Journal of Engineering Faculty, 2(2): 185-192, 2024

Sivas Cumhuriyet University
Journal of Engineering Faculty

| cumfad.cumhuriyet.edu.tr | Founded: 2023 Available online, ISSN Publisher: Sivas Cumhuriyet Üniversitesi

Bekenbey AI: Innovative Solutions at the Intersection of Deep Learning and Law

Erdicem Yucesan[1,a], Mehmet Ali Erkan[2,b], Ali Deveci[3,c], Ihsan Tolga Medenia[2,d]

[1]*Ankara Yıldırım Beyazıt University, Ankara,06760, Türkiye*
[2]*Middle East Technical University, Ankara, 06800, Türkiye*
[3]*Hacettepe University, Ankara,06800, Türkiye*

Research Article

History

Received: 25/11/2024
Accepted: 10/12/2024

ABSTRACT

This research introduces a cutting-edge integration of generative artificial intelligence (AI) within the realm of law, creating a sophisticated application tailored for legal professionals, organizations, and the public. The Bekenbey AI model show cased in this study is distinguished by its substantial potential, with key performance metrics such as accuracy, precision, recall, F1- score, ROUGE, and BLEU scores illustrating its adeptness at legal analytics. The model demonstrates exceptional precision and adaptability across various legal sectors and frameworks, establishing it as an indispensable asset for modern legal challenges. The findings suggest that the Bekenbey AI proficiently handles and interprets legal texts, significantly aiding the progression of legal systems. The model's efficiency escalates with the expansion of dataset sizes, emphasizing its capacity for extensive data analysis. Ongoing enhancements are focused on increasing the model's precision and extending its functionality to a wider array of legal contexts. To the best of our knowledge, this study represents the first instance of combining the domains of artificial intelligence and law using real data.

Keywords: Advanced Generative AI, Deep Learning Innovations, Bekenbey AI Model, AI Applications in Law, Legal Framework Analysis, Legal Document Examination

a erdicemyc@gmail.com ORCID
b mail ORCID
c ybilge@medicine.ankara.edu.tr ORCID
d mail ORCID

How to Cite: Yucesan E, Erkan MA, Deveci A, Medenia IT (2024) Bekenbey AI: Innovative Solutions at the Intersection of Deep Learning and Law, Journal of Engineering Faculty, 2(2): 185-192

Introduction

Deep Learning (DL), a sophisticated branch of machine learning, utilizes multi-layered neural networks to learn data representation through hierarchical levels of abstraction [1]. This advanced representation enhances proficiency in complex tasks including image recognition, natural language processing (NLP), and intricate data generation, thus broadening the functional scope of creative AI systems across diverse applications [2]. According to a study conducted in 2024, 70% of leading law firms in the United States predicted in 2023 that generative artificial intelligence would create value-added work for their clients. This percentage indicates a growing convergence between GenAI and the legal system [3].

The legal system, a structured regime of rules and principles enforced by institutions, governs societal behavior. AI's integration into this framework promises enhancements in legal analytics, predictive modeling, and document management through its capacity to analyze extensive datasets, recognize patterns, and facilitate complex decision-making processes. Particularly, DL models, equipped with advanced NLP capabilities, play a crucial role in interpreting and generating legal language with increasing precision and subtlety. These capabilities lead to substantial improvements in drafting legal documents, predicting case outcomes based on historical data, and conducting in-depth analyses of legal texts [4]. In recent years, the intersection of law and Generative AI has become a focal point for research, driven by the potential of AI to transform legal processes [5]. Moreover, researchers are increasingly focused on the use of AI for predictive legal analysis. By applying Deep Learning techniques, these systems can analyze large datasets of legal cases to identify trends, forecast outcomes, and suggest strategic approaches for legal practitioners. This predictive capability is especially valuable in litigation, where anticipating the likely outcome of a case can influence legal strategy and decision-making. This study explores the potential transformative impacts of integrating Artificial Intelligence (AI) technologies, particularly Generative

Artificial Intelligence (GAI) and Deep Learning (DL), into the legal sector. GAI, employing mechanisms such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), possesses the capability to generate new, high-quality data that accurately reflects the distribution of input datasets. This capability enables the production of realistic outputs in various formats—text, images, and audio—highlighting the versatility and expansive potential of creative models in data generation [6].

