

Generative Artificial Intelligence in Legal Drafting

Thippaluri Yahid Basha

Department of Information Technology
Velagapudi Ramakrishna Siddhartha
Engineering College
Vijayawada, India
yahidbashat@gmail.com

Barma Kalyani

Department of Information Technology
Velagapudi Ramakrishna Siddhartha
Engineering College
Vijayawada, India
kalyanibarma2@gmail.com

Yelisetti Sandeep

Department of Information Technology
Velagapudi Ramakrishna Siddhartha
Engineering College
Vijayawada, India
y.sandeepit@vrsiddhartha.ac.in

Abstract—"Lexi" is a light of clarity in a world where legal complexity frequently makes comprehension difficult. It is a tool that uses the revolutionary potential of generative Artificial Intelligence (AI) to completely change the process of producing legal documents. To simplify legal language and improve the accessibility and comprehension of legal documents, this paper introduces Lexi, a revolutionary tool. Through the integration of cutting-edge AI technology, Lexi not only improves legal drafting productivity but also supports legal communications that are clear and understandable. This innovation marks a paradigm change in legal documentation by emphasizing readability and ease of use and opening the door to a more diverse legal environment.

Keywords—Generative Artificial Intelligence, Legal Drafting, AI-powered Tools, Legal Jargon Simplification, Fine-Tuning, Prompt Engineering, NLP.

I. INTRODUCTION

A maze of intricacy and ineffectiveness afflicts the current state of legal document writing. This conventional method, which is based on extensive resources and has a strong foundation in specialist expertise, becomes time-consuming and very expensive. These difficulties point to the urgent need for innovation—a revolutionary strategy that may simplify these complex procedures and increase the effectiveness, affordability, and accessibility of legal services for a larger range of people.

This is where the generative AI's promise shines through as a sign of advancement in the legal industry. The ability to use this cutting-edge technology might lead to a revolution in the way that legal writing is done. In addition to helping lawyers work more efficiently, artificial intelligence's capacity to automate and improve the preparation of legal documents marks the beginning of the democratization of legal services. Beyond the realm of legal specialists, this technology provides a useful tool for enterprises, entrepreneurs, and even individuals, enabling them to handle legal papers with more confidence and simplicity. The use of AI in legal drafting is a big step toward lowering barriers that have long stood between the public and legal knowledge and making the law more understandable and accessible.

The widespread use of intricate legal language, which frequently obscures the genuine meaning and intent of legal documents, is a significant obstacle in this field. This language barrier alienates those without a legal background and causes widespread misunderstandings. Artificial intelligence (AI)-driven solutions such as Lexi become essential in solving this

problem. Their design aims to streamline the drafting process and, more crucially, convert complex legal jargon into easily comprehensible phrases for the public. This method is essential for improving the accessibility and comprehensibility of legal materials and promoting a more diverse knowledge of legal issues.

Lexi stands as a testament to the power of Artificial Intelligence in bridging the gap between complex legal language and the need for clarity in legal communication, propelling the legal field into a new era of transparency and inclusivity.

A. Problem Statement

In the current legal landscape, the drafting of legal documents is a complex, time-intensive, and costly process, often laden with intricate jargon inaccessible to non-specialists. This situation makes legal texts more difficult to understand and accessible, which can result in inefficiencies and possibly misinterpretations. The goal of this work is to harness the capabilities of generative AI to streamline the legal drafting process, making it more efficient and accessible. By translating difficult legal terminology into easier, more intelligible words, this initiative intends to democratize legal paperwork, guaranteeing it is not only more efficient for legal experts but also more understandable for the general public. The objectives of this work include:

1) Developing an AI-powered tool

Putting together a cutting-edge AI system that can expedite and simplify the preparation of legal documents.

2) Legal Jargon Simplification

To make legal papers more accessible to non-lawyers, complicated legal concepts and jargon are simplified into language that is simple to grasp.

3) Customization of Legal papers

Developing capabilities for the AI tool to modify legal papers according to unique user requirements, assuring relevance and correctness.

4) Enhancing Accessibility

Making legal documentation more user-friendly and accessible, particularly for individuals and small businesses who may lack legal expertise or resources by making this tool

easy to access and use, through better user experience and interaction principles based design of website for the users.

