

Adoção da IA em uma empresa demanda uma cultura que aceite riscos

Adoção da IA em uma empresa demanda uma cultura que aceite riscos (IA é probabilístico e, portanto, existe sempre a possibilidade de erros, e o que varia é o grau de assertividade dos modelos) e propensão a aceitar um mind set “data-driven”, apoiando decisões e ações nos processos. Se a empresa não estiver aberta, dificilmente veremos projetos, por mais fantástico que pareçam, serem aceitos de forma ampla pela organização. Provavelmente ficarão nos protótipos, não entrando em produção e não gerando valor.

Uma organização, por mais reativa que seja, pode aos poucos incorporar conceitos e princípios que permitam maior flexibilidade e inovação, e com isso as experimentações com IA começam a chamar atenção, facilitando sua disseminação pela empresa. Um dos pontos essenciais à essas mudanças é educação no tema. Disseminar conhecimento sobre conceitos e permitir que mais e mais pessoas nas empresas usem IA, principalmente agora que existem soluções fáceis de utilizar como sistemas LLM (Large Language Models) como chatGPT e outros, acessados via simples browsers.

Os primeiros projetos devem mostrar claramente que têm objetivos e expectativas realistas. Criar um overhype vai levar à frustração e conseqüentemente, a aumentar a reação contrária. Recentemente o WSJ publicou uma entrevista interessante com o CIO da Goldman Sachs, “Goldman Sachs CIO Tests Generative AI”. Uma pesquisa da KPMG, feita nos EUA, apontou que cerca de 65% dos executivos entrevistados acreditam que a IA generativa terá um grande impacto em sua organização nos próximos três a cinco anos, e 77% disseram que teria um impacto maior na sociedade do que qualquer outra tecnologia emergente durante esse período. Mas 60% disseram que estão a um ou dois anos de antes de implementar sua primeira solução de IA generativa devido a barreiras como talento, custo e privacidade de dados.

Setores que lidam com muita informação como bancos será bastante afetada e a velocidade nesse cenário competitivo é essencial. Assim, o Goldman Sachs quer reduzir esse tempo para adoção, mas por ser um setor altamente regulado, precisa ser cauteloso quanto à aderência às regulações e questões de privacidade, direito autoral e valores éticos. Uma das primeiras atividades que o banco iniciou e que ajuda a resolver um problema real, que é escassez de talentos de desenvolvimento de sistemas, seria entender como usar adequadamente ferramentas LLM como o Copilot como tornar os desenvolvedores mais produtivos. Com a automação de código (ou pelo menos grande parte dele) pode-se obter alguns ganhos de eficiência bastante rápidos. Tirando tarefas repetitivas, de nível júnior dos seniores, permite que esses se concentrem nos objetivos do projeto e no que realmente é importante para o cliente.

O banco tem diversas outras provas de conceitos em andamento. Nada ainda está

em fase de produção. Uma dessas iniciativas, é a classificação e categorização de documentos. No setor financeiro, os bancos recebem milhões de documentos que geralmente estão na forma de contratos legais para produtos de qualquer tipo, como um título, um empréstimo, um derivativo, etc. Precisam extrair as informações para torná-las compreensíveis para uma máquina, para depois tomar decisões com base nos dados processados. E uma pergunta básica que sempre é feita ao receber um documento é: que tipo de documento é esse? O Goldman já estava usando a IA tradicional para fazer isso, mas agora estão testando os LLMs como uma maneira potencial de subir a um nível mais elevado, como fazendo resumos desses documentos para facilitar ações e tomadas de decisão.

Os resultados dessas experimentações parecem promissores. Quando se trata de classificar documentos, as acurácias se mostram animadoras, pelo menos tão bons quanto os humanos. Quando se trata de codificação, até o momento, a prova de conceito não é grande o suficiente para ter certeza sobre esses números, mas os primeiros resultados mostram que até 40% do código escrito pela IA é aceito pelo desenvolvedor. Já é um bom resultado, pois mostra que pelo menos quase metade do código não precisa mais ser escrito pelo desenvolvedor.

Na opinião deles, a maior restrição agora é que há muitas coisas que ainda não sabemos. Precisam adaptar as estruturas de controle para garantir que não haja nenhuma ameaça para o banco ou para os clientes. Isso exige cautela. Um problema mais fundamental é que ainda não existem maneiras de especificar requisitos formalmente e verificar se esses requisitos são atendidos. Portanto, atualmente é impossível saber se o comportamento de um programa gerado por IA corresponde ao que ele deveria fazer. Um problema relacionado é que o código que essas ferramentas de IA produzem não é necessariamente otimizado para nenhum atributo específico, como escalabilidade. Embora seja possível conseguir isso com os prompts corretos, isso levanta a questão de como compor esses prompts”

Os defensores da IA argumentarão corretamente que os humanos cometem os mesmos erros. Mas podemos pensar em três razões pelas quais o código gerado por IA deve ser mantido em um padrão mais alto. Primeiro, porque o processo de treinamento utiliza o corpus de todo o código acessível ao público e porque não há critérios diretos para julgar a qualidade, você simplesmente não sabe o quão bom é o código gerado pela IA. Em segundo, as pessoas tendem a acreditar que o código gerado por computador estará livre de problemas, então elas podem ser menos cuidadosas. E, em terceiro, como as pessoas que usam essas ferramentas não criaram o código elas mesmas, elas podem não ter as habilidades para depurá-lo ou otimizá-lo. A mensagem que tiramos é que o uso eficaz da IA exigirá que as empresas desenvolvam controles de qualidade e validação que identifiquem as falhas e vulnerabilidades e as removam. Não use o código sem passar por esses crivos.

