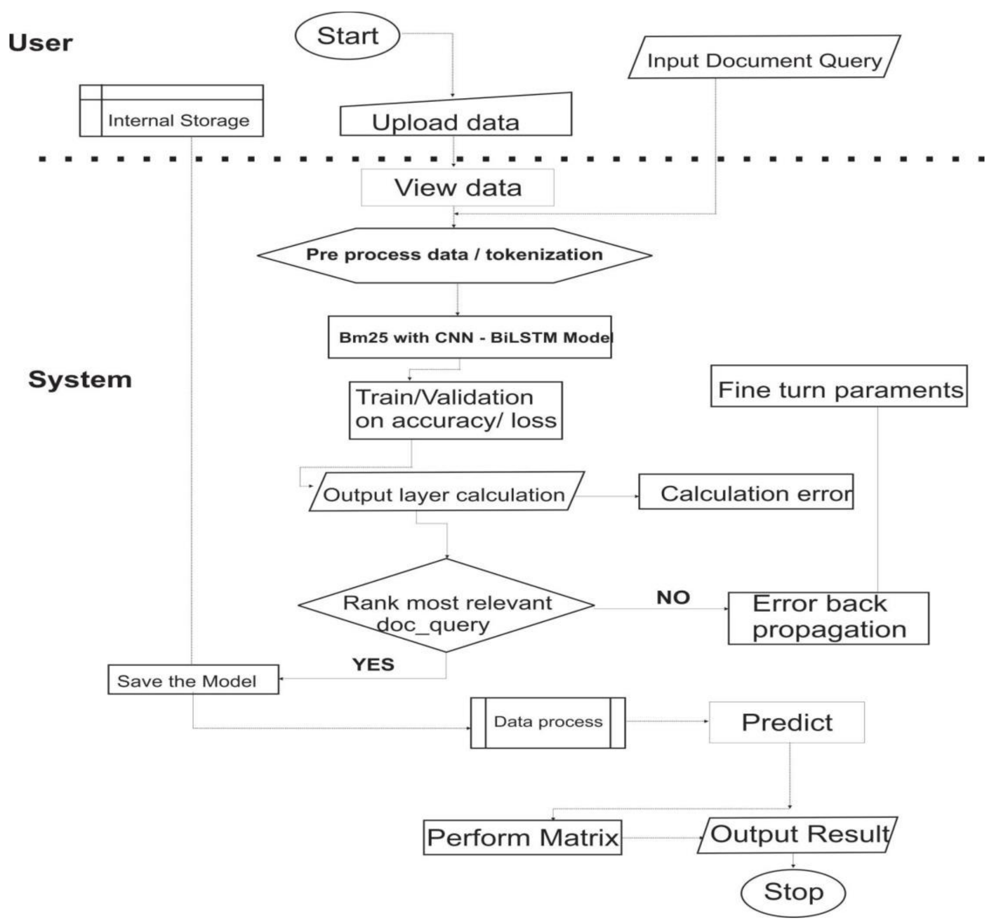

Figure 1: Swim-learn Diagram of the hybrid model

The Swim-lane Diagram of the hybrid model visualizes the process flow of the system which show the combination of different stages to achieve its goal

5. RESULTS AND DISCUSSION

The results demonstrate that integrating BM25 with CNN-BiLSTM in a hybrid model significantly enhances the performance of legal document retrieval systems in contract law cases compare to other deep learning models which is described in Table 1.

Table 1: A Comparative Predictive Function Based on Performance

The analysis of the results in table 1 reveals that the hybrid model, which integrates BM25,

Metrics	MLP	LSTM	CNN + BiLSTM
Precision	0.9487179487179488	0.8823529411764706	0.9722222222222222
Recall	0.9259259259259259	0.7777777777777777	0.9629629629629629
F1 – score	0.9305555555555555	0.7619047619047619	0.9658994032395567

Convolutional Neural Network (CNN), and Bidirectional Long Short-Term Memory (BiLSTM), outperforms the individual MLP and LSTM models across various evaluation metrics. The figure 2 presented the graphical analysis of the comparative models.

