

COMPARANDO O DESEMPENHO DE CHATGPT, GOOGLE GEMINI E META AI EM PERGUNTAS SOBRE GESTÃO DE PROJETOS

COMPARING THE PERFORMANCE OF CHATGPT, GOOGLE GEMINI AND META AI ON PROJECT MANAGEMENT QUESTIONS

LEANDRO KEI OGURA

USP - UNIVERSIDADE DE SÃO PAULO

ANDERSON ANTÔNIO DE LIMA

CENTRO UNIVERSITÁRIO SENAC

Comunicação:

O XIII SINGEP foi realizado em conjunto com a 13th Conferência Internacional do CIK (CYRUS Institute of Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

COMPARANDO O DESEMPENHO DE CHATGPT, GOOGLE GEMINI E META AI EM PERGUNTAS SOBRE GESTÃO DE PROJETOS

Objetivo do estudo

Analisar comparativamente a acurácia de três chatbots de inteligência artificial gratuitos (ChatGPT, Google Gemini e Meta AI) na resolução de questões específicas sobre gestão de projetos, identificando padrões de dificuldades em categorias temáticas distintas.

Relevância/originalidade

Constitui o primeiro estudo empírico comparativo simultâneo entre três plataformas de IA conversacional aplicadas ao domínio da gestão de projetos. Preenche lacuna significativa na literatura sobre efetividade de ferramentas de IA generativa em contextos organizacionais específicos.

Metodologia/abordagem

Emprega delineamento de levantamento de campo mediante aplicação de 147 questões de múltipla escolha, baseadas em Oguz (2022). Utiliza abordagem metodológica mista, combinando análise quantitativa (teste-z) com avaliação qualitativa do detalhamento das respostas obtidas.

Principais resultados

ChatGPT e Meta AI demonstraram taxa de conformidade de 93,2%, enquanto Google Gemini alcançou 90,48%. A análise estatística revelou ausência de diferenças significativas entre os desempenhos ($p>0,05$). Consenso em 82,55% das questões pelos três sistemas.

Contribuições teóricas/metodológicas

Desenvolve framework metodológico para avaliação sistemática de chatbots em domínios especializados. Evidencia limitações dos modelos de linguagem em processamento matemático e raciocínio estratégico organizacional. Propõe protocolo de análise de adequação ao padrão para validação.

Contribuições sociais/para a gestão

Fornece diretrizes empíricas para tomada de decisão organizacional na adoção de ferramentas de IA. Demonstra viabilidade técnica dos chatbots como instrumentos de apoio gerencial. Identifica características qualitativas distintivas para seleção contextualizada das plataformas disponíveis.

Palavras-chave: Gestão de Projetos, Inteligência Artificial, Chatbot

COMPARING THE PERFORMANCE OF CHATGPT, GOOGLE GEMINI AND META AI ON PROJECT MANAGEMENT QUESTIONS

Study purpose

To comparatively analyze the accuracy of three free artificial intelligence chatbots (ChatGPT, Google Gemini, and Meta AI) in solving specific project management questions, identifying patterns of difficulties across distinct thematic categories and performance variations.

Relevance / originality

Constitutes the first empirical comparative study simultaneously examining three conversational AI platforms applied to project management domain. Addresses significant gap in literature regarding effectiveness of generative AI tools in specialized organizational contexts and decision-making processes.

Methodology / approach

Employs field survey design through application of 147 multiple-choice questions based on Oguz (2022). Utilizes mixed methodological approach, combining quantitative analysis (z-test) with qualitative evaluation of response detail and comprehensiveness across the three platforms.

Main results

ChatGPT and Meta AI demonstrated 93.2% conformity rate, while Google Gemini achieved 90.48%. Statistical analysis revealed no significant differences between performances ($p > 0.05$). Consensus achieved in 82.55% of questions by all three systems, indicating substantial agreement.

Theoretical / methodological contributions

Develops methodological framework for systematic evaluation of chatbots in specialized domains. Evidences language model limitations in mathematical processing and organizational strategic reasoning. Proposes pattern-matching analysis protocol for validation in domain-specific applications and accuracy assessment.