Bekenbey AI model aims to leverage law-based AI technologies to address prevalent challenges in the legal sector, offering innovative solutions that benefit legal professionals and the broader community[7]. The paper is organized as follows. In Section II, we provide background information on DL and ANN, introduce Legislation and Court, and give an overview of Integration of Artificial Intelligence into Legislative and Court Systems. Section III explores related studies in the literature. The proposed approach is detailed in Section IV. Section V and Section VI outlines the experimental settings and discusses the results. Finally, Section VII concludes the study.

Background

Deep Learning and Artificial Neural Networks

Deep Learning (DL) represents a fundamental domain within machine learning, distinguished by its utilization of Artificial Neural Networks (ANNs) comprising multiple layers [8]. One of the key strengths of DL lies in its capacity to autonomously extract and learn hierarchical representations from complex and large-scale and datasets. Unlike conventional machine learning models, which necessitate manual feature engineering by domain experts, DL models indicate a remarkable proficiency in automatically uncovering intricate patterns within data, particularly in the context of unstructured datasets such as audio, image, and textual information. This capability is most prominently exemplified by the success of Convolutional Neural Networks (CNNs), which have emerged as the benchmark for image recognition and various other tasks in computer vision. Similarly, Recurrent Neural Networks (RNNs) and transformers have essentially transformed the field of natural language processing, facilitating significant advancements in language modeling, machine translation, and other related areas [9]. Transformer-based models like BERT and GPT have advanced deep learning in NLP by using self-attention mechanisms to understand contextual information, overcoming the limitations of sequential processing in older architectures like RNNs. This improvement enhances performance in tasks such as language modeling, machine translation, and text generation. The evolution of deep learning and ANNs has also broadened AI applications, including in the legal industry, where AI technologies are enhancing efficiency and accuracy in legal research, document review, and contract analysis. AI systems can now rapidly sift through vast amounts of legal documents, extracting related information and providing lawyers with insights that would have taken much longer to achieve manually. This capability not only improves the speed of legal processes but also reduces costs for clients, making legal services more accessible [10].

Legislation and Court

Legislation and court systems are fundamental components of the legal framework that governs societal behavior and resolves disputes [11]. Legislation refers to the body of laws enacted by a legislative body, while courts are institutions responsible for adjudicating legal disputes based on these laws [12]. The evolution of these systems reflects the complexity and dynamism of legal governance. The advent of artificial intelligence (AI) presents new opportunities for enhancing legislative and judicial processes, though it also introduces challenges that must be addressed [13]. 1) Legislation: The process of enacting laws by a legislative body, which includes drafting, debating, and passing legal statutes [14]. Legislation serves as the foundation for legal norms and regulations governing various aspects of society. 2) Court: An institution responsible for adjudicating legal disputes, interpreting laws, and ensuring justice. Courts operate at various levels, including trial-, appellate-, and supreme-courts, each with specific roles and functions [15].

Related Work

The integration of AI into various domains has driven significant advancements in methodologies across sectors. The deployment of Generative AI (GAI) and Deep Learning models, such as Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and transformers, has notably transformed data-driven research. These models autonomously learn complex patterns in large datasets, enabling the generation of high-quality synthetic data that closely mirrors the original data.