II. BACKGROUND ESSENTIALS

A. Artificial Intelligence

Artificial Intelligence (AI) is a multifaceted field of computer science dedicated to creating systems capable of performing tasks that typically require human intelligence. These tasks include learning, reasoning, problem-solving, perception, and language understanding. AI is characterized by its ability to mimic cognitive functions, adapt to new inputs, and improve from experience. It spans various approaches and technologies, such as machine learning, neural networks, and deep learning, significantly impacting numerous sectors including healthcare, finance, and, notably, legal drafting, where it offers innovative solutions to complex challenges.

B. Generative AI

A subset of artificial intelligence technology known as "generative AI" is capable of creating new material by learning from preexisting data, such as text, photos, music, and code. Generative Adversarial Networks, Variational Autoencoders are two deep learning approaches it uses to produce new, high-quality outputs that can approximate human creativity. This technology can automate and reinvent procedures by generating legal papers, changing conventional ways and improving efficiency and accessibility. It has significant ramifications for several sectors, including legal drafting.

C. Natural Language Processing

A subfield of artificial intelligence called natural language processing (NLP) is concerned with how computers and human languages interact. To enable computers to comprehend, interpret, and produce meaningful human language, it entails programming them to process and analyze vast volumes of natural language data. NLP makes a wide range of applications possible, including sentiment analysis, language translation, and information extraction. It does this by fusing computational linguistics with machine learning and deep learning models. NLP is essential for automating and improving document generation and analysis procedures in legal drafting.

D. Fine-Tuning

In the context of artificial intelligence and machine learning, fine-tuning refers to the process of further training or "tuning" a pre-trained model on a smaller, more focused dataset after it has previously been trained on a larger, more diversified dataset. The model may adjust and specialize its knowledge to specific activities or domains, such as legal language in document drafting, thanks to this additional training step. By improving the model's accuracy and efficiency in addressing domain-specific problems, fine-tuning makes it more applicable and useful for the intended applications.

E. Prompt Engineering

In the realm of artificial intelligence, prompt engineering is essential, particularly for models that produce content based on

inputs. It entails carefully planning and refining input prompts to successfully direct AI models—such as chatbots and generative AI—to generate desired results. This method is essential for improving the precision, applicability, and caliber of AI-generated responses, and it is widely used in applications that call for contextual responses and sophisticated understanding, including legal AI systems that provide legal advice or write papers. In certain fields, the performance and applicability of AI models can be greatly impacted by efficient rapid engineering.

III. RELATED WORK

This section meticulously examines various studies and advancements in the realm of AI applications in legal documentation. It draws upon the knowledge and findings of previous research to establish a comprehensive understanding of the current state of AI in legal drafting. These references not only validate the necessity and relevance of Lexi but also highlight the evolution and potential future directions of AI in transforming legal documentation processes. This careful examination of related work not only anchors Lexi within the broader scientific dialogue but also elucidates the gaps and opportunities that this project aims to address.

Legal-BERT [1]: The adaption of the BERT model for legal domain NLP problems are explored in the paper "Legal-BERT: The Muppets Straight Out of Law School"[1]. It methodically looks at three approaches to applying BERT in specialist fields such as law: using the original BERT, doing additional pre-training on domain-specific corpora, and starting from scratch when pre-training BERT on legal texts. The study shows that using BERT directly is not as effective as pre-training further or pre-training from scratch on domain-specific corpora. The culmination of this study is the release of BERT with a legal language tuning called Legal-BERT, which provides insightful information about domain adaptation for specialized fields.[1]

Natural Language Processing in the Legal Domain [2]: With approximately six hundred NLP and law papers published in the last ten years, the paper "Natural Language Processing in the Legal Domain [2]" offers a thorough review of them all. Major trends in the field are highlighted by the report, including rising paper counts, task counts, and language coverage. It observes a rise in methodological sophistication as well as compliance with criteria for data accessibility and code replication. The authors compiled an almost comprehensive collection of pertinent publications that documented the advancements and difficulties in Legal NLP. This highlighted the increasing methodological convergence of Legal NLP with general NLP and the professional standards of the larger scientific community.[2]

Gracenote.ai: Legal Generative AI for Regulatory Compliance [3], explores the use of large language models (LLMs) in regulatory compliance and law, with a particular emphasis on creating AI tools such as an expert system based on LLMs, a tool for generating obligations, and a tool for horizon scanning. The method improves accuracy and lowers hallucinations in AI-generated legal content by combining the LangChain framework, GPT-4 text embeddings, and rapid

engineering. The article lays the foundation for future legal and compliance applications by emphasizing LLMs as basic engines in specialized tools. Subsequent investigations will focus on improving capabilities in many languages and jurisdictions, enhancing legal reasoning, and creating resilient LLM-based systems.[3]