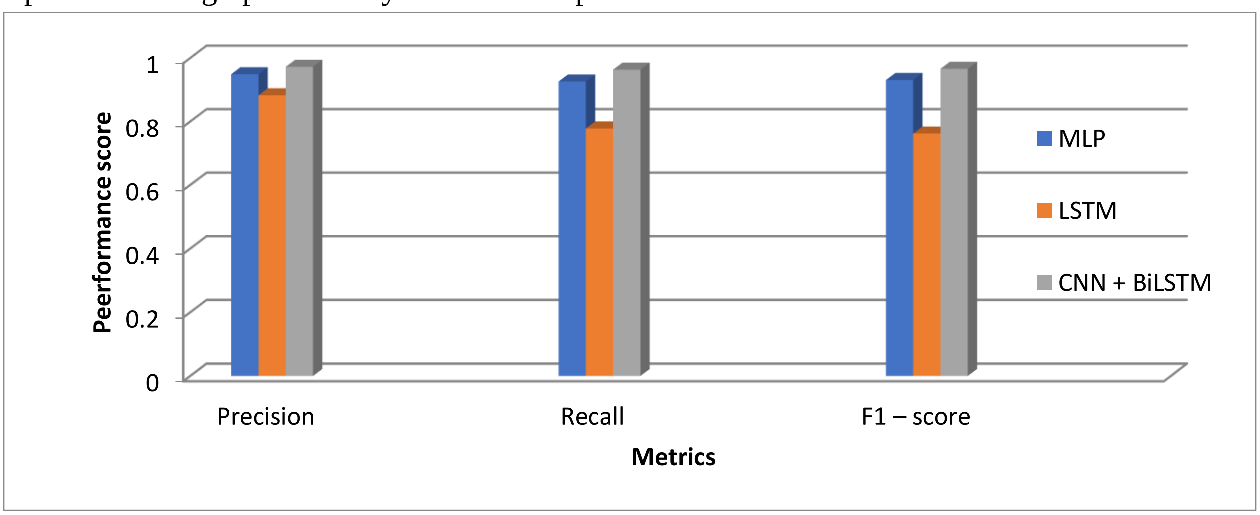

Figure 2: Graphical analysis of results

In the figure 2, precision, measures the accuracy of the model in identifying relevant documents, shows that the hybrid model achieves the highest precision of 0.9722, indicating it is the most effective at correctly identifying relevant documents while minimizing false positives. The MLP model follows with a precision of 0.9487, while the LSTM model has the lowest precision at 0.8824.

Recall, which measures the model's ability to retrieve all relevant documents, also demonstrates the superiority of the hybrid model. The hybrid model achieves the highest recall of 0.9630, meaning it retrieves most of the relevant documents in the dataset. The MLP model has a recall of 0.9259, and the LSTM model has the lowest recall at 0.7778. The F1-score, which provides a balanced measure of a model's performance by considering both false positives and false negatives, further highlights the advantages of the hybrid model. The hybrid model has the highest F1-score of 0.9659, indicating the best overall performance by effectively balancing precision and recall. The MLP model follows with an F1-score of 0.9306, while the LSTM model has the lowest F1-score at 0.7619. These results suggest that the proposed hybrid model, which leverages the strengths of BM25, CNN, and BiLSTM, is the most effective in accurately identifying and retrieving relevant legal documents in the context of contract litigation. The integration of these advanced techniques enhances the model's ability to capture the nuanced contextual information present in legal texts, leading to improved performance compared to the individual MLP and LSTM models. The figure 3 present the results of the average score for the model output for document score analysis.