The datasets employed in these studies often comprise extensive collections of unstructured data, such as text, images, and audio, sourced from diverse origins. For instance, transformer-based models have revolutionized the processing of textual datasets, including Common Crawl and OpenWeb-Text, in natural language processing tasks [16]. AI has significantly impacted the social sciences and humanities. Research by Binns has explored AI's ethical considerations, focusing on fairness and bias mitigation. In these fields, AI employs models like transformers and RNNs to analyze textual and behavioral data, providing insights into human behavior and social trends [17]. In the legal domain, AI's application has witnessed a marked increase, particularly in areas such as legal document generation, predictive analytics, and case outcome prediction. One of the seminal studies in the field of legal prediction is the research conducted by Katz, Bommarito, and Blackman (2017), which utilized advanced machine learning models to forecast the outcomes of Supreme Court cases with notable accuracy. Their study primarily employed a random forest classifier as the central machine learning technique for predicting case outcomes. The model achieved a remarkable accuracy rate of 71.9% at the individual Justice vote level over a historical period spanning from 1816 to 2015. Additionally, the model demonstrated an accuracy of 70.2% in predicting the overall outcomes of cases, underscoring the practical utility of AI in the domain of legal forecasting [18]. Predictive analytics in the legal field employs machine learning models to forecast legal case outcomes based on historical data, offering valuable insights to legal practitioners [19]. The automation of legal text creation, facilitated by advanced DL models, represents a significant breakthrough. For example, applying GANs in generating synthetic legal documents illustrates AI's potential to handle complex, structured data with high precision [20]. Abimbola, de La Cal Marin, and Tan (2024) explored deep learning in sentiment analysis for Canadian maritime case law, developing a framework to enhance legal analytics. Their study highlights the automation of legal document extraction and the integration of sentiment analysis with advanced deep learning models [21]. These advancements highlight AI's transformative potential in improving the efficiency and accuracy of legal processes. The Bekenbey AI project stands out as a pioneering initiative, integrating Generative AI (GAI) and Deep Learning (DL) models into a comprehensive system specifically designed for legal document generation and predictive legal analytics. Unlike previous studies that focused on general applications or single-task models, Bekenbey AI employs a hybrid approach, combining GANs, VAEs, and transformers to create contextually accurate legal documents and provide predictive insights for real-time legal strategies. To our knowledge, this is the first study that successfully integrates GANs, VAEs, and transformers into a unified hybrid model specifically tailored for the legal domain.

Proposed Model: Bekenbey AI

In an era where technology integration in law is crucial, our Bekenbey AI model represents a transformative approach. Utilizing advanced NLP and DL techniques, Bekenbey is designed to streamline complex legal processes and enhance legal document management and analysis [22]. This system integrates various NLP techniques, including tokenization, stopword removal, lemmatization, and stemming, to improve the parsing and comprehension of complex legal texts. It also utilizes text classification algorithms like Logistic Regression, SVM, and Random Forests to efficiently categorize legal documents across different domains [23]. The core functionality of Bekenbey AI is supported by advanced deep learning architectures, including Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and the innovative Bidirectional Encoder Representations from Transformers (BERT) model. These technologies are essential for processing, understanding and generating legal language and structures, enabling the system to perform deep analyses and generate outputs that closely resemble human legal reasoning. To further augment its capabilities, Bekenbey integrates Generative AI technologies, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) to create realistic and contextually appropriate legal documents.

This feature not only accelerates document creation but also provides unparalleled customization, adapting to specific case requirements or legal stipulations [24]. In addressing the challenges of data management within legal processes, the model utilizes both SQL and NoSQL databases, including PostgreSQL and MongoDB, ensuring robust and flexible data handling. Compliance with stringent security standards and privacy regulations, such as the General Data Protection Regulation (GDPR), is achieved through sophisticated encryption and anonymization techniques. The backend infrastructure of Bekenbey is meticulously designed using Python, with frameworks such as Django and Flask. APIs are developed using FastAPI to ensure efficient integration and communication within the legal tech ecosystem. This configuration is designed to meet the high demands for security and operational efficiency that are critical in legal applications. Moreover, the model significantly reduces the time and costs associated with traditional legal procedures while enhancing the accuracy and accessibility of legal services, marking a significant milestone in legal technology. It not only streamlines legal workflows but also provides powerful predictive analytics and decision-support tools, enhancing the ability of legal professionals to manage cases with greater efficacy and confidence [25].

Figure 1 illustrates the architecture of this comprehensive data processing and artificial intelligence modeling system. This system encompasses data collection, preprocessing, various modeling layers and output processes. In the Data Collection section, data is systematically gathered from diverse sources, including legal databases, government and corporate websites, academic and research resources, digital libraries, and archives. During the Data Preprocessing stage, the collected data undergoes cleaning, tokenization, stemming, and vectorization [26]. Once transferred to the database server, the data is processed through various embedding layers, employing techniques such as Word Embeddings, Word2Vec and BERT. The CNN Layer and RNN/LSTM Layer process the text using deep learning techniques such as ReLU, pooling layers, and GRU. The Transformer Layer handles tasks such as text translation and summarization using advanced techniques like the attention mechanism and multi-head attention. The Classification Layer classifies the data utilizing densely connected layers, softmax, and cross-entropy loss techniques. Finally, the Output Layer encompasses model training, validation, and deployment processes. The system is optimized with AI Ring nodes supported by load balancing and message queuing mechanisms [27] [28][29] [30].