Jennifer Somali Angeyo and Dr. Peter Jehopio's study "An Artificial Neural Network Model for Regulating Legislative Drafting Practices in Uganda"[4] describes a novel method of drafting laws with artificial neural networks (ANNs). The study examines how environmental factors, such as social, political, economic, and cultural elements, affect legislation drafting with a particular focus on Uganda. It uses artificial neural networks (ANNs) to forecast if legislation will succeed and provides insights into how these elements might affect the legislative process. To improve the efficacy and efficiency of the legislative process, this strategy represents a novel junction of technology and legislative processes.[4]

IV. ARCHITECTURE

The chapter titled "Architecture Diagram" provides an organized model of our AI-driven legal drafting system, outlining the complex network of interconnected parts that make up the entire system given in Fig. 1. This section will dissect the architecture and look at how each module fits into the overall operation of the system to make sure that readers have a thorough grasp of its workings.

A. User Interaction Layer

With a safe authentication process, this crucial layer acts as a portal for both current and potential users to engage with the system. Through a registration process, new users are guided to ensure that every contact is secure and personalized. Returning users can enjoy a personalized experience that remembers their past interactions and preferences thanks to the login module, which optimizes the system's relevance and effectiveness for specific legal drafting requirements.

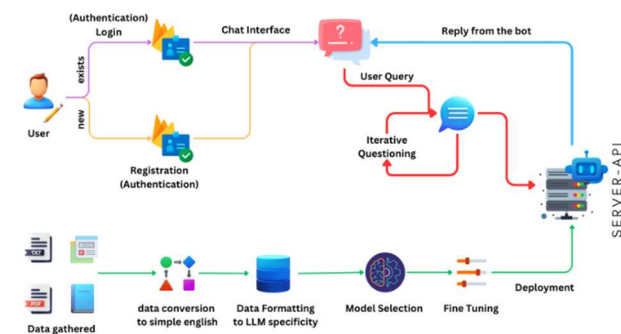


Fig. 1. Architecture Diagram

B. Chat Interaction

This part serves as the conversational front end of the system, allowing users to express their questions. The chat interface's elegant and user-friendly design makes use of advanced UI/UX principles to promote a smooth flow of conversation. It acts as a bridge between the user and the AI's processing powers by capturing user inputs, which might vary

from straightforward questions to intricate demands for legal drafting.

C. Iterative Questioning Mechanism

An advanced algorithm that functions as the system's clarifying agent, this mechanism engages users in an iterative questioning process. By probing for additional details, it refines the user's initial query, ensuring the AI understands the context and nuances of the request. This iterative process is vital for reducing ambiguity and honing the user's precise legal drafting needs.

D. Data Handling and Conversion

At this juncture, the system processes the collected data, which may include existing legal documents and user queries. It translates this data into simplified English and formats it to match the specific requirements of the language model, thereby setting the stage for effective AI processing. This conversion is essential for ensuring that the AI's responses are both accurate and accessible to users without legal expertise.

E. AI and Machine Learning Components

The heart of the system is where the AI's capabilities are defined and enhanced. Model selection involves choosing the appropriate AI framework that is best suited for legal language processing, while fine-tuning is an iterative process of adjusting the model's parameters to the legal drafting domain. This ensures the system's output is not only precise but also contextually relevant to the field of law.

F. Deployment Infrastructure

This component is the technical backbone that operationalizes the AI model. It encompasses the deployment mechanisms that integrate the finely tuned AI with the necessary servers and Application Programming Interfaces, ensuring that the system is robust, responsive, and scalable. It handles the complexities of managing AI workloads and enables real-time processing of user queries for immediate legal drafting assistance.

Each of these components is meticulously designed to contribute to a cohesive, intelligent system that prioritizes user experience, security, and the delivery of precise, legally sound documents.

V. REQUIREMENTS

A. Data

The foundation of every AI-powered tool for legal drafting is its data. An extensive collection of current legal papers is essential for Lexi to comprehend and produce legal documents in an efficient manner. This dataset must include a broad range of legal terms and ideas, which Lexi must be trained to translate into plain English and simplify. To fine-tune AI, especially for language models, the data must be specifically formatted to distinguish responses from instructions. The model can effectively identify the task at hand and produce the corresponding output thanks to this structured methodology.

