Problem Statement

This problem statement is designed to introduce beginners to the exciting field of Generative Artificial Intelligence (GenAI) through a series of hands-on exercises.

Participants will learn the basics of GenAI, perform simple Large Language Model (LLM) inference on a CPU, and explore the process of fine-tuning an LLM model to create a custom Chatbot.

Unique Idea Brief (Solution)

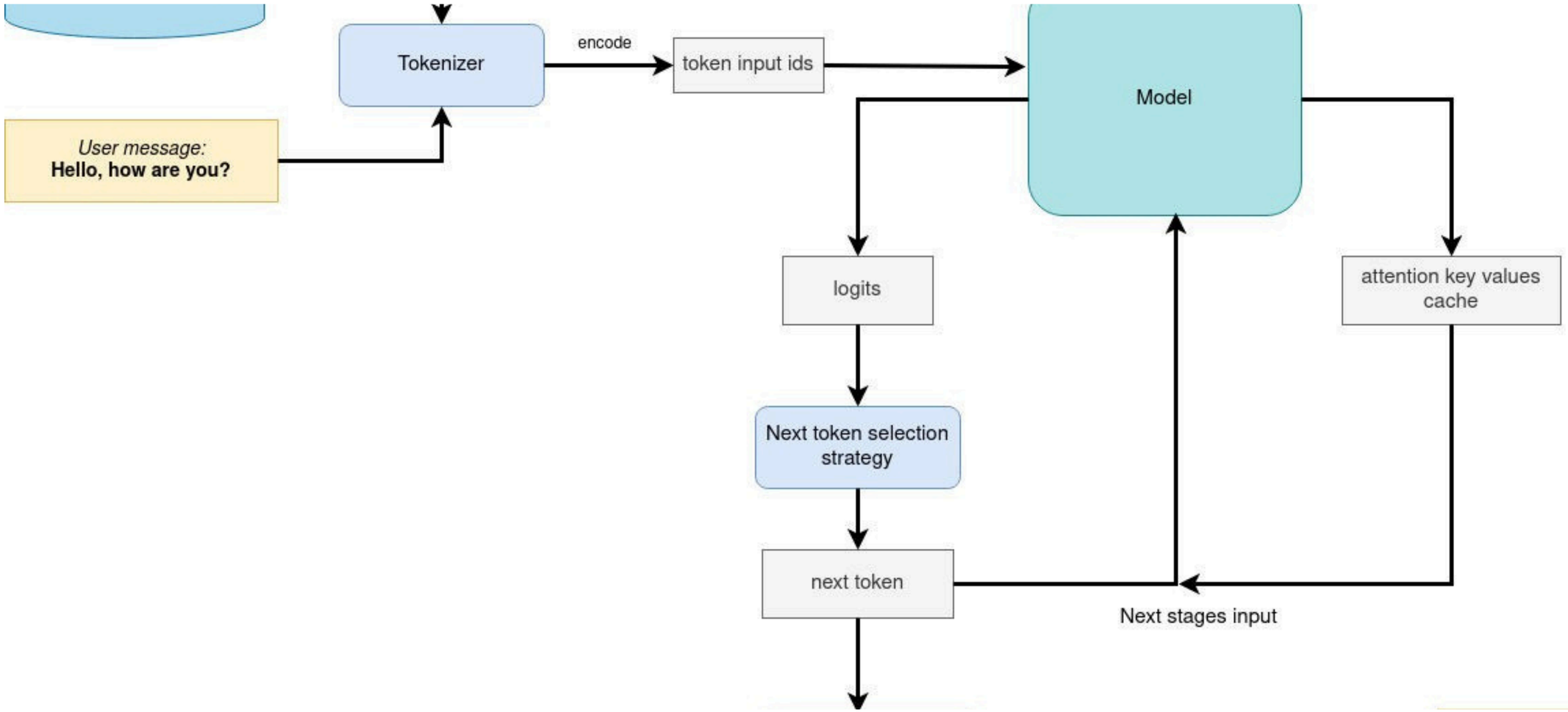
The answer is setting up the environment, GenAI for beginners, LLM inference with OpenVINO™ on CPUs , training LLMs and generate a chatbot.