Experiments

Dataset and Preprocessing

The experiments utilized datasets compiled from multiple sources, including legal databases, government and corporate websites, academic resources, and digital libraries. The datasets comprised a mix of structured data (e.g., legal codes, statutes) and unstructured data (e.g., case law texts, legal opinions). All datasets, anonymized by the Torun Law and Consulting, were subjected to a comprehensive preprocessing pipeline that included cleaning, tokenization, stemming, and vectorization. This step ensured that the data was uniformly formatted and suitable for input into the model.

Experimental Setup

The model was implemented in Python, utilizing TensorFlow and PyTorch for deep learning, along with NLTK and SpaCy for natural language processing tasks. The setup also incorporated tools for hyperparameter tuning and performance monitoring. The model architecture included embedding layers (Word2Vec, BERT), CNN, RNN/LSTM, and Transformer layers. Each was carefully configured to optimize performance for tasks like text classification, summarization, and translation. The experiments involved tuning key hyperparameters, including learning rate, batch size, number of epochs, and embedding dimensions. A combination of grid search and random search techniques was employed to identify the optimal settings for each hyperparameter. The model was trained on a diverse set of legal documents, with training data split into training, validation, and test sets. The training process involved iterative updates of model weights using backpropagation, with loss minimization achieved through the use of cross-entropy loss for classification tasks. Model performance was validated at regular intervals using the validation set. The validation process involved calculating accuracy, precision, recall, and F1-score metrics.

Performance Metrics

Accuracy, Precision, Recall and F1-Score:

Accuracy measures a model's overall correctness by calculating the proportion of true results (both true positives and true negatives) among the total number of cases. For classifying legal documents, Precision, Recall, and F1-Score are crucial, especially with imbalanced datasets. Precision evaluates the proportion of correctly classified documents among predicted positives, minimizing false positives. Recall measures the model's ability to identify all relevant documents, reducing missed critical information. F1-Score, the harmonic mean of Precision and Recall, balances false positives and false negatives, making it particularly useful in legal document classification, where both over-inclusion and under-inclusion can have significant consequences.

ROUGE and BLEU Scores:

For text summarization tasks, the quality of the generated summaries was evaluated using ROUGE and BLEU scores, which compared the generated outputs to reference summaries

Computation Time and Memory Usage:

The scalability of the model was assessed by measuring computation time and memory usage across different dataset sizes, providing insights into the model's efficiency.

