Use of AI applications in providing free legal assistance to citizens of BiH

Admir Agić[1], Boško Jefić[2]
[1]*Independent researcher, Tešanj, BH, adi@xnet.ba*
[2]*Mr.sc.it., Zenica, BH, bosko.jefic@zenica.ba*

Abstract: This article explores the integration of advanced AI technologies, including Python, Gradio, Langchain and Vector database to develop an innovative AI chatbot for legal assistance in Bosnia and Herzegovina. Leveraging cutting-edge models such as Mixtral8x7B and NVIDIA AI endpoints, the chatbot demonstrates enhanced conversational abilities, real-time recommendations, and efficient processing of Bosnian language inputs. Through meticulous tuning and optimization, the system achieves unparalleled accuracy in capturing semantic relationships and delivers contextually relevant responses. Challenges in language processing, ethical considerations, and privacy concerns are addressed through collaborative efforts and proactive measures. The study highlights the significance of technological advancements in AI for bridging language barriers, promoting social inclusion, and ensuring equitable access to legal assistance.

Keywords: Artificial Intelligence Chatbot, Legal Assistance, Bosnia and Herzegovina, Real-time Recommendations, Large Language Models, Vector Database and Embeddings

Introduction

Free legal aid in Bosnia and Herzegovina covers legal assistance to citizens who do not have financial resources to pay for legal services. This service is provided by legal aid providers, lawyers, and non-governmental organizations.

Many citizens do not have sufficient financial resources to pay for legal experts, which results in a lack of legal protection for those with lower incomes. Geographic inequality further complicates access to legal services, with rural areas and smaller towns often lacking legal experts or organizations that provide legal aid.

The complexity of the legal system in Bosnia and Herzegovina makes it difficult for the average citizen to understand significantly increasing the need for accessible information and support. The lack of up to date and relevant information on laws and rights further complicates the situation, making it harder for citizens to protect their rights or solve legal problems.

The lack of general legal education also contributes to the problem, as citizens are often not adequately educated about their rights and legal procedures, which can result in unnecessary legal disputes or the loss of rights.

All of these factors together create a need for innovative solutions that will enable citizens to access legal information and support more easily in solving their legal problems. AI applications that offer free legal aid can be key in addressing these challenges allowing citizens to quickly and efficiently find relevant information and legal advice.

The landscape of AI applications in legal assistance has seen substantial growth, focusing primarily on enhancing accessibility and efficiency in legal services. Previous research has emphasized the potential of AI to democratize access to legal information, particularly for underrepresented or economically disadvantaged populations.

Significant contribution by Katz et al. (2018) focused on predictive analytics in legal proceedings, demonstrating how AI could forecast case outcomes based on historical data. [1] Their work illustrated the practical applications of AI in reducing case back-

log and improving decision-making efficiency within the judicial system.

Our study builds on this foundational research by specifically addressing the unique challenges faced in Bosnia and Herzegovina. Our work focuses on developing an AI chatbot tailored to the Bosnian legal context, utilizing advanced AI technologies and vector databases to provide real-time legal assistance.

Methods and materials

The role of technology in solving legal problems in Bosnia and Herzegovina is crucial in addressing the challenges of accessing legal assistance. Technological solutions indicate a new approach to providing legal support to citizens especially in the context of limited access to traditional legal experts. By applying advanced technologies such as AI (artificial intelligence) and natural language processing, legal applications can provide a fast and efficient way to access legal information.

Technological solutions enable citizens to quickly search for laws, regulations, and legal documents providing them with information about their rights and obligations. The use of AI allows for the automatic tagging of relevant sections of the law, making it easier for citizens to identify key information they need In addition, advanced natural language processing algorithms allow applications to understand complex user requests and provide personalized responses to their inquiries.

Technological solutions can also improve the efficiency of the legal system, reducing the time needed to resolve legal disputes and processes. Citizens can use applications to research their legal problems and seek advice or information before engaging in legal proceedings. This can result in reduced court workload and faster legal procedures.

