

Agente conversacional baseado em inteligência artificial generativa aplicado ao setor financeiro

Vanessa da Silva¹, Carlos Amaral Hölbig¹

¹Programa de Pós-Graduação em Computação Aplicada (PPGCA)
Universidade de Passo Fundo (UPF)
Campus 1 – BR 285 – 99.052-900 – Passo Fundo – RS – Brasil

172815@upf.br, holbig@upf.br

Abstract. *This work presents the development of a conversational agent based on generative AI for the financial sector, focused on natural language queries to internal structured databases. The solution aims to reduce information overload, provide contextual and secure answers, and support operational efficiency. Its practical implementation will take place in a financial institution, in the context of a chatbot designed to solve internal questions.*

Resumo. *Este trabalho apresenta o desenvolvimento de um agente conversacional baseado em IA generativa para o setor financeiro, com foco em consultas em linguagem natural a bases internas estruturadas. A solução busca reduzir o excesso de informações, oferecer respostas contextualizadas e seguras e apoiar a eficiência operacional. Sua implementação prática ocorrerá em uma instituição financeira, no contexto de um chatbot para resolução de dúvidas internas.*

1. Introdução

Grandes empresas possuem, em geral, um grande número de funcionários que precisam estar constantemente envolvidos em atividades que garantam uma elevada eficiência operacional. No contexto do setor financeiro, tal exigência se torna crítica, visto que a movimentação do mercado obriga os profissionais a acompanharem e analisarem, de maneira constante, métricas e indicadores de desempenho que impactam diretamente no resultado financeiro das operações. Entretanto, a interpretação quantitativa de todo esse grande conjunto de dados pode se tornar uma tarefa penosa e exaustiva, especialmente quando realizada por profissionais recém-contratados, que ainda não conhecem todos os detalhes dos processos e das ferramentas da organização. Um estudo de [Firmino et al. 2022] analisou estratégias de presença empreendedoras e de inovação nas dez maiores instituições financeiras brasileiras, revelando que todas estão incorporando tecnologia em seus processos e produtos.

A necessidade de acesso ágil e assertivo aos principais indicadores da carteira de clientes sob a responsabilidade das agências incrementa, ainda mais, a complexidade das tarefas que envolvem tanto os gerentes de carteira quanto a gestão da própria agência. Além das atividades de natureza negocial, esses profissionais ainda precisam enfrentar o desafio de identificar e interpretar sinais em um mar de informações cada vez mais sobrecarregado. Isso se torna ainda mais complicado quando se considera a diversidade e a dispersão das fontes de informação, que vão desde bases analíticas detalhadas até dashboards, cujo design e organização nem sempre atendem às necessidades dos profissionais que lidam diretamente com os clientes.

2. Fundamentação teórica

Em face da alta exposição a um grande volume de dados que cresce continuamente, é fundamental contar com soluções tecnológicas que possam filtrar, de maneira seletiva e contextual, apenas as informações mais pertinentes para o momento específico do usuário. Essa demanda surge do fenômeno denominado information overload, em que a excessiva quantidade de dados pode prejudicar a capacidade de tomar decisões de forma eficiente e oportuna.

Nesse contexto, surgem tecnologias sofisticadas que permitem a comunicação em linguagem natural entre o usuário e diversas bases de dados corporativas. Essas soluções, baseadas em técnicas de Processamento de Linguagem Natural (NLP), conseguem entender e interpretar perguntas feitas em linguagem humana, fornecendo respostas precisas e contextualizadas. Essa habilidade permite, por exemplo, que um funcionário do departamento financeiro acesse de forma rápida dados e indicadores específicos da carteira de clientes, sem precisar revisar manualmente uma grande quantidade de registros ou dashboards pouco focados.

No âmbito das FinTechs, as aplicações que utilizam o Processamento de Linguagem Natural (NLP) têm se revelado instrumentos eficientes para possibilitar que os usuários analisem o rendimento de variados ativos financeiros, entendam as tendências do mercado e façam escolhas de investimento embasadas, mesmo sem ter um conhecimento avançado em programação ou análise de codificação especializada [Yildirim et al. 2024]. Essas soluções possibilitam o conceito de “conversação de dados” (*data conversation*), que permite ao usuário se comunicar diretamente com um sistema computacional, fazendo perguntas em linguagem natural para obter, por exemplo, dados específicos sobre a carteira de clientes de uma instituição financeira.