Social / management contributions

Provides empirical guidelines for organizational decision-making in AI tool adoption. Demonstrates technical feasibility of chatbots as managerial support instruments. Identifies distinctive qualitative characteristics enabling contextualized platform selection based on organizational needs and user requirements.

Keywords: Project Management, Artificial Intelligence, Chatbot

COMPARANDO O DESEMPENHO DE CHATGPT, GOOGLE GEMINI E META AI EM PERGUNTAS SOBRE GESTÃO DE PROJETOS

1 Introdução

Há mais de sete décadas, Turing (1950) propôs a pergunta: “As máquinas podem pensar?” e descreveu o jogo da imitação, um teste que nos permitiria entender se uma máquina pode demonstrar um nível de inteligência indistinguível daquele demonstrado por um ser humano. Desde então, o campo da inteligência artificial (IA) se expandiu consideravelmente. Tlili et al. (2023) corroboram essa perspectiva, sugerindo que estamos à beira de uma nova era em que a IA terá um impacto profundo em nossas vidas.

De acordo com Meredith et al. (2015), a gestão de projetos é uma área fundamental para o sucesso organizacional, demandando precisão, planejamento cuidadoso e tomada de decisão baseados em informações confiáveis.

A importância deste estudo se justifica pelo fato de os chatbots inteligência artificial serem ferramentas novas e promissoras para a gestão de projetos. Segundo Belharet et al. (2020), tecnologias de IA vão transformar sistemas de gestão empresarial em plataformas que incrementam as habilidades humanas e as grandes organizações devem estar preparadas para mudanças significativas em suas metodologias e abordagens de gerenciamento de projetos. Estudos pioneiros como o de Minelle e Stolfi (2023) demonstram que o ChatGPT é capaz de auxiliar gerentes de projeto. No entanto, ainda há carência de pesquisas que comparem a efetividade de diferentes chatbots em tarefas específicas à área de gestão de projetos. Uma revisão da literatura mostrou que não foram encontradas pesquisas que avaliassem simultaneamente três chatbots quanto à precisão em responder perguntas específicas sobre gestão de projetos. Essa lacuna evidencia a relevância deste estudo, que busca explorar um novo horizonte ao comparar o desempenho de três ferramentas amplamente utilizadas: ChatGPT, Google Gemini e Meta AI.

Diante da relevância da temática, este estudo tem o objetivo de identificar qual chatbot de IA gratuito apresenta melhor desempenho ao lidar com perguntas sobre gestão de projetos, além de identificar potenciais tipos de perguntas ou temas específicos onde os chatbots encontram dificuldades. O conhecimento gerado será crucial para informar a escolha e o uso dessas ferramentas nas organizações, otimizando a eficiência na gestão de projetos e contribuindo para o sucesso de iniciativas organizacionais.

2 Referencial Teórico

A fim de estabelecer um referencial teórico, esta seção explora a evolução histórica da IA, traçando a evolução dos chatbots, desde seus primórdios até os modelos de linguagem de grande escala atuais. Paralelamente, busca mapear a evolução da gestão de projetos, com um olhar também para a convergência entre a gestão de projetos e a inteligência artificial.

2.1 Inteligência Artificial (IA), seu surgimento e evolução.

De acordo com Russell e Norvig (2021), a Inteligência artificial (IA) é um campo abrangente e em constante evolução, que historicamente foi definido como um campo dedicado à criação de máquinas capazes de simular o pensamento ou comportamento humano, ou que demonstrem racionalidade. A racionalidade é entendida como a capacidade de agir de forma correta, onde o que é considerado correto é definido pelo objetivo que fornecemos. Essa racionalidade pode se manifestar tanto por meio de processos internos quanto de

comportamentos externos. Ainda segundo Russel e Norvig, a definição mais difundida é a que define a IA como o campo que foca no estudo e criação de agentes que agem da forma correta. Este entendimento é chamado de modelo padrão da IA.