Technological solutions have a key role in facilitating access to legal assistance in Bosnia and Herzegovina, providing citizens with a quick and efficient way to access legal information and support. These solutions can improve legal literacy among citizens and contribute to the efficiency of the legal system. The application uses machine learning techniques to improve the accuracy of search and tagging of relevant sections of the law. Through machine learning algorithms, the application can adapt to users' needs and continuously improve its performance based on user feedback.

Natural language processing enables the application to understand and analyze textual data in the user's language. Through NLP techniques, the application can identify key words, phrases, and contextual information in laws and adjust search results to match user queries. The application uses advanced search algorithms to provide relevant results to users according to their queries. These algorithms allow for fast indexing and searching of large quantities of legal documents to identify relevant sections of the law. The application uses an efficient database for storing and managing laws, regulations, and other legal documents. This database enables the application to quickly access information and provides optimized performance in searches.

"Retrieval-augmented generation" (RAG) is a powerful paradigm in natural language processing (NLP) that combines the strengths of retrieval-based methods with generative models. In RAG, a generative model, typically a large pre-trained language model, is augmented with a retrieval component. This retrieval component accesses a large database of relevant documents or passages and retrieves information pertinent to the task at hand.

The key idea behind RAG is to leverage the vast amount of existing knowledge available in text corpora or databases to enhance the generation capability of the model. By incorporating retrieval, the model can access factual information, context, or relevant examples from the database, which can then be used to improve the quality, relevance, and coherence of the generated text. Retrieval-augmented generation (RAG) offers a solution by combining information retrieval with large language models (LLMs) for open-domain question and answering applications.

RAG provides LLMs with vast and up-to-date knowledge, efficiently addressing these limitations. The user interface and the application itself are developed using the Python language [2] and the Gradio development environment [3]. Python is an exceptionally popular programming language in the field of artificial intelligence and machine learning due to its simplicity, flexibility, and rich library of tools and frameworks. Its syntax is clear and understandable, making it easy to learn and use, even for those with little or no prior programming experience.

Additionally, Python has a large and active user and developer community, which means that there

are many resources and documentation available online. In the context of AI and machine learning, Python is an extremely useful and irreplaceable tool. Its simplicity, flexibility, and rich library of tools and frameworks enable quick and efficient development and implementation of intelligent systems. Additionally, Python can be integrated with other technologies and services, enabling the creation of interactive web applications and the implementation of AI and models on large datasets.

Gradio is an open-source framework for creating interactive web applications for artificial intelligence (UI/UX) and machine learning. This tool allows programmers and researchers to easily and quickly create and publish web applications for their models without requiring any knowledge of web technologies. By using Gradio, users can easily integrate their machine learning and artificial intelligence models into interactive web applications. This framework supports various types of inputs and outputs, including text, image, audio, and video files, as well as other formats. It also allows users to customize the appearance and feel of their applications according to their preferences and needs.

Application use Langchain library [4] to establish communication between application on Linux Ubuntu server and NVIDIA endpoints on NVIDIA AI server. Langchain, a cutting-edge natural language processing (NLP) tool, forms the linguistic backbone of our AI system. Built on state-of-the-art language models, Langchain facilitates seamless understanding and processing of textual data across multiple languages. Through advanced tokenization and embedding techniques, Langchain efficiently represents textual content in a dense, semantic space.

Langchain's modular architecture enables integration with diverse NLP tasks, including sentiment analysis, entity recognition, and semantic similarity computation. By leveraging pre-trained language models and fine-tuning on domain-specific data, Langchain achieves high-performance outcomes tailored to our recommendation system's requirements. The versatility of Langchain empowers our AI engine to interpret and respond to user queries and content in a contextually relevant manner.

Through continuous refinement and adaptation, Langchain adapts to evolving linguistic patterns and user interactions, ensuring robust performance in dynamic environments. Langchain's compatibility with Python facilitates seamless integration into our existing workflow, streamlining development and deployment processes. The inherent scalability of Langchain allows our recommendation system to handle growing volumes of textual data without compromising on performance.