JavaScript Object Notation (JSON) objects containing "inputs" and "responses" keys, which direct the model during its generating process, are an example of this format.

A. Software Requirements

Several software components are needed for the creation and optimization of Lexi. Code execution and model iterations require a Jupyter Notebook environment. Gradient & React JS handles the front-end user interface, offering a smooth user experience. MongoDB provides a dependable and expandable chat data storage solution. Because Firebase Auth is dependable and secure, it is the service of choice for user authentication. To keep Lexi available and operating at peak efficiency, Axios is used for API requests, and the Replicate platform is chosen for hosting and model deployment.

B. Hardware Requirements

A balance between processing capacity and efficiency is needed to run complex AI models like the Large Language Model Meta AI (LLaMA-2-7b-chat) [7]. A Graphic Processing Unit with at least 28GB of Video Random Access Memory (VRAM) is advised for running the Large Language Model Meta AI (LLaMA-2-7b) [7] model in full precision, according to information from the Hugging Face forums and other AI resources. However, the Video Random Access Memory (VRAM) demand is reduced by half to 14GB by employing half-precision (16-bit floating points). Use hardware with greater memory during training and fine-tuning, where memory demands are higher, or use strategies like tensor parallelism or model quantization to effectively manage memory requirements.

VI. IMPLEMENTATION AND RESULTS

This portion of the paper explores the empirical core of our research, providing a comparative comparison of the outputs produced by different models. An overview of the website's user interface, including information on the flow of operations and user experience, is given in this part. Combining the theoretical foundations of Lexi with observable outcomes and user interactions helps to illustrate the useful advantages and innovations that this system offers in the legal drafting process.

B. Model Comparison

The training and validation loss curves for the Llama 2 7B [7] language model, which has been refined using a dataset of legal documents, are shown in Fig. 2. The graphic displays the model's performance during the fine-tuning process across a span of four epochs. Both graphs show a steady downward trend, suggesting that as the model learns from the training data, it becomes more capable of producing accurate legal documents.

The validation loss curve stays marginally higher throughout every epoch than the training loss curve. This behavior makes sense because the validation set offers an objective assessment of the model's generalization ability on untested data. The loss curves' convergence indicates that the refined model is successfully adjusting to the legal field while

striking a decent balance between its performance on the training set and its ability to generalize to new situations.

Table 1 compares the performance metrics of three different models the domain-specific Fine-tuned model (Lexi), the base model (Llama 2 7b chat [7]), and Generative Pre-trained Transformer - 3.5 in the comparative study of our models. Specifically designed for legal document drafting, the Fine-tuned model is distinguished by its domain-specific skills that neither the base model (Llama 2 7b chat [7]) nor Generative Pre-trained Transformer - 3.5 has. With an average token count of about 512, it produces outputs that are typified by straightforward legal language, making it more comprehensible and maybe accounting for its greatest ease of comprehension score of 9.5. There may be a trade-off between language complexity and user comprehensibility because the base model and GPT-3.5, in comparison, generate texts with advanced and expert levels of legal jargon, respectively. This leads to longer documents with higher token counts and lower ease of understanding scores.

C. Signup & Login

The website's "Signup & Login" part serves as a vital entry point for user interaction with the Lexi platform. The purpose of this carefully thought-out part is to give users a safe and easy way to enter. The simple signup procedure allows new users to establish a free account by providing basic information like their complete name, email address, and password as shown in Fig 3. By using the login link, current users may quickly access their accounts.

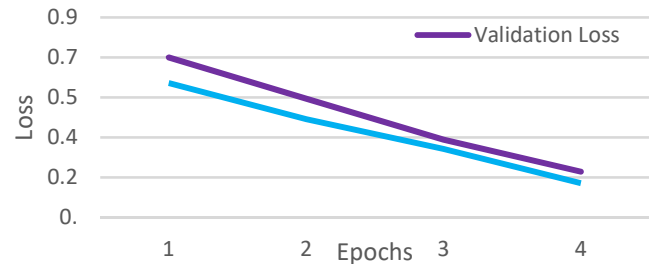


Fig 2. Loss and Dilution versus the number of training epochs

TABLE I. Comparative Analysis of Various Models.