McCorduck (2004) considera a conferência Dartmouth Summer Research Project on Artificial Intelligence de 1956 como o marco de nascimento da IA como campo formal de pesquisa. Essa conferência estabeleceu o termo 'Inteligência Artificial' e impulsionou o progresso da área por meio da colaboração entre pesquisadores de renome.

Desde então, a IA percorreu um longo caminho, passando por diferentes abordagens, sendo as principais a abordagem simbólica, que é baseada em regras algorítmicas explícitas e a abordagem conexionista, baseada em redes neurais.

Segundo Garnelo e Shanahan (2019), a abordagem da IA Simbólica (ou clássica) foi dominante até o final dos anos 80. Eles denotam que esta abordagem tem como princípio central a ideia de que a inteligência resulta da manipulação de representações abstratas, ou seja, símbolos. Um sistema de IA simbólica funciona ao executar operações de raciocínio lógico sobre representações linguísticas. Os sistemas contêm instruções explícitas, em linguagem de programação, que guiam suas operações.

Wooldridge (2021) afirma que do fim dos anos 70 até o início dos anos 90, houve diversos momentos de otimismo e outros de crise na área de pesquisa da IA até que se tornou evidente que o modelo simbólico tinha limitações significativas e seria improvável que fosse possível criar sistemas capazes de simular a racionalidade humana somente através desta abordagem.

A abordagem conexionista, também chamada de abordagem de redes neurais surgiu antes do surgimento do próprio campo da IA e vem evoluindo desde então, alcançando resultados surpreendentes neste século e tornando-se a abordagem dominante.

Haykin (1999) define redes neurais artificiais como sistemas computacionais inspirados na estrutura e no funcionamento do cérebro humano, compostos por unidades de processamento interconectadas chamadas neurônios artificiais. As redes neurais simulam o aprendizado humano ao ajustar a força das conexões entre os neurônios com base em dados de experiências, permitindo que uma máquina aprenda sem que seja explicitamente programada.

2.2 Ascensão dos chatbots com IA Generativa: ChatGPT, Gemini e Meta AI

Al-Amin et al. (2024) afirmam que o termo 'chatbot', uma fusão de 'chat' e 'robot', tem origem na sua função inicial como sistema de computador de diálogo baseado em texto, buscando simular a linguagem humana. Segundo Shum et al. (2018), os primeiros chatbots, como ELIZA criado por Joseph Weizenbaum em 1966, e o Parry, criado por Kenneth Colby em 1975 eram sistemas baseados em regras criadas manualmente e só funcionam bem em contextos específicos. O próprio Weizenbaum (1966) descreve ELIZA como um programa de computador para o estudo da comunicação em linguagem natural entre a máquina e o ser humano. Um programa que funciona através de regras que buscam palavras-chave na mensagem do seu interlocutor humano, e gera respostas baseadas nas palavras-chave. Essas respostas criam uma ilusão de que o programa compreende o que está sendo dito pelo seu interlocutor.

Cao et al. (2023) afirmam que nos anos recentes ocorreram tremendos avanços em IA generativa. Isso levou ao desenvolvimento de uma ampla gama de aplicações, desde processamento de imagens a processamento de linguagem natural.

Al-Amin et al. (2024) conta que no início da década de 2020 vimos uma rápida expansão do mercado de chatbots de IA, com o surgimento do ChatGPT, da OpenAI e do Gemini, inicialmente chamado de Bard, da Google. Estes chatbots são treinados, ou seja aprendem, a

partir de uma extensa quantia de dados textuais, o que os habilita a reconhecer padrões de linguagem. Ambos utilizam uma arquitetura de rede neural chamada de transformer. A arquitetura transformer lida com processamento de linguagem natural e habilita os chatbot a gerar respostas coerentes que levam em conta o contexto da conversa.