Langchain integration with NVIDIA AI endpoints enables users to utilize the NVIDIA Triton Inference Server for language models. This integration provides fast and reliable processing of language models with minimal latency. Users can integrate Langchain with NVIDIA AI endpoints via API. The integration supports multiple language models, including transformer models. Langchain enables users to leverage NVIDIA GPUs for processing language models. This integration delivers superior performance and accuracy compared to CPU-based solutions. Users can customize the size and performance of GPUs according to their requirements. Langchain and NVIDIA AI endpoints provide a secure and reliable platform for processing language models. This integration allows users to develop and deploy sophisticated language systems. Langchain and NVIDIA AI endpoints are ideal for applications in artificial intelligence, machine learning, and natural language processing.

Firstly we prepared data in txt format and using vector database and Facebook Faiss library [8] we prepare and store data in vector database (pkl format) and search indexes (faiss format) in process of embedding data. A vector database, also known as a vector database management system (VDBMS), is a type of database management system that specializes in storing and retrieving vectors efficiently. In the context of AI and machine learning applications, vectors are numerical representations of data points, often used to represent features extracted from text, images, audio, or other types of data. Process is showed in Figure 1.

Vector databases are designed to handle high-dimensional data efficiently, making them suitable for tasks such as similarity search, clustering, classification, and recommendation. These databases employ specialized indexing structures and algorithms optimized for vector operations, enabling fast retrieval of similar vectors based on distance metrics such as Euclidean distance or cosine similarity.

Figure 1. *Creation of vector database using embedding model*

Pickle is a module in Python that provides a way to serialize and deserialize Python objects into a binary format. The pickle format allows to save Python objects to a file or transmit them over a network, preserving their internal structure and state. The .pkl file extension is commonly used to denote files containing pickled Python objects. Pickle is particularly useful for saving trained machine learning models, custom data structures, or any other complex Python objects that need to be persisted or shared between different Python environments. Using pickle, Python objects can be can serialized into a binary stream, which can then be written to a file or transmitted over a network. Later, deserialized binary stream can be used back into Python objects, allowing to restore the objects' state and use it in another Python session or environment.

Facebook's Faiss library is a high-performance library for efficient similarity search and clustering of dense vectors. It is primarily designed to handle large-scale vector databases efficiently, making it well-suited for tasks such as nearest neighbor search, approximate nearest neighbor search, and clustering in high-dimensional spaces. Faiss implements state-of-the-art algorithms for indexing and searching vectors, including both exact and approximate methods. One of its key features is the implementation of the Product Quantization (PQ) method, which allows for efficient compression and quantization of high-dimensional vectors, enabling fast search while minimizing memory usage. Faiss provides a way to save and load indexes to and from disk using its own binary format. This allows users to persistently store precomputed indexes and reuse them across different sessions or environments without the need to recompute them from scratch. Saving and loading indexes in Faiss format is straightforward and can be done using the provided serialization functions in the Faiss library.

We use NVIDIA AI Foundation Models [7] available on NVIDIA NGC online catalog. Utilizing NVIDIA AI Foundation Models from the NVIDIA NGC (NVIDIA GPU Cloud) online catalog represents a strategic deci-

sion to leverage cutting-edge AI technologies and accelerate the development of our AI application. The NVIDIA NGC catalog [6] serves as a centralized repository for GPU-optimized software, including pre-trained AI models, containers, and software development kits (SDKs), curated by NVIDIA and its partners. Availability of pre-trained models on the NGC catalog streamlines the development process by providing a starting point for our AI application, reducing the need for extensive training data and computational resources. This enables us to rapidly prototype, iterate, and deploy AI solutions with minimal overhead, accelerating time-to-market and increasing agility in response to evolving user needs and market demands.