Model/ Metric	Fine- Tuned	Base model	Generative Pre-trained Transformer 3.5
Domain specific	Yes	No	No
Legal jargon level	Basic	Advanced	Expert
Token count	≈ 512	> 800	>1000
Ease of understanding	9.5	7	6

D. Chat Interface

The website's "Chat Interface" component, as shown in Fig. 4, serves as the interactive hub of the Lexi platform, allowing users to communicate with the AI directly to complete legal writing tasks. Users can enter their requests using a natural language conversational interface. Users can request, for example, that Lexi construct a typical rental agreement or a notice of rent increase per Indian rental rules. After processing these requests, the AI creates personalized legal papers in real-time, taking into account the user-specified input parameters. With its simple and intuitive design, this interface makes legal services more accessible to all people by reducing the complexities of legal drafting to a straightforward interaction with an AI assistant.

VII. CONCLUSION

The result of our research, Lexi, a generative artificial intelligence system, redefines the interface between law and technology and represents a revolutionary advance in legal documentation. Our findings show that Lexi can produce legal papers that are accurate and straightforward for users, all while maintaining an unmatched level of efficiency. By converting technical jargon into understandable language, Lexi democratizes legal understanding and drastically cuts down on

the time and expertise needed to generate legal documents. This work promises a more efficient and accessible legal environment for professionals and the general public, and it sets a new standard for artificial intelligence in legal services. It serves as a beacon for future improvements in the field.

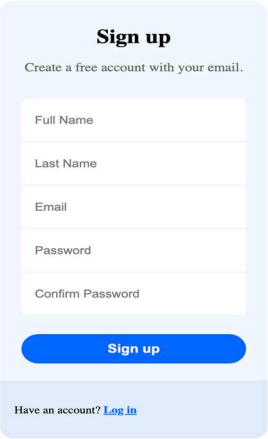


Fig 3. Sign up screen of the web page.