Our method focuses on Intel hardware where we can train AI models efficiently, optimize the model, then test inference and generate a chatbot and even do some deployments to demonstrate how this can be done in practice for instance chatbot development.

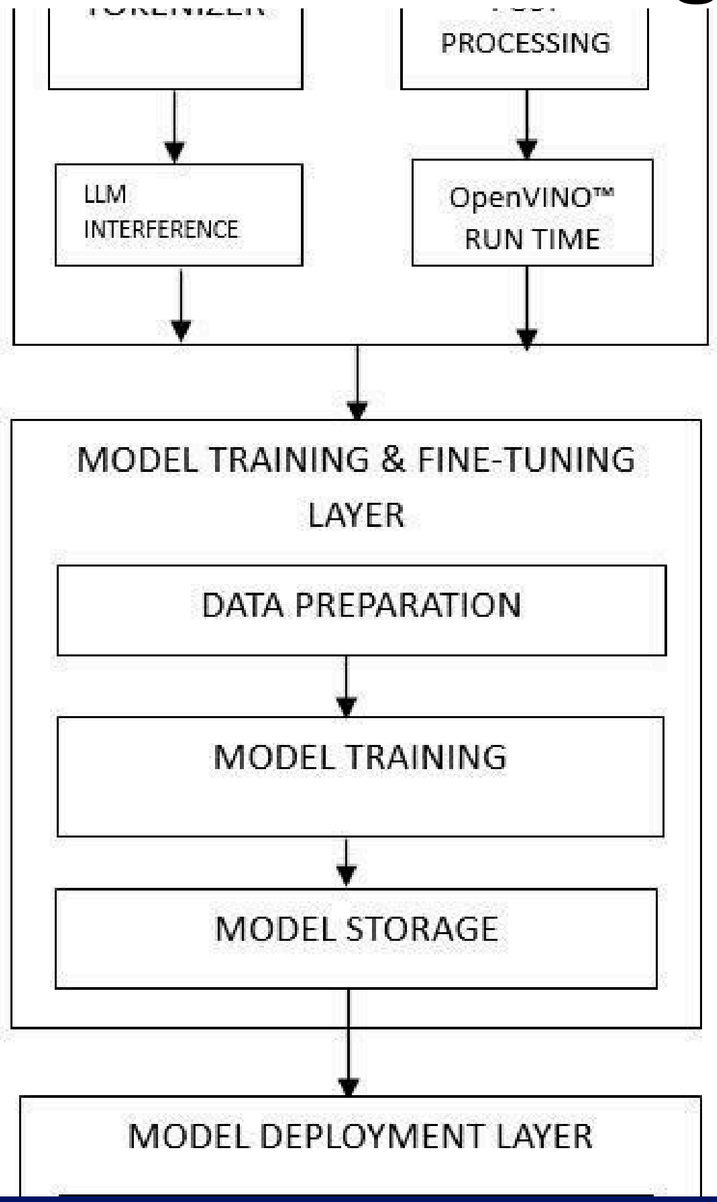
Features Offered

Environment Setup: Easy and cross-platform setup with essential software.
GenAI Training: Foundational knowledge and practical examples for understanding GenAI.
LLM Inference Optimization: Efficient LLM inference on CPUs using OpenVINO™.
Chatbot Creation: Interactive chatbot creation with the optimized model.
Performance Evaluation: Thorough evaluation and feedback mechanisms to improve model performance.

Process flow



Architecture Diagram



Technologies used

Programming Language: Python.