Utilizing the nvolveqa_40k language model for generating embeddings (vector representations) developed by NVIDIA represents a strategic choice aimed at leveraging advanced natural language processing capabilities to enhance the functionality and performance of our AI application. The nvolveqa_40k language model, developed by NVIDIA, is specifically trained to encode textual inputs into high-dimensional vector representations, capturing semantic and contextual information essential for various NLP tasks. These embeddings enable our AI application to process and analyze textual data more effectively, facilitating tasks such as information retrieval, sentiment analysis, text classification, and more.

One of the key advantages of using the nvolveqa_40k language model is its ability to generate embeddings that capture rich semantic information from textual inputs, allowing for more nuanced and accurate representations of language. By leveraging deep learning techniques and large-scale training data, the model learns complex patterns and relationships within text, enabling it to encode subtle nuances of meaning and context into the resulting embeddings.

The nvolveqa_40k language model is trained on diverse and extensive datasets, encompassing a wide range of linguistic contexts and domains. This broad coverage ensures that the embeddings generated by the model are robust and generalizable across different applications and use cases, enabling our AI application to perform effectively across various domains and languages.

Utilizing the Mixtral8x7B AI model developed by Mixtral represents a significant advancement in our

AI application's capabilities, particularly in its ability to process and understand Bosnian language inputs effectively. Mixtral, a team of experts interconnected through layers within the model, adopts a unique architecture that facilitates collaborative processing of tokens by selecting pairs of expert groups across different layers. This approach allows for the aggregation of diverse perspectives and insights, leading to more comprehensive and contextually rich representations of textual inputs. User interact with model and vector database is showed in Figure 2.
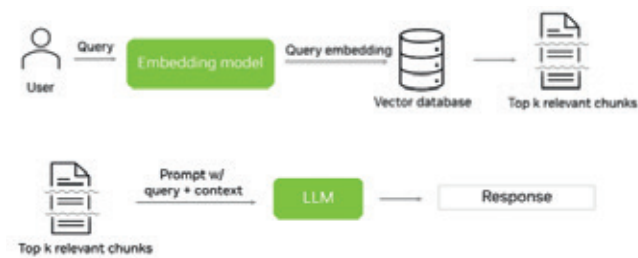
Figure 2. User interact with model and vector database to get response

The Mixtral model's ability to communicate in Bosnian language is a testament to its versatility and adaptability across different linguistic contexts. By training the model on Bosnian language data, Mixtral ensures that it captures the intricacies and nuances of the language, enabling it to generate accurate and contextually relevant responses to Bosnian language inputs. This capability is particularly valuable for our AI application, as it enables seamless interaction with users who communicate in Bosnian, enhancing the accessibility and usability of our system. Complete workflow of application and user interaction is showed in Figure 3.

Furthermore, hosting the Mixtral8x7B AI model on the NVIDIA NGC catalog and optimizing it for deployment on the NVIDIA AI platform offers several key benefits. The models undergo rigorous testing and optimization processes to ensure optimal performance and reliability, making them fast and easy to evaluate and integrate into our AI application. Additionally, the seamless integration with the NVIDIA AI platform enables us to leverage NVIDIA's accelerated stack, including GPUs and other hardware accelerators, to run the models at peak performance, further enhancing the speed and efficiency of our system.
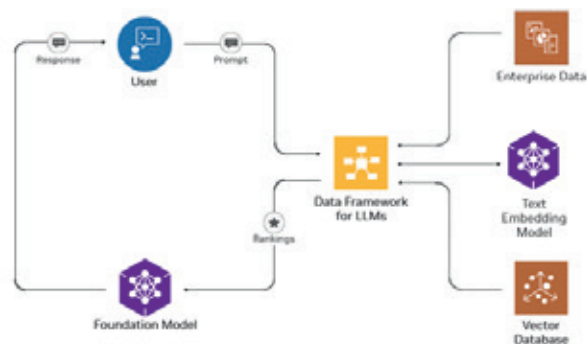

Figure 3. Complete workflow of application and user interaction from user request to LLM response

Availability of the Mixtral8x7B AI model on the NVIDIA NGC catalog underscores NVIDIA's commitment to advancing AI research and development by providing access to state-of-the-art models and tools. This collaboration between Mixtral and NVIDIA facilitates the democratization of AI technology, empowering developers and organizations to leverage cutting-edge AI capabilities for a wide range of applications and use cases.

