brillio

Redefining Operational Efficiency with Innovative Generative AI Solutions for a Fortune 500 Pharma Enterprise

Revolutionizing user experience with a one-stop tool through real-time query responses and enhanced scalability, security, and workflows.

With a workforce exceeding 34,000 professionals and research and development facilities spanning the globe, this biopharmaceutical enterprise ranks among the Fortune 500 powerhouses. It asserts its position as one of the largest global corporations, dedicated to the relentless pursuit of discovering, developing, and delivering groundbreaking medicines to address the critical needs of patients grappling with severe diseases.

In the pharmaceutical sector, time is of the essence. Recognizing this imperative, is committed to a culture of perpetual advancement across all operational dimensions, steadfastly integrating state-of-the-art technologies and fostering agility to ensure optimal performance.

However, the organization encountered a number of challenges, particularly with excessively long user query durations, significantly impeding operational efficacy.

Revolutionizing User Experience: Leveraging Gen AI for Instant Query Resolutions

In pursuit of enhanced efficiency and streamlined operational processes, the company's regulatory business team expressed a desire for a comprehensive solution: a single tool capable of promptly addressing user queries without the customary three-day delay. They had envisioned an interactive Q&A bot, proficient at leveraging enterprise data to furnish precise responses to specific business inquiries. With this conversational bot at their disposal, authorized users could access information instantaneously.

To achieve this objective, the company embarked on a quest to find a partner to spearhead this implementation. With deep expertise, seasoned talent, and a demonstrated track record of success in the AI space, Brillio emerged as the ideal partner.

Two separate solution approaches were devised to tackle the challenge and develop LLM modules, one tailored for unstructured data and another specifically designed for structured data.

The existing process mandated business teams that needed information regarding a drug, such as registration statuses or obtaining various details, to submit a request, which would require over **72 hours for completion.**

Demystifying Unstructured Data

In the case of unstructured data, a dual strategy was employed, consisting of a Retriever and a Generator. For the Retriever component, OpenSearch was chosen for its open-source nature, scalability, and cost-effectiveness, facilitating the extraction of relevant context for a given user query. This contextual information is subsequently utilized to generate responses. As for the Generator, an AzureOpenAI model—text-davinci-003—was utilized. This model harnesses the user query along with the relevant context retrieved from OpenSearch to craft natural language responses.

Handling Structured Data

For structured data, the approach required an innovative solution. Given the sheer size of the input structured data, constructing a single generator posed challenges due to token limit errors and scalability issues. To overcome these roadblocks, a unique generator-generator approach was proposed and implemented, built upon the AzureOpenAI model with gpt-35-turbo, with the first generator converting a given question into an SQL query, while the second generating a natural language response by integrating the user query with the output of the SQL query.

To ensure a seamless user experience, an intelligent classification system was employed to categorize input user queries as structured, unstructured, follow-up, or a combination of both, triggering the respective pipelines automatically. Memory management was handled using the langchain framework, with a fallback mechanism incorporated to activate alternate pipelines in case of misclassification, ensuring robustness and reliability in addressing user queries.

Securing Scalability: Multi-Cloud Components Enhance Data Management

While natural language responses are generated by Azure OpenAI models hosted within the Azure cloud, access to bots is tightly controlled through the creation of a security group, only allowing authorized users to submit queries.

For scalability, the Teams UI was chosen as the user interface due to its ability to effortlessly handle load issues, while the remaining components of this architecture, including the OpenSearch index and AzureOpenAI models, inherently possess

scalability features, ensuring seamless expansion as demands increase.

These intelligent bots seamlessly navigate between various tabular datasets, addressing both direct inquiries and follow-up questions. When presented with a query lacking context, the model refrains from speculation and instead provides a straightforward response: "No information is available for the given query."

To enhance data security and scalability, multi-cloud components were leveraged.

The Impact of One-Stop Tool Integration: Streamlined Processes and Real-Time Responses

The implementation through the integration of a one-stop tool has brought forth numerous benefits for the client, revolutionizing the user experience and enhancing operational efficiency.

Users now have instant access to information from documents and tabular datasets, resulting in streamlined processes and reduced turnaround times through intelligent enablement and automation. With an initial build time to market of just 12-14 weeks and the capability for quick enhancements based on requirements, the implementation ensures agility and responsiveness to evolving needs.

Additionally, the chat operates effortlessly within the familiar MS Teams environment, requiring no additional tools or specialized training. This accessibility empowers users of all technical proficiencies to leverage the tool effectively. The integration of the Bot into Teams enables users to enjoy real-time responses, eliminating the need for lengthy waits and follow-up requests. Queries are resolved in under 30 seconds, enhancing productivity and completely transforming the user experience.

brillio

ABOUT BRILLIO

At Brillio, our customers are at the heart of everything we do. We were founded on the philosophy that to be great at something, you need to be unreasonably focused. That's why we are relentless about delivering the technology-enabled solutions our customers need to thrive in today's digital economy. Simply put, we help our customers accelerate what matters to their business by leveraging our expertise in agile engineering to bring human-centric products to market at warp speed. Born in the digital age, we embrace the four superpowers of technology, enabling our customers to not only improve their current performance but to rethink their business in entirely new ways. Headquartered in Silicon Valley, Brillio has exceptional employees worldwide and is trusted by hundreds of Fortune 2000 organizations across the globe.

https://www.brillio.com/
Contact Us: info@brillio.com