

Santander melhora a qualidade e eficiência no serviço a cliente com solução de IA desenvolvida em parceria com a NVIDIA

- *Santander desenvolve modelos de inteligência artificial especializados em tarefas bancárias, com o objetivo de melhorar o desempenho e a eficiência.*
- *Parceria com a NVIDIA é reforçada com um acordo de três anos, que abrange áreas como a prevenção da fraude, os agentes de IA e a tomada de decisões em tempo real.*

Lisboa, 21 de outubro de 2026. **NOTA DE IMPRENSA**

O Banco Santander está a aprofundar a sua colaboração com a NVIDIA para desenvolver soluções de inteligência artificial mais eficientes e adaptadas às necessidades da banca. Através do Santander AI Lab, o Banco está a testar e a aperfeiçoar modelos e agentes de IA destinados a aplicações bancárias, permitindo identificar o modelo mais adequado para cada tarefa e implementá-lo da forma mais eficiente.

O trabalho está a ser desenvolvido no Santander AI Lab, que explora e testa tecnologias emergentes de IA antes da sua adoção mais ampla em todo o Banco. No âmbito desta iniciativa, o Santander tem vindo a evoluir progressivamente da utilização de grandes modelos de uso geral para modelos mais especializados, concebidos para tarefas bancárias específicas. O objetivo é obter melhores resultados em operações bancárias específicas, com maior eficiência no processamento.

Os primeiros testes já demonstram o potencial desta estratégia. Em casos de uso experimentais selecionados, esta abordagem permitiu reduzir os custos de processamento de cerca de 5 euros por interação para aproximadamente 0,03 euros, mantendo um elevado nível de desempenho. Num projeto-piloto de atendimento ao cliente, o Santander conseguiu também melhorar a qualidade das conversas e reduzir os custos de serviço em mais de 50%.

No âmbito do Lab, o Santander criou o Model Arena, uma plataforma que avalia modelos de pesos abertos (*open-weight*) em tarefas bancárias e identifica o modelo mais adequado para cada caso de uso. O Banco está também a especializar modelos com recurso ao NVIDIA NeMo Customizer. Os testes comparativos realizados no atendimento ao cliente demonstraram que modelos mais pequenos e especializados podem alcançar um elevado desempenho em tarefas

bancárias específicas. O Santander desenvolveu ainda inteligência especializada para processos centrais da atividade bancária com base no NVIDIA Kumo, uma família de modelos fundacionais de pesos abertos para dados estruturados, que permite às empresas transformar rapidamente estes dados em informação preditiva necessária à prestação de serviços suportados por agentes de IA.

O Lab está a industrializar a inferência na infraestrutura tecnológica do Banco através do NVIDIA Dynamo-Triton, a solução que está na origem da redução dos custos operacionais anteriormente referida. Em paralelo, está a desenvolver o Model Arena para conceber, simular, avaliar e supervisionar agentes bancários, complementando este trabalho com iniciativas de otimização de desempenho e testes adversariais a sistemas de IA conversacional em ambientes controlados. O objetivo é estabelecer um processo rigoroso de experimentação que permita otimizar e testar as tecnologias emergentes de IA face aos exigentes requisitos do setor bancário, antes de as capacidades que demonstrem valor serem implementadas à escala em todo o Grupo Santander.

A partir desta experiência, o Santander está a evoluir de modelos de pesos abertos para modelos especializados e agentes bancários sujeitos a mecanismos de governação, concebidos para proporcionar maior controlo e segurança, bem como custos operacionais mais baixos à escala.

Para dar continuidade a este trabalho, o Santander e a NVIDIA estabeleceram um acordo-quadro de colaboração com a duração de três anos, que abrange áreas como modelos fundacionais, agentes de IA, capacidades de prevenção da fraude e de combate ao branqueamento de capitais, tomada de decisões em tempo real e tecnologias avançadas de simulação, incluindo a utilização do conceito de “digital twin” no contexto do funcionamento do Banco.

Iñaki Bernal, responsável global de Inteligência Artificial do Santander, afirmou: “A NVIDIA é um parceiro fundamental nas infraestruturas, no software e na investigação, impulsionando de forma significativa os mais recentes avanços em inteligência artificial. Ao conjugarmos estas capacidades com a escala do Santander em áreas de negócio, clientes e geografias, acreditamos estar numa posição única para transformar, com sucesso, esses avanços em capacidades bancárias concretas.”

Aser Blanco, responsável global pelo setor bancário na NVIDIA, afirmou: “A banca é um dos setores mais intensivos em dados. Os dados estão no centro de todos os processos e, agora, os

bancos precisam de passar a estar centrados na inteligência do que é possível fazer com esses dados. Transformar anos de dados em inteligência útil e proprietária não depende da dimensão do modelo, mas do treino e do aperfeiçoamento. Neste novo contexto, o custo por resultado de inteligência é uma métrica fundamental. O Santander está a demonstrar como esta abordagem funciona em produção, com modelos mais pequenos, especializados com os dados do próprio Banco, avaliados em tarefas bancárias reais e executados com a eficiência necessária para operar à escala de um banco global."