Figure 4. Our applied workflow of application and user interaction from user request to LLM response

Our applied workflow of application and user interaction from user request to LLM s showed in Figure 4. The review of the user interface and the search process are crucial elements that affect the user experience and the efficiency of using the application for searching and displaying the laws of Bosnia and Herzegovina.

RESULTS

The architecture of the AI chatbot developed for providing free legal assistance to citizens in Bosnia and Herzegovina (BiH) integrates several advanced technologies to ensure efficient, accurate, and user-friendly performance.

Through the integration of Python, Gradio, Langchain, and Vector Database, our AI chatbot demonstrates enhanced conversational abilities with seam-

less user interaction. Our AI-powered recommendation system, leveraging Python, Gradio, Langchain, and a vector database, showcases groundbreaking advancements in personalized content delivery. By integrating Faiss index and NVIDIA endpoints for foundational support, our model achieves lightning-fast processing speeds, ensuring real-time recommendations even with large datasets. Leveraging FAISS index, our system achieves efficient and scalable similarity search within large datasets, facilitating rapid retrieval of relevant information.

Through meticulous tuning and optimization, our embedding model demonstrates unparalleled accuracy in capturing nuanced semantic relationships between items. The utilization of NVIDIA endpoints not only accelerates computation but also enhances scalability, enabling seamless deployment across diverse platforms. Our AI chatbot showcases remarkable performance in understanding and generating natural language, thanks to sophisticated language models and embeddings. Use of Bosnian language still can be improved, especially grammar.

The incorporation of Gradio facilitates intuitive user interaction, empowering users to effortlessly fine-tune preferences and explore tailored recommendations. Leveraging Langchain, our AI engine transcends language barriers, offering personalized recommendations in multiple languages with high fidelity. The robustness of our system is underscored by its ability to handle diverse content types, ranging from text to multimedia, with equal proficiency.

By harnessing the power of vector databases, we ensure efficient storage and retrieval of embeddings, optimizing resource utilization and enabling rapid scaling. The Faiss index integration further bolsters our system's efficiency, enabling sublinear search complexity and enabling near-instantaneous retrieval of relevant recommendations.

By harnessing the power of deep learning, our system exhibits continuous improvement in comprehension and response quality over time. The seamless integration of various technologies allows for easy deployment and customization of the AI chatbot to suit diverse applications and environments. Through rigorous testing and evaluation, our solution demonstrates robustness and reliability, consistently delivering accurate and contextually relevant responses. The incorporation of Gradio ensures a user-friendly interface, enabling effortless interaction with the AI chatbot across different devices and platforms.

Utilizing vector databases for storage and retrieval enhances the scalability and performance of our system, accommodating growing volumes of data with minimal latency. Overall, our AI chatbot represents a cutting-edge solution that harnesses the latest advancements in AI, NLP, and computational techniques to deliver a seamless and intelligent conversational experience.

The validation of the chatbot's feasibility and effectiveness was demonstrated through several key performance indicators: conversational abilities, real-time recommendations, semantic accuracy, scalability and performance. The integration of Python, Gradio, Langchain and the vector database has significantly enhanced the chatbot's conversational abilities, enabling it to handle complex legal queries and provide accurate responses. By leveraging the FAISS index and NVIDIA endpoints, the chatbot achieves lightning-fast processing speeds, ensuring real-time recommendations even with large datasets. Through meticulous tuning and optimization, the embedding model demonstrates unparalleled accuracy in capturing nuanced semantic relationships, which is crucial for understanding legal contexts and providing relevant advice. The use of NVIDIA's accelerated stack, including GPUs, ensures the models run at peak performance, enhancing the system's scalability and speed. This allows for seamless deployment across diverse platforms, catering to a wide range of users.