Portanto, as preocupações com privacidade e aderência às regras e compliance são essenciais. Um bom ponto de partida para entender o assunto é o estudo da IBM, “A Policymaker’s Guide to Foundation Models”. Recomendo enfaticamente sua leitura.

Além disso, existe o fator talentos. Há poucas pessoas que conhecem IA em geral, e há ainda menos pessoas que conhecem uma particularidade da IA, que são os LLMs. Então, um dos principais desafios é como treinar e capacitar o pessoal que estará envolvido com essas tecnologias.

É um bom case para ser estudado. Vemos nitidamente que o fator “change management” se torna relevante. O uso de IA de forma massiva implica em mudanças de processos, práticas, eliminação de tarefas e criação de outras e, sem dúvida, a capacitação do pessoal interno e eventualmente aquisição de novos talentos.

Os projetos de IA, com ou sem LLM, precisam de pessoal técnico e engajamento dos líderes e profissionais dos setores de negócios. É necessário identificar que áreas de negócio poderiam ser as primeiras a serem usadas em provas de conceitos e isso, obviamente, depende do engajamento de seus líderes. As expectativas precisam ser realistas e os resultados obtidos devem ser amplamente divulgados, para que a organização aprenda com os acertos e erros. As ações de IA devem focar não apenas em resultados de curto prazo, mas ter a visão mais estratégica, de futuras possíveis utilizações, que, claro, só acontecerão se as primeiras apresentarem valor palpável para o negócio.

Um dos grandes desafios das empresas é a capacitação de talentos. Como usar sistemas LLM de forma adequada e segura? Existem vários desafios que dificultam a entrada em produção de aplicativos baseados em sistemas LLM, como:

1. A ambigüidade das linguagens naturais, que pode proporcionar uma experiência agradável para o usuário, mas uma experiência dolorosa para o desenvolvedor.
2. A natureza estocástica dos LLMs que pode levar à inconsistência na experiência do usuário e a falhas silenciosas, que passam despercebidas em leituras rápidas.
3. A rapidez com que as tecnologias LLM estão evoluindo, o que torna muito difícil tomar decisões de negócios: análise de custo-latência, construção (usando código aberto) x compra (usando APIs pagas), solicitação x ajuste fino x técnicas mais recentes precisam ser feitas constantemente. Recomendo a leitura do paper “Building LLM applications for production” para uma visão prática e mais aprofundada desses desafios.

Além disso, existe, disseminada pela mídia, o receio que a IA vai eliminar empregos. Volta e meia nos deparamos com afirmativas bombásticas que os robôs e a IA acabarão com os empregos e destruirão a humanidade. Em 1965, o prêmio Nobel de Ciências Econômicas, Herbert Simon predisse: “machines will be capable within 20 years

of doing any work a man can do.” Ele projetou o cenário para 1985! De lá para cá vários livros e estudos abordaram o mesmo assunto. E 60 anos depois ainda não aconteceu. E, aliás, ainda está bem longe de acontecer.

Exercitarmos o pensamento crítico é essencial. Um tecno-otimismo sem visão dos desafios do mundo real é fantasioso. LLMs são excelentes ferramentas, mas são geradores de texto.

Não são projetados para avaliar a veracidade do texto que eles inserem e produzem. Eles também não usam qualquer tipo de pensamento crítico, e é por isso que trabalhos que exigem pensamento crítico e acurácia das respostas não poderão ser substituídos por LLMs. Serão úteis para ajudar em muitas tarefas, e aumentar a produtividade. Mas são serão substitutos.

A preocupação com capacitação e a forma como IA é divulgada na empresa vai afetar em muito o entusiasmo ou rejeição. Os funcionários têm expectativas diversas, estão em pontos diferentes de suas trajetórias profissionais e, portanto, interesses diversos. Uma educação estilo “one size fits all” nem sempre funciona. É um ponto que merece muita atenção. A educação em IA demanda investimentos e tempo. Se a empresa está realmente engajada em adotar IA de forma mais intensa, esse cuidado é essencial.

IA não é obviamente apenas tecnologia. Demanda liderança engajada, cultura “data-driven” e aberta às experimentações e erros. A tecnologia é a parte mais fácil da equação. O maior desafio está na mobilização e engajamento dos funcionários e líderes, para que experimentem e adotem essas ferramentas.