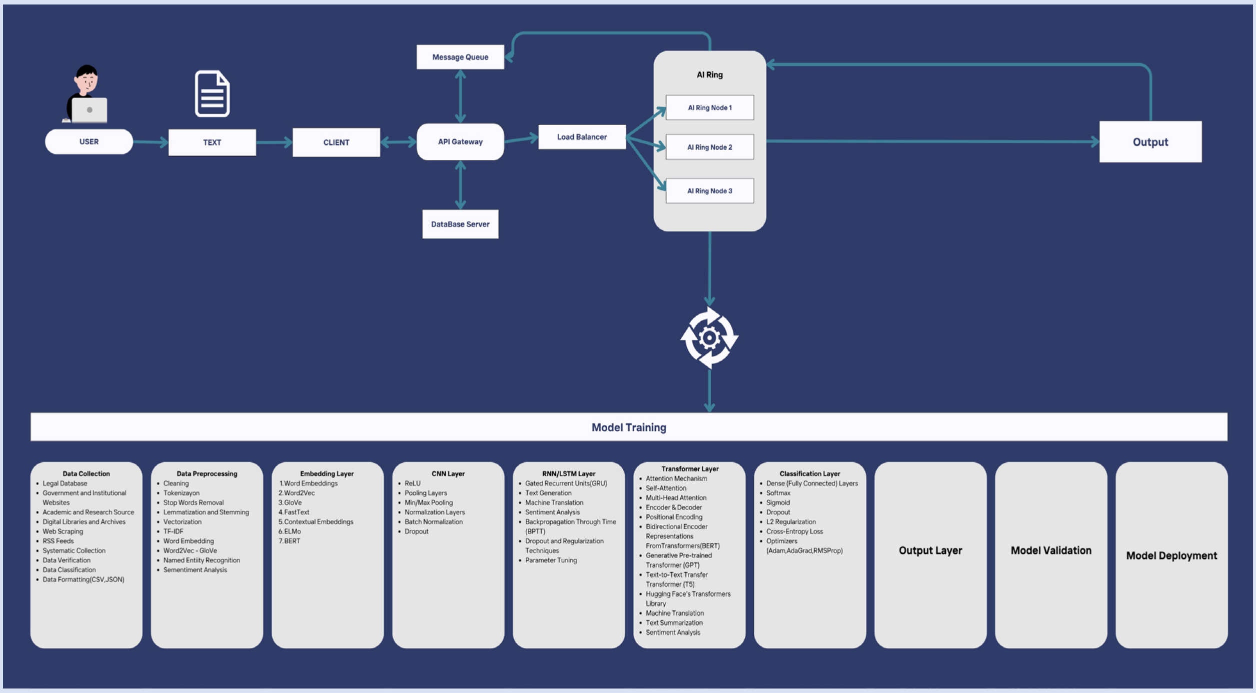

Fig. 1. Flow chart of the Bekenbey AI model.

Experimental Results

In this study, the performance of the proposed model was thoroughly evaluated using a comprehensive set of metrics, including accuracy, precision, recall, and F1-score. These metrics were calculated across different sample sizes, allowingfor an in-depth analysis of the model's behavior as the amount of training data increased. Figure 2 and Figure 3 illustrates the accuracy performance, and the computational time and memory usage (for ten different process) of the Bekenbey AI model respectively, while Table I presents its performance in terms of precision, recall, and F1-score. (Here, NoS: Number of Samples) Initially, the accuracy rate achieved with 4 samples was 45.65%, which increased to 88.73% with 50 samples.

This rise in accuracy highlights the significant improvement in the model's performance when trained with more data. Notably, a marked increase in accuracy was observed after 20 samples, indicating that the model possesses a higher learning capacity beyond a certain amount of data, resulting in more accurate predictions. An examination of the data revealed a significant increase in accuracy rates as the sample size increased. For instance, the accuracy rate obtained with 4 samples was 45.65%, which rose to 53.78% with 10 samples and further to 70.01% with 20 samples. These increases clearly demonstrate the dependency of accuracy on the sample size. The correlation between accuracy rate and sample size was calculated using the Pearson correlation coefficient. The resulting coefficient, $r = 0.957$, indicates a very strong positive correlation between these two variables. This finding confirms that increasing the sample size significantly enhances the model's accuracy. Table II provides a detailed breakdown of the performance of the Bekenbey AI model in varying sample sizes (NoS: Number of Samples), as measured by the ROUGE-1(R-1), ROUGE-2(R-2), ROUGE-L(R-L) and BLEU score. The results demonstrate a consistently high level of performance across all metrics, indicating that the model effectively captures both lexical and syntactic features of the target text.

Precision also demonstrated positive trends with increasing sample sizes. Initially, the model exhibited lower precision with smaller datasets due to a higher proportion of false positives. For example, with 4 samples, the precision was 46%, indicating that nearly half of the positive predictions were incorrect. However, as the dataset grew, precision improved significantly, reaching 89% with 50 samples, suggesting enhanced reliability in positive predictions. Recall showed a similar improvement. At the smallest sample size of 4, recall was 45.00%, reflecting the model's initial difficulty in identifying all positive instances. As the sample size increased, recall steadily improved, reaching 88.00% with 50 samples. This increase indicates that the model becomes more effective at capturing all relevant data points as the training set grows. The F1-score, starting at 45.50% with 4 samples, also improved with more data, reaching 88.50% at 50 samples. The rise in the F1-score, alongside improvements in precision and recall, highlights the model's growing effectiveness in balancing the identification of positive instances and minimizing false positives as more data is processed.