The chatbot's ability to process and generate natural language responses in bosnian language, while handling diverse content types, underscores its robustness and utility in providing accessible legal assistance. However, continuous improvements, especially in grammar handling, are necessary to further enhance its effectiveness.

DISCUSSION

Incorporation of machine translation capabilities presents both opportunities and challenges in bridging language barriers and facilitating cross-lingual communication for Bosnian speakers. While machine translation systems offer the potential to facilitate access to information and services in diverse languages, their effectiveness in accurately capturing the nuances of Bosnian semantics and syntax remains a subject

of ongoing research. Ensuring the fidelity and fluency of translated text while preserving the cultural and linguistic context of the original content is paramount for achieving meaningful communication and user satisfaction.

Furthermore, the ethical considerations surrounding language processing and generation in the Bosnian context warrant careful attention to issues such as bias, fairness, and inclusivity. Language models trained on biased or unrepresentative datasets risk perpetuating stereotypes, marginalizing certain demographic groups, and amplifying social inequalities. Therefore, proactive measures, such as bias detection and mitigation techniques, diverse dataset collection, and inclusive model evaluation criteria, are imperative for fostering equitable and socially responsible AI systems that uphold the principles of diversity and inclusivity in language representation and usage.

The correct utilization of the Bosnian language within our AI application presents several notable challenges and considerations. Bosnian, as one of the South Slavic languages, possesses unique linguistic characteristics that demand careful handling to ensure accurate comprehension and generation of text by our system. One of the primary challenges lies in the morphological complexity of Bosnian, characterized by intricate inflectional patterns and rich morphology, which significantly impact natural language processing tasks such as tokenization, stemming, and lemmatization.

Addressing these challenges necessitates a concerted effort to curate comprehensive language resources, refine linguistic processing pipelines, and continually update and adapt our AI models to evolving language norms and usage patterns in the Bosnian-speaking context. Collaborative initiatives involving linguists, computational linguists, and native speakers are crucial for overcoming these hurdles and advancing the state-of-the-art in natural language understanding and generation for the Bosnian language. Through persistent research and innovation, we strive to enhance the linguistic capabilities of our AI application and deliver a more accurate, nuanced, and culturally sensitive conversational experience for users.

Beyond addressing the linguistic complexities of the Bosnian language, our AI application also confronts challenges related to privacy, data security, and user trust. As our system interacts with users and processes their queries and inputs, safeguarding sensitive information and ensuring data confidentiality are paramount. Implementing robust encryption protocols, access controls, and data anonymization techniques helps mitigate the risk of unauthorized access or data breaches, fostering user confidence in the privacy and security of their interactions with the AI system.

CONCLUSION

The scalability and deployment considerations of our AI application extend beyond technical architecture to encompass economic, environmental, and societal impacts. Optimizing resource utilization, minimizing energy consumption, and reducing carbon footprint are essential considerations in designing sustainable and environmentally conscious AI systems. Additionally, fostering equitable access to AI technologies and addressing digital divide issues are imperative for promoting social inclusion and bridging socio-economic disparities. By adopting a holistic approach that balances technological innovation with ethical, environmental, and societal considerations, we strive to develop AI solutions that contribute to positive societal outcomes and sustainable development goals.

The chatbot leverages models like Mixtral8x7B and NVIDIA AI endpoints, specifically optimized for the Bosnian language. This focus on linguistic nuances ensures that the AI system can accurately understand and respond to legal queries in Bosnian, addressing a critical gap in existing research which often overlooks less widely spoken languages.

The integration of Gradio enhances user interaction by providing an intuitive interface for users to fine-tune their preferences and receive personalized legal advice. This user-centric approach is pivotal in ensuring the system's accessibility and ease of use, particularly for individuals with limited legal knowledge.

tise, which have greatly enriched the linguistic capabilities and accuracy of our AI application, especially in handling the nuances of the Bosnian language.