Nesse contexto, os chatbots se destacam, também conhecidos como inteligência artificial conversacional (IA conversacional) ou chatbots de IA. Esses programas de computador são projetados para simular conversas humanas [Susana et al. 2024]. Quando incorporadas a bases de dados empresariais e equipadas com recursos de interpretação semântica, essas ferramentas podem oferecer respostas exatas e personalizadas, aprimorando os processos de consulta e análise no setor financeiro.

3. Agentes conversacionais no contexto financeiro

A proposta deste pesquisa a respeito do problema identificado envolve a criação de um agente conversacional baseado em arquiteturas GPT. Modelos como GPT (*Generative Pre-trained Transformer*) e Gemini, incluindo ChatGPT (desenvolvido pela OpenAI), estão na vanguarda do desenvolvimento de chatbots [Gupta and Gupta 2025].

Levando em conta as demandas do setor financeiro em relação à segurança e conformidade regulatória, a implementação desse agente conversacional ocorrerá em um ambiente supervisionado pela própria instituição. Para isso, a API da OpenAI será empregada, garantindo conformidade com as políticas corporativas de governança de dados da empresa onde esta aplicação está sendo desenvolvida. Essa estratégia tem como objetivo não só diminuir os riscos associados à manipulação e a exposição de dados sensíveis, mas também reduzir a frequência de alucinações do modelo e estabelecer mecanismos de contenção (*guardrails*) que garantam respostas consistentes, seguras e adequadas ao

contexto organizacional. Na Fig. 1 é apresentada a arquitetura básica da solução, onde o usuário envia uma pergunta sobre alguma informação da instituição, a qual é recebida pelo chatbot e transformada em uma requisição direcionada à IA. Antes de gerar a resposta, a IA acessa e combina informações provenientes de bases internas estruturadas, garantindo que a saída seja contextualizada e confiável. Após o processamento, a resposta é retornada ao chatbot, que aplica filtros de segurança e compliance antes de entregá-la ao usuário. Esse fluxo permite oferecer interações mais precisas, seguras e alinhadas ao contexto organizacional, reduzindo alucinações do modelo e ampliando a eficiência na disponibilização de dados e informações.

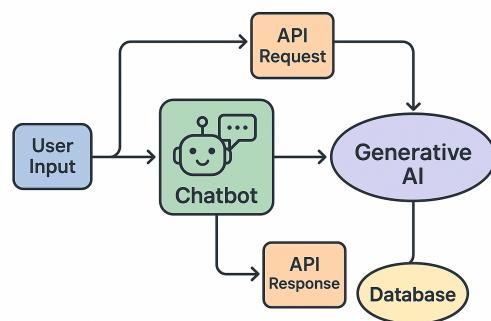


Figura 1. Fluxograma da solução proposta

4. Considerações finais

Esta pesquisa encontra-se em andamento, em fase de criação da base de dados e início das fases de implementação, de validação e de testes. Sua implementação terá caráter prático, desenvolvida em uma instituição financeira, no contexto de um chatbot voltado à resolução de dúvidas internas, a partir do tratamento e da disponibilização das informações que circulam dentro da própria organização.

Referências

- Firmino, T. T., da Silva, A. G. F., da Silva, J. D., and Nóbrega, K. C. (2022). Critérios para avaliação da presença de estratégias empreendedoras e de inovação: um estudo em instituições financeiras brasileiras. *Revista de Gestão e Secretariado*, page 72–93.
- Gupta, E. and Gupta, V. (2025). A comparative analysis of esp32 and esp8266 for ai-powered applications. In *2025 International Conference on Next Generation Communication Information Processing (INCIP)*, pages 992–996.
- Susana, D., Vijaykumar, N., and Menon, S. (2024). Ai chatbots in banking: Exploring key factors influencing customer satisfaction through the extended e-squal model. In *2024 International Conference on Intelligent Innovative Practices in Engineering Management (IIPEM)*, pages 1–6.
- Yildirim, S., Santur, Y., and Aydogan, M. (2024). Nlp in fintech: Developing a lightweight text-to-chart application for financial analysis. In *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, pages 1–6.