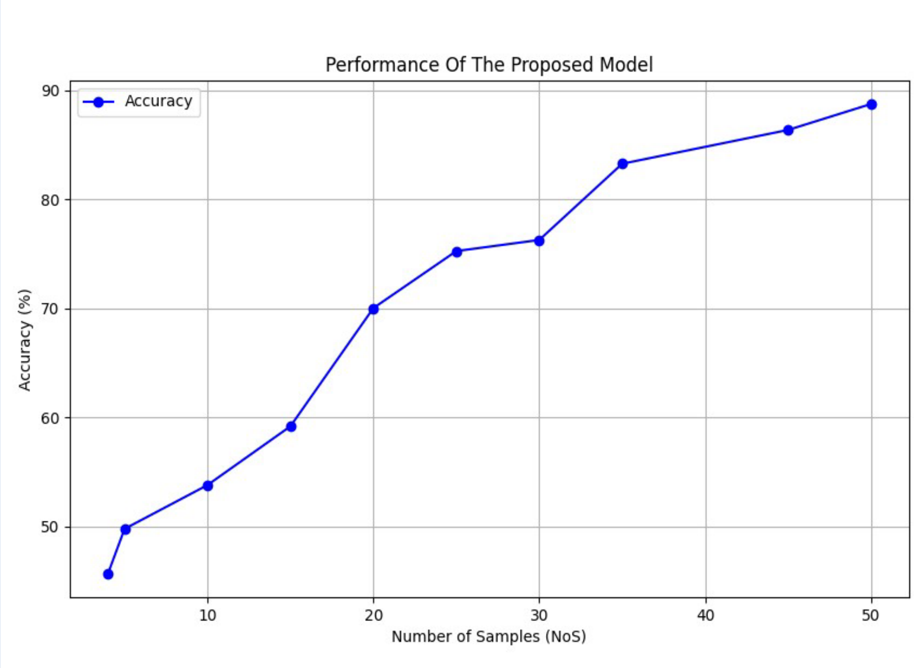

Fig. 2. The performance of Bekenbey AI model in terms of accuracy.

Table I Precision, Recall, and F1-Score Metrics for Different Sample Sizes

NoS	R-1(%)	R-2(%)	R-L(%)	BLEU(%)
4	85.50	78.30	84.20	76.50
5	87.00	80.10	85.80	78.20
10	89.30	83.50	88.20	81.60
15	91.50	86.20	90.10	84.50
20	93.00	88.00	91.70	86.80
25	94.50	89.80	93.20	88.90
30	95.20	90.70	94.00	89.80
35	96.50	92.30	95.30	91.50
45	97.00	93.00	96.00	92.30
50	97.50	93.80	96.50	93.00

Table 2 Rouge and Bleu Scores for Different Sample Sizes

NoS	Precision (%)	Recall (%)	F1-Score (%)
4	46.00	45.00	45.50
5	50.00	49.00	49.50
10	54.00	53.00	53.50
15	60.00	59.00	59.50
20	71.00	70.00	70.50
25	76.00	75.00	75.50
30	77.00	76.00	76.50
35	84.00	83.00	83.50
45	87.00	86.00	86.50
50	89.00	88.00	88.50

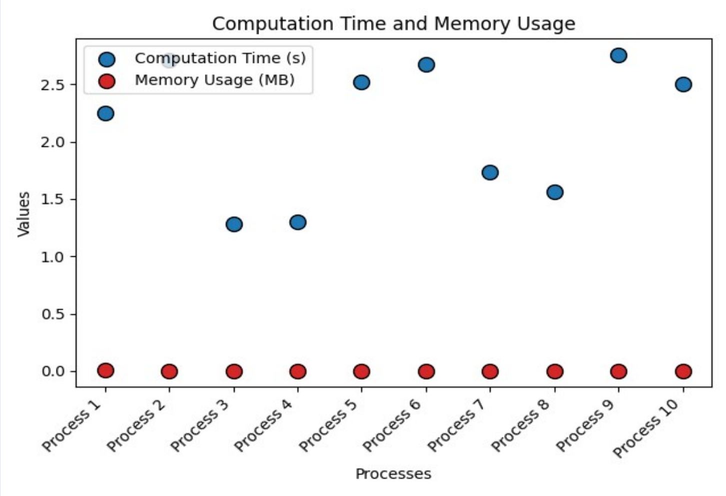

Fig. 3. Computation time and memory usage.