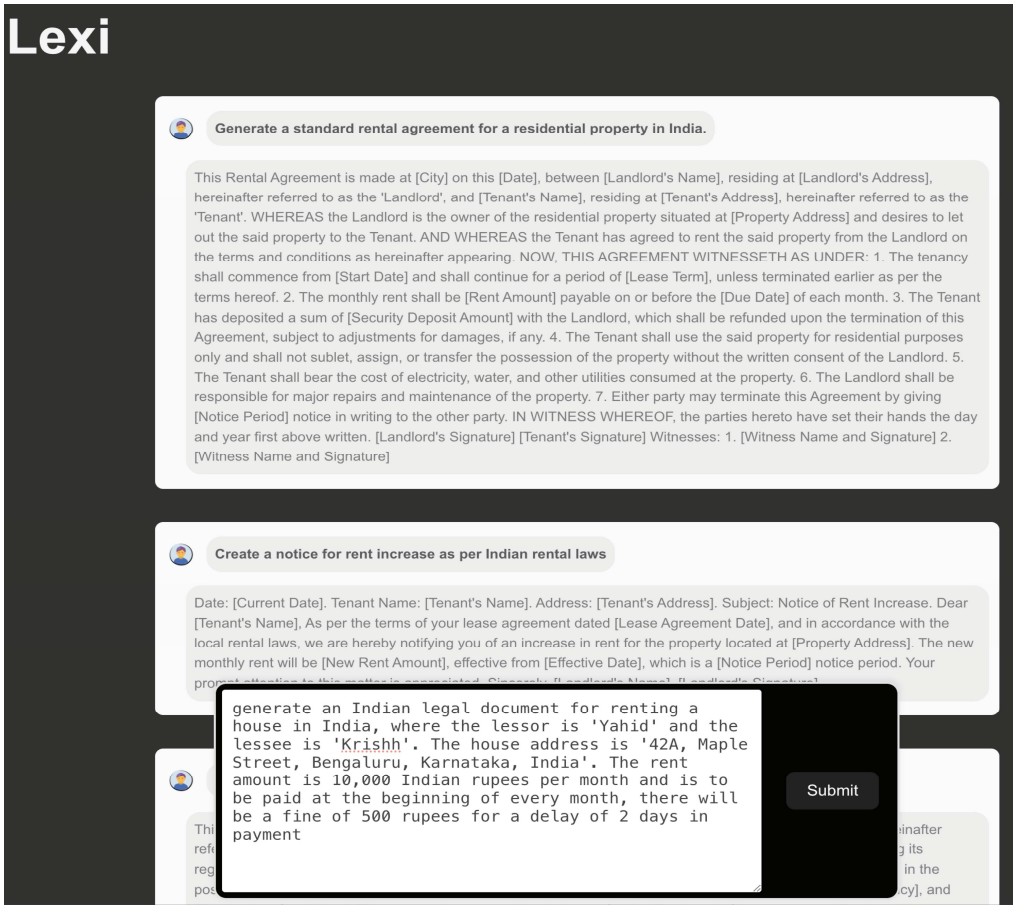


Fig 4. The Chat Interface of the web page

VIII. FUTURE WORK

The plan calls for expanding AI's legal document knowledge beyond rental rules to serve a broader range of industries and legal requirements. To improve the system's usability and accessibility, future research will provide features that will let users export created documents straight to websites like Google Docs. Preserving the integrity of legal information—that is, making sure that the accuracy and significance of the content are retained even when language becomes simpler—will be a crucial component of further progress. With continued development, Lexi can be an even more useful tool that helps close the gap between legal knowledge and general use.

REFERENCES

- [1] Chalkidis, Ilias, Manos Fergadiotis, Prodromos Malakasiotis, Nikolaos Aletras, and Ion Androutsopoulos. "LEGAL-BERT: The muppets straight out of law school." *arXiv preprint arXiv:2010.02559* (2020).
- [2] Katz, Daniel Martin, Dirk Hartung, Lauritz Gerlach, Abhik Jana, and Michael J. Bommarito II. "Natural Language Processing in the Legal Domain." *arXiv preprint arXiv:2302.12039* (2023).
- [3] Ioannidis, Jules, Joshua Harper, Ming Sheng Quah, and Dan Hunter. "Gracenote. ai: Legal Generative AI for Regulatory Compliance." (2023).
- [4] Angeyo, Jennifer Somali, and Peter Jehopio. "An Artificial Neural Network model for regulating legislative drafting practices in Uganda." In *International Conference on Computing and ICT Research*, pp. 51-67. 2010.
- [5] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- [6] Yu, Yu, Chao-Han Huck Yang, Jari Kolehmainen, Prashanth G. Shivakumar, Yile Gu, Sungho Ryu, Roger Ren, Qi Luo et al. "Low-rank adaptation of large language model rescoring for parameter-efficient speech recognition." In *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 1-8. IEEE, 2023.
- [7] Touvron, Hugo, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov et al. "Llama 2: Open foundation and fine-tuned chat models." *arXiv preprint arXiv:2307.09288* (2023).
- [8] Zhong, Haoxi, Chaojun Xiao, Cunchao Tu, Tianyang Zhang, Zhiyuan Liu, and Maosong Sun. "How does NLP benefit legal system: A summary of legal artificial intelligence." *arXiv preprint arXiv:2004.12158* (2020).
- [9] Song, Dezhaoh, Sally Gao, Baosheng He, and Frank Schilder. "On the effectiveness of pre-trained language models for legal natural language processing: An empirical study." *IEEE Access* 10 (2022): 75835-75858.
- [10] Ding, Ning, Yujia Qin, Guang Yang, Fuchao Wei, Zonghan Yang, Yusheng Su, Shengding Hu et al. "Parameter-efficient fine-tuning of large-scale pre-trained language models." *Nature Machine Intelligence* 5, no. 3 (2023): 220-235.
- [11] Fui-Hoon Nah, Fiona, Ruilin Zheng, Jingyuan Cai, Keng Siau, and Langtao Chen. "Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration." *Journal of Information Technology Case and Application Research* 25, no. 3 (2023): 277-304.
- [12] Chowdhery, Aakanksha, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham et al. "Palm: Scaling language modeling with pathways." *Journal of Machine Learning Research* 24, no. 240 (2023): 1-113.
- [13] Chang, Yupeng, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen et al. "A survey on evaluation of large language models." *ACM Transactions on Intelligent Systems and Technology* 15, no. 3 (2024): 1-45.
- [14] Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.
- [15] Zhang, Renrui, Jiaming Han, Chris Liu, Peng Gao, Aojun Zhou, Xiangfei Hu, Shilin Yan, Pan Lu, Hongsheng Li, and Yu Qiao. "Llama-adapter: Efficient fine-tuning of language models with zero-init attention." *arXiv preprint arXiv:2303.16199* (2023).