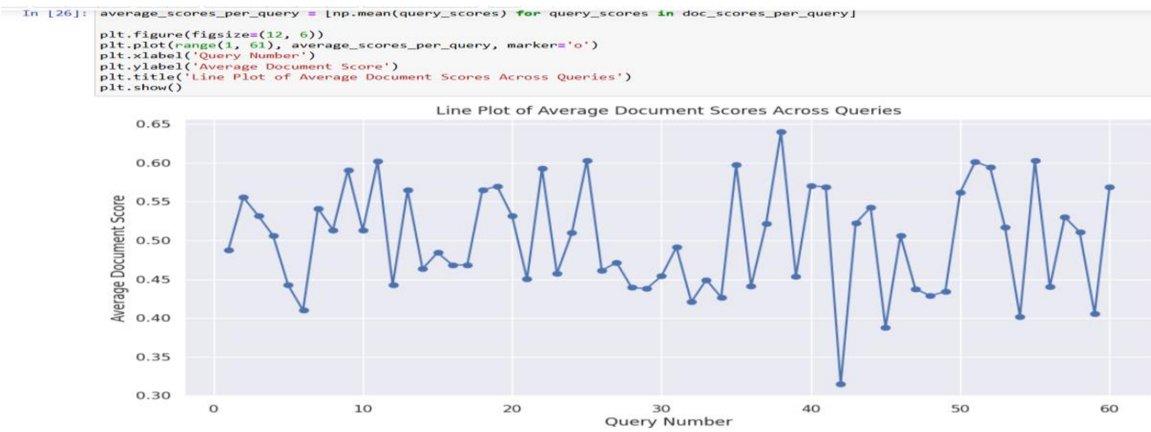

Figure 3: line plot of average doc-score_ query output.

The model's output, shown in Figure 3, is a dataset of 3257 files including previous case documents, some of which are relevant to the questions in task 1 and some of which are also helpful for task 2. This shows that query number 38 has the highest /most relevant document score on precedents/statutes cases out of the 60 queries uploaded, while query number 42 has the lowest relevant document within the text; this shows that the model captures semantic relationships and contextual information at optimal accuracy. The integration of these techniques offers a comprehensive solution that combines deep learning techniques with traditional keyword-based retrieval to improve the retrieval of legal documents.

5. SUMMARY

Based on the findings of the study, it is concluded that document analysis models combining BM25 with CNN-BiLSTM architectures offer substantial advantages in legal contract litigation. Using the frequency of query phrases in each document and their rarity throughout the document collection, the Best Matching 25 (BM25) determines a relevance score for each precedent & statute case document in a collection concerning a query. Convolution filters are applied across the spatial dimensions of the document by Convolutional Neural Networks (CNN) to extract hierarchical characteristics from input documents. To identify patterns and structures in the input data and make them useful for tasks, it makes use of shared weights and local connections. In addition the foundation of Bidirectional Long Short-Term Memory (BiLSTM) analysis of the forward and backward input sequences of the process. Which can record dependencies in both temporal directions, which is crucial for recording long-range dependencies. These models demonstrate superior accuracy, precision, and efficiency compared to traditional methods, leading to more effective contract analysis, case review, and legal research. Therefore, adopting these advanced models can significantly enhance the capabilities and performance of legal professionals in handling contract litigation cases. Further research and development efforts should focus on optimizing model parameters, expanding training datasets, and addressing domain-specific challenges to maximize the potential impact of these models in the legal domain.

6. CONCLUSION

Findings of this study suggest that intelligent legal contract models which combine BM25 with CNN-BiLSTM architectures hold great promise for revolutionizing legal contract litigation practices. By leveraging the synergies between traditional information retrieval techniques and deep learning methodologies, these models offer a powerful framework for enhancing the efficiency, accuracy, and effectiveness of legal document analysis, ultimately leading to improved outcomes for legal practitioners and stakeholders. The model's ability to identify relevant case law and legal precedents within contracts significantly enhances legal research efficiency, reducing the cost and time required for case review by 97%.

7. REFERENCES

Ahmad, S., Asghar, M. Z., Alotaibi, F. M., & Al-Otaibi, Y. D. (2022). A hybrid CNN + BILSTM deep learning-based DSS for efficient prediction of judicial case decisions. Expert Systems with Applications, 209, 118318. https://doi.org/10.1016/j.eswa.2022.118318

Cormack, G. V., & Grossman, M. R. (2014). Evaluation of machine-learning protocols for technology-assisted review in electronic discovery. https://doi.org/10.1145/2600428.2609601

Castano, S., Ferrara, A., Furiosi, E., Montanelli, S., Picascia, S., Riva, D., & Stefanetti, C. (2024). Enforcing legal information extraction through context-aware techniques: The ASKE approach. Computer Law and Security Report/Computer Law & Security Report, 52, 105903. https://doi.org/10.1016/j.clsr.2023.105903

Dyevre, A. (2021). Text-mining for Lawyers: How Machine Learning Techniques Can Advance our Understanding of Legal Discourse. Erasmus Law Review, 14(1). https://doi.org/10.5553/elr.000191

Francia, O. a. A., Núñez-Del-Prado, M., & Alatrista-Salas, H. (2022). Survey of Text Mining Techniques Applied to Judicial Decisions Prediction. Applied Sciences, 12(20), 10200. https://doi.org/10.3390/app122010200