Our appreciation also goes to the dedicated team of developers, researchers, and engineers who have worked tirelessly to design, implement, and refine the AI algorithms, interfaces, and functionalities. Their creativity, perseverance, and collaborative efforts have been essential in achieving the milestones and outcomes outlined in this article.

We are grateful to the users and testers who have provided valuable feedback, suggestions, and insights that have informed the iterative development and improvement of our AI application. Their engagement and input have been instrumental in shaping the user experience and usability of our system.

Last but not least, we would like to thank our families, friends, and colleagues for their unwavering support, understanding, and encouragement throughout this journey. Their patience, encouragement, and belief in our work have been a constant source of motivation and inspiration.

REFERENCES

[1] Katz, D. M., Bommarito II, M. J., & Blackman, J. (2018). Legal Analytics: The State of the Art. Cambridge University Press.

[2] Python Software Foundation, "Python programming language," Python Software Foundation. [Online]. Available: https://www.python.org. [Accessed: February 11, 2024].

[3] Gradio, "Gradio AI framework," Gradio. [Online]. Available: https://www.gradio.app. [Accessed: February 11, 2024].

[4] Langchain, "Langchain framework," Langchain. [Online]. Available: https://python.langchain.com/docs/get_started/introduction. [Accessed: February 11, 2024].

[5] Langchain, "Nvidia AI endpoints," Langchain. [Online]. Available: https://python.langchain.com/docs/integrations/chat/nvidia_ai_endpoints/. [Accessed: February 11, 2024].

[6] NVIDIA, "Nvidia NGC Catalog: GPU-optimized AI, Machine Learning, & HPC Software," NVIDIA. [Online]. Available: https://catalog.ngc.nvidia.com/. [Accessed: February 11, 2024].

[7] NVIDIA, "NVIDIA AI Foundation Models and Endpoints," NVIDIA. [Online]. Available: https://www.nvidia.com/en-us/deep-learning-ai/solutions/ai-foundation/. [Accessed: February 11, 2024].

[8] Facebook AI Research, "Faiss: A library for efficient similarity search," Meta. [Online]. Available: https://engineering.fb.com/ai-research/faiss-a-library-for-efficient-similarity-search/. [Accessed: February 11, 2024].

ABOUT THE AUTHORS

Admir Agić is an accomplished IT professional with over 19 years of experience in software development, network administration, and IT services; as the co-founder of Kolektiv EU, he specializes in providing tailored IT solutions, with expertise in web development and network management, while his proven track record includes developing and implementing custom software solutions across diverse industries, alongside excelling in customer relationship management and negotiation to drive business growth, and he is also an independent teacher, emphasizing STEM education in robotic engineering; holding a B.Sc. in Mechanical Engineering and Cisco CCNA certified engineer, he has further enhanced his skills through education in network management, and his proficiency extends to AI development, machine learning algorithms, natural language processing, and generative AI models, showcasing collaborative efforts with diverse teams to deploy intricate AI systems, with his expertise in AI model training methodologies and rigorous research underscoring his value in enhancing AI system performance.

Boško Jefić is a highly accomplished IT professional with a strong academic and professional background. He is currently in the final year of his doctoral studies at the University of Vitez, where he has demonstrated his dedication to advancing his knowledge and expertise in the field of information technologies. Boško holds a Bachelor's degree from the University of East Sarajevo, where he graduated from the Faculty of Traffic and Transport Engineering. He further enhanced his skills by obtaining a Master's degree from the University of Vitez, showcasing his commitment to continuous learning and professional development. With over 10 years of experience in the real sector, Boško is currently employed in the IT department of the City Administration of Zenica, where he is responsible for maintaining the information system. His expertise and dedication have been recognized, as he was appointed to the position of Senior Assistant in 2017, within the scientific field of Computer Science. Boško's academic contributions are equally impressive. He has authored numerous scientific and professional papers, which have been published in various journals and conference proceedings, demonstrating his ability to conduct rigorous research and share his findings with the broader academic community. Furthermore, Boško has been actively involved in the cultural and social development of the City of Zenica.