Segundo (Brown et al., 2020), os modelos de linguagem treinados em uma quantidade imensa de dados, os chamados Large Language Models (LLMs) ou modelos de linguagem de grande escala, são capazes de responder perguntas sobre uma ampla gama de domínios de forma adaptável. Cao et.al (2023) denota que o surgimento desses grandes modelos de linguagem como o ChatGPT mudou o paradigma da indústria e do meio acadêmico sobre como desenvolver sistemas de aprendizado de máquina da próxima geração.

De acordo com Labadze et al. (2023), o ChatGPT e o Google Gemini estão entre os mais impactantes chatbots baseados em IA.

O ChatGPT foi lançado em novembro de 2022 pela OpenAI, enquanto o Google Gemini foi lançado em março de 2023 pelo Google com o nome Bard e posteriormente rebatizado como Google Gemini. Já o Meta AI foi lançado mais recentemente, em abril de 2024. Desde então, as capacidades e funcionalidades dos três vêm sendo aprimoradas.

Rane et al. (2024) explicam que, quando consideramos ChatGPT e Gemini, cada um desses LLMs apresenta características distintas que determinam sua aplicabilidade. O Gemini se destaca pela integração com a Pesquisa Google, garantindo maior precisão factual, mostrando-se particularmente valioso para pesquisa, análise e tomada de decisão baseada em evidências, principalmente em contextos que demandam amplo escrutínio de dados e planejamento estratégico. Por outro lado, o ChatGPT brilha na fluidez conversacional e expressão criativa, tornando-o ideal para aplicações de engajamento com o cliente, marketing e recursos humanos. Com isso, a escolha entre Gemini e ChatGPT depende das necessidades do usuário. É importante notar que ambos Gemini e ChatGPT evoluem constantemente e novas atualizações podem mudar essas comparações ao longo do tempo.

2.3 Surgimento e evolução de Projetos e Gestão de Projetos

Entre o final do século XIX e o início do século XX, a revolução industrial catalisou a evolução das práticas de gestão de projetos. O aumento da complexidade dos empreendimentos industriais gerou uma demanda crescente por métodos mais eficazes de planejamento e controle. Segundo Morris (1994), durante esse período, houve uma transição das abordagens informais para métodos mais científicos. O desenvolvimento de ferramentas como o diagrama de Gantt, por exemplo, foi fundamental para melhorar o planejamento e o controle de projetos, respondendo às novas demandas por eficiência. Ainda com base em Morris (1994), a gestão de projetos moderna, mais formal e estruturada, surgiu entre as décadas de 1930 e 1950, com uma relação próxima ao desenvolvimento da engenharia de sistemas na indústria de defesa dos EUA. Moder et al. (1983) indicam também que o desenvolvimento da computação e teorias gerais de sistemas também prepararam o terreno para novas metodologias. Os autores indicam que duas das metodologias mais influentes desenvolvidas nesse período foram o PERT (Program Evaluation and Review Technique) e o CPM (Critical Path Method).

O PERT foi desenvolvido para gerenciar o desenvolvimento do míssil balístico Polaris para a Marinha dos EUA. Este método introduziu a ideia de análise probabilística do tempo de conclusão das atividades, permitindo uma melhor gestão de incertezas. Meredith et al. (2015) indicam que o CPM, método desenvolvido pela DuPont Corporation, foca na identificação do caminho crítico em um projeto, melhorando também a gestão do tempo e a priorização de atividades. A adoção destas ferramentas em diversas indústrias consolidou a mudança em direção a uma abordagem sistemática e analítica para a gestão de projetos.

2.4 O cenário atual da gestão de projetos e tendências para o futuro

A disciplina de gestão de projetos continua a evoluir, adaptando-se ao cenário dinâmico do Século XXI. Maylor (2010) indica que o desenvolvimento e adoção de softwares de gestão de projetos vieram com o potencial de mudar a forma com que se trabalha em projetos.

Bergmann e Karwowski (2017) denotam que o gerenciamento de projeto ágil (Agile) emergiu recentemente como uma abordagem inovadora para a gestão de projeto que prioriza flexibilidade, desenvolvimento iterativo e estreita colaboração com o cliente. A metodologia Agile de gestão de projetos surgiu como um novo paradigma, em contraponto aos modelos mais tradicionais, comumente denominados waterfall. Nesse contexto, surgem também os modelos híbridos de gestão de projetos, incorporando elementos de métodos Agile e Waterfall. Dionisio (2022) explica que o gerenciamento híbrido combina diferentes metodologias para alcançar os objetivos do projeto.