Giwa, D. C., & Kodjovi, D. (2023). Artificial intelligence and the future of the Administration of Law and Justice in Nigeria. Social Science Research Network. https://doi.org/10.2139/ssrn.4566440

Gomede, E., PhD. (2023, September 2). Understanding the BM25 ranking Algorithm - Everton Gomede, PhD - medium. Medium. https://medium.com/@evertongomede/understanding-the-bm25-ranking-algorithm-19f6d45c6ce

Grimm, P. W. (2021). Artificial Intelligence as Evidence. Northwestern Pritzker School of Law Scholarly Commons. https://scholarlycommons.law.northwestern.edu/njtip/vol19/iss1/2/

Hassan, M. U. (2022). Technology Assisted Review of Legal Documents. RIT Scholar Works. https://scholarworks.rit.edu/theses/11395/

Ifeanyi-Reuben, N. J., & Benson-Emenike, M. E. (2018). An Efficient Feature Selection Model for Igbo Text. Social Science Research Network. https://doi.org/10.2139/ssrn.3389715

IR-LEGPREC. (2021, October 13). Kaggle. https://www.kaggle.com/datasets/sarthakjohnsonprasad/irlegprec/data

Jelali, S. E., Fersini, E., & Messina, E. (2015). Legal retrieval as support to eMediation: matching disputant's case and court decisions. Artificial Intelligence and Law, 23(1), 1–22. https://doi.org/10.1007/s10506-015-9162-1

Katz, D., Bommarito, M. J., & Blackman, J. (2017). A general approach for predicting the behaviour of the Supreme Court of the United States. PLOS ONE, 12(4), e0174698. https://doi.org/10.1371/journal.pone.0174698

Keeling, R., Chhatwal, R., Huber-Fliflet, N., Zhang, J., Wei, F., Zhao, H., Shi, Y., & Qin, H. (2019). Empirical Comparisons of CNN with Other Learning Algorithms for Text Classification in Legal Document Review. In arXiv (Cornell University). Cornell University. https://doi.org/10.1109/bigdata47090.2019.9006248

Locke, D., & Zuccon, G. (2022, February 15). Case law retrieval: problems, methods, challenges and evaluations in the last 20 years. arXiv.org. https://arxiv.org/abs/2202.07209

Millidge, B., Seth, A. K., & Buckley, C. L. (2021). Predictive Coding: a Theoretical and Experimental Review. ResearchGate. https://www.researchgate.net/publication/353510198_Predictive_Coding_a_Theoretical_and_Experimental_Review

Park, K., Hong, J. S., & Kim, W. (2020). A Methodology Combining Cosine Similarity with Classifier for Text Classification. Applied Artificial Intelligence, 34(5), 396–411. https://doi.org/10.1080/08839514.2020.1723868

Pisner, D., & Schnyer, D. M. (2020). Support vector machine. In Elsevier eBooks (pp. 101–121). Elsevier BV. https://doi.org/10.1016/b978-0-12-815739-8.00006-7

Rhanoui, M., Mikram, M., Yousfi, S., & Barzali, S. (2019). A CNN-BILSTM model for Document-Level Sentiment Analysis. Machine Learning and Knowledge Extraction, 1(3), 832–847. https://doi.org/10.3390/make1030048

Saxena, S. (2023, October 25). What is LSTM? Introduction to Long Short-Term Memory. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/#:~:text=Bidirectional%20LSTMs%20(Long%20Short%2DTerm,based%20on%20the%20preceding%20context.

Spoonmore, R. (2021). How can NLP using Cosine Similarity tell if a question is answered? A glimpse at analyzing an unlabeled dataset. www.linkedin.com. https://www.linkedin.com/pulse/how-can-nlp-using-cosine-similarity-tell-question-robert-spoonmore/

Sampath, K., & Thenmozhi, D. (2022a). PReLCaP : Precedence Retrieval from Legal Documents Using Catch Phrases. Neural Processing Letters/Neural Processing Letters, 54(5), 3873–3891. https://doi.org/10.1007/s11063-022-10791-z

Wei, F., Qin, H., Ye, S., & Zhao, H. (2018). Empirical Study of Deep Learning for Text Classification in Legal Document Review. https://doi.org/10.1109/bigdata.2018.8622157

Yang, E., MacAvaney, S., Lewis, D. D., & Frieder, O. (2021). Goldilocks: Just-Right Tuning of BERT for Technology-Assisted Review. ResearchGate.