While the computational time varies depending on the content and length of the legal document being analyzed, there is no significant change observed in memory usage. This outcome is not unexpected, as text comparison inherently incurs a temporal cost, but it does not result in a change in the memory allocated within the same window size during a given unit of time. The ROUGE-1 scores, reflecting unigram matches, start at 85.5% for a sample size of 4 and increase to 97.5% as the sample size reaches 50, indicating improved alignmentwith the reference text as data exposure grows. ROUGE-2 scores, accounting for bigram matches, rise from 78.3% to 93.8%, highlighting the model's ability to maintain contextual coherence and phrase-level dependencies, crucial for fluent text generation. ROUGE-L scores, which measure the longest common subsequence, show a similar trend, starting at 84.2% and reaching 96.5%, indicating the model's effectiveness in preserving the reference text's structure. BLEU scores, which assess text similarity across various n-grams, also improve from 76.5% to 93%, confirming the model's strong performance in generating accurate and contextually appropriate text.

This study demonstrates that the proposed model exhibitshigh performance with large datasets and that accuracy improves with an increasing number of samples. These findings indicate that the model could be an effective tool for large-scale data analysis and applications. However, achieving higher accuracy rates and enhancing the model's overall performance will require the evaluation of additional strategies and improvements. Future work should aim to further enhance the model's performance and expand its applicability across a broader spectrum of applications.

Conclusion

In this study, we integrated generative AI with legal domains to develop an application serving citizens, organizations, and legal professionals. The proposed Bekenbey AI model demonstrates notable performance in accuracy, precision, recall, F1-score, ROUGE, and BLEU scores, highlighting its potential to enhance legal systems. The model's strong performance across various metrics indicates its effectiveness in processing and analyzing legal texts. Its ability to generalize across different legal fields and adapt to various legal systems further enhances its practical utility and relevance to current legal challenges. To our knowledge, this is the first study to integrate GANs, VAEs, and transformers into a unified hybrid model for the legal domain, using real-world data. Our future work(s) will focus on improving the model's capabilities by: i) Analyzing different generative models in legal contexts to identify the most effective approaches. ii) Conducting comparative analyses to evaluate the strengths and limitations of various models. iii) Testing the proposed model on different datasets and application domains to assess performance and adaptability. iv) Exploring advanced techniques and strategies to enhance accuracy and overall performance.

References

[1] T. Räz and C. Beisbart, "The importance of understanding deep learning,"Erkenntnis, vol. 89, no. 5, pp. 1823–1840, 2024.

[2] "Reward Value-Based Goal Selection for Agents' Cooperative Route Learning Without Communication in Reward and Goal Dynamism – SN Computer Science — link.springer.com." https://link.springer.com/article/ 10.1007/s42979-020-00191-2. [Accessed 22-08-2024].

[3]"GenAI: predicted impact at top U.S. law firms, statista.com." https://www.statista.com/statistics/1456037/ genai-predicted-impact-at-top-us-law-firms/, 2024. [Accessed 20-08-2024].

[4]“Balancing the scale: navigating ethical and practical challenges of artificial intelligence (AI) integration in legal practices - Discover Artificial Intelligence — link.springer.com.” https://link.springer.com/article/10.1007/s44163-024-00121-8. [Accessed 22-08-2024].

[5] C. V. Chien and M. Kim, “Generative ai and legal aid: Results from a field study and 100 use cases to bridge the access to justice gap,” Loyola of Los Angeles Law Review, forthcoming, 2024.

[6] “Generative Legal Minds - Harvard Law School Center on the Legal Profession — clp.law.harvard.edu.” https://clp.law.harvard.edu/knowledge-hub/magazine/issues/generative-ai-in-the-legal-profession/ generative-legal-minds/. [Accessed 22-08-2024].

[7]“AI and law: ethical, legal, and socio-political implications AI&SOCIETY-link.springer.com.” https://link.springer.com/article/10.1007/s00146- 021-01194-0. [Accessed 22-08-2024].

[8] O. A. Montesinos López, A. Montesinos López, and J. Crossa, “Fundamentals of artificial neural networks and deep learning,” in Multivariate statistical machine learning methods for genomic prediction, pp. 379–425, Springer, 2022.

[9] W. Q. Yan, “Convolutional neural networks and recurrent neural networks,” in Computational Methods for Deep Learning: Theory, Algorithms, and Implementations, pp. 69–124, Springer, 2023.