Sobre o futuro da disciplina de gestão de projetos, é evidente que o campo continuará a evoluir. A integração de tecnologias emergentes como a Inteligência artificial tem o potencial de transformar as práticas de gestão de projetos. Taboada et al. (2023) alegam que o gerenciamento de projetos se beneficiaria da inteligência artificial para alcançar objetivos de projeto, melhorando o desempenho do projeto e, conseqüentemente, alcançando uma maior taxa de sucesso. Fridgerisson et al. (2021) afirmam ser evidente que a IA pode acompanhar cronogramas, revisar previsões e manter padrões de referência.

2.7 Uso de Chatbots de IA para auxiliar Gestores de Projetos

Os chatbots de inteligência artificial, que utilizam grandes modelos de linguagem, apresentam um grande potencial no que diz respeito ao auxílio a gestores de projetos na realização de tarefas em diversas fases dos projetos. Bainey (2024) diz que o ChatGPT pode ajudar a automatizar fluxos de trabalho, iniciar documentos fornecendo rascunhos e modelos, prover conclusões a partir de dados, identificar riscos, ajudar na tomada de decisões, tudo isso, é claro, com supervisão humana. Karamthulla et al. (2024) citam alguns exemplos de como o processamento de linguagem natural dos chatbots pode automatizar tarefas. Os exemplos citados são fornecer resumo de minutas de reunião, extrair de itens de ação, identificar informações relevantes a partir de documentação, detectar erros e inconsistências. Lakshminarasimham (2024) indica que a tomada de decisões com a ajuda de inteligência artificial melhora a qualidade das decisões ao fornecer conclusões a partir de dados e previsões melhores do que a de métodos tradicionais. Hezz e Kunz (2024) mostram alguns exemplos práticos de como os chatbots de IA podem ajudar a trazer ideias à tona, gerar conclusões, fornecer conselhos e instruções, criar documentação e diagramas, criar material de comunicação, ajudar no estudo e treinamento do usuário através de simulações e analisar sentimentos.

Através deste referencial teórico, conseguimos notar como estamos em um ponto de convergência entre as tecnologias de inteligência artificial, em especial os chatbots de IA e a prática de gestão de projetos. A integração de ferramentas de IA generativa na gestão de projetos representa um novo capítulo na evolução da disciplina, abrindo um caminho promissor para maiores níveis de eficiência na área. Esta pesquisa se insere nesse contexto, investigando o desempenho de três modelos específicos de chatbot (ChatGPT, Gemini e Meta AI) em tarefas relacionadas à gestão de projetos. Ao identificar qual chatbot apresenta melhor performance, o estudo contribuirá para a compreensão prática da aplicação da IA na gestão de projetos.

3 Metodologia

Esta pesquisa adota o delineamento de levantamento de campo para comparar a acurácia do ChatGPT, Google Gemini e Meta AI em responder questões sobre gestão de projetos. O estudo foi conduzido com o objetivo de analisar, de forma qualitativa e quantitativa, o desempenho dessas ferramentas de IA em termos de precisão e detalhamento das respostas.

O instrumento de coleta de dados utilizado foi um questionário composto por 147 perguntas de múltipla escolha, extraído da literatura especializada em gestão de projetos. As perguntas para o questionário foram extraídas do livro “Project Management, Navigating the Complexity with a Systematic Approach” de Abdullah Oguz que está licenciado sob uma licença Atribuição-Não Comercial do padrão Creative Commons 4.0. Esta licença permite rearranjar, adaptar e criar a partir deste trabalho, desde que não seja para fins comerciais e que o devido crédito seja atribuído. O questionário é composto de 147 perguntas de múltipla escolha, com 4 alternativas cada, cobrindo uma ampla