NLP Framework: Hugging Face Transformers, Tokenizers

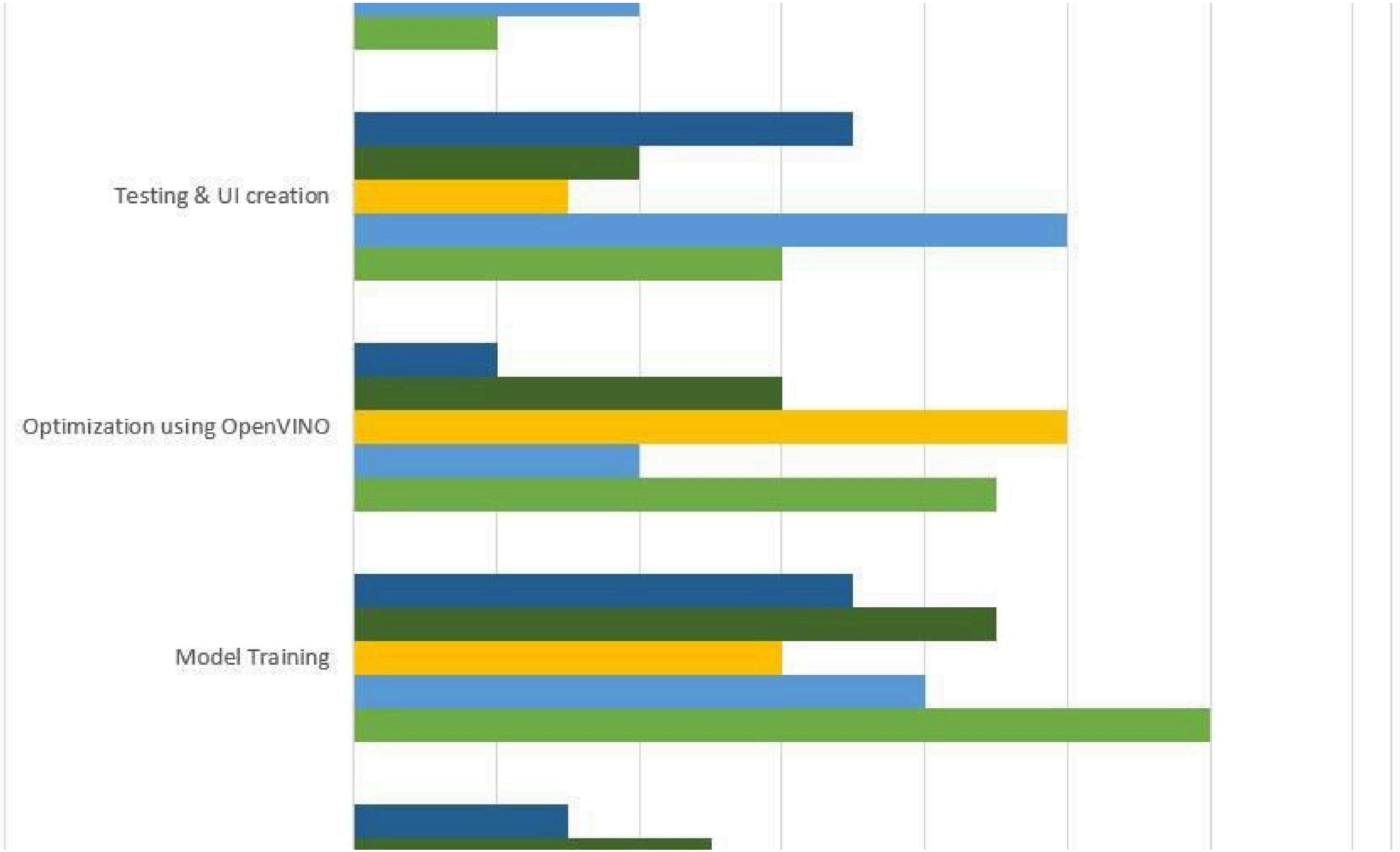
Model Optimization: Intel® OpenVINO™ Toolkit.

Data Manipulation: Numpy and Pandas.

Deep Learning Frameworks : PyTorch.

Creating User Interface: Streamlit

Team members and contribution:



Conclusion

In conclusion, this project not only equipped participants with the necessary skills and knowledge to work with GenAI and language models but also highlighted the practical applications of AI in addressing significant societal challenges.

The integration of Intel® OpenVINO™ tools further enriched the learning experience by demonstrating the power of optimized AI solutions on widely accessible hardware.

Continuous research and development in this field are essential to keep up with the evolving landscape of AI and its applications.