[10] M. Kauffman and M. Soares, “Ai in legal services: New trends in ai-enabled legal services. soca, 14, 223–226,” 2020.

[11] S. Kumari, “From chaos to order: The role of law in society,” Indian Journal of Integrated Research in Law Volume IV Issue II| ISSN, vol. 2583,p. 0538, 2024.

[12] L. Kern, H. P. George, L. L. Evanovich, J. M. Hodnett, and J. Freeman, “A review of us policy guidance and legislation on restraint and seclusion in schools: Considerations for improvement,” Exceptional Children, p. 00144029241247032, 2024.

[13] Y. K. Dwivedi, L. Hughes, E. Ismagilova, G. Aarts, C. Coombs, T. Crick, Y. Duan, R. Dwivedi, J. Edwards, A. Eirug, et al., “Artificial intelligence (ai): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy,” International journal of information management, vol. 57, p. 101994, 2021.

[14] F. H. Easterbrook, “What does legislative history tell us,” Chi.-Kent L. Rev., vol. 66, p. 441, 1990.

[15] R. D. González, V. F. Vásquez, and H. Mikkelson, “Fundamentals of court interpretation,” Theory, policy, and practice, 1991.

[16] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al., “Language models are unsupervised multitask learners,” OpenAI blog, vol. 1, no. 8, p. 9, 2019.

[17] R. Binns, “Fairness in machine learning: Lessons from political philosophy,” in Conference on fairness, accountability and transparency, pp. 149–159, PMLR, 2018.

[18] D. M. Katz, M. J. Bommarito, and J. Blackman, “A general approach for predicting the behavior of the supreme court of the united states,” PloS one, vol. 12, no. 4, p. e0174698, 2017.

[19] N. Aletras, D. Tsarapatsanis, D. Preo¸tiuc-Pietro, and V. Lampos, “Predicting judicial decisions of the european court of human rights: A natural language processing perspective,” PeerJ computer science, vol. 2, p. e93, 2016.

[20] Q. Chen, Q. Wu, J. Chen, Q. Wu, A. van den Hengel, and M. Tan, “Scripted video generation with a bottom-up generative adversarial network,” IEEE Transactions on Image Processing, vol. 29, pp. 7454–7467, 2020.

[21] B. Abimbola, E. D. L. C. Marin, and Q. Tan, “Enhancing legal sentiment analysis: A cnn-lstm document-level model,” Preprints, 2024.

[22] “Introduction for artificial intelligence and law: special issue “natural language processing for legal texts” - Artificial Intelligence and Law — link.springer.com.” https://link.springer.com/article/10.1007/s10506-019-09251-2. [Accessed 22-08-2024].

[23] “Evaluating Text Classification in the Legal Domain Using BERT Embeddings — link.springer.com.” https://link.springer.com/chapter/10.1007/ 978-3-031-48232-8_6. [Accessed 22-08-2024].

[24] “A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU — arxiv.org.” https://arxiv.org/abs/ 2305.17473. [Accessed 22-08-2024].

[25] “Solving NoSQL database governance & compliance with CI/CD — liquibase.com.” https://www.liquibase.com/blog/solving-nosql-database-governance-and-compliance-challenges-with-ci-cd.[Accessed 22-08-2024].

[26] “Big data preprocessing: methods and prospects - Big Data Analytics —bdataanalytics.biomedcentral.com.” https://bdataanalytics.biomedcentral.com/articles/10.1186/s41044-016-0014-0. [Accessed 22-08-2024].

[27] “Papers with Code - Enriching Word Vectors with Subword Information — paperswithcode.com.” https://paperswithcode.com/paper/ enriching-word-vectors-with-subword. [Accessed 22-08-2024].

[28] “Deep CNN-LSTM with Word Embeddings for News Headline Sarcasm Detection — link.springer.com.” https://link.springer.com/chapter/10.1007/978-3-030-14070-0_69. [Accessed 22-08-2024].

[29] “Language Models and Contextualised Word Embeddings -davidsbatista.net.” https://www.davidsbatista.net/blog/2018/12/06/Word_Embeddings/. [Accessed 22-08-2024].

[30] “What Are Word Embeddings? | IBM — ibm.com.” https://www.ibm.com/topics/word-embeddings. [Accessed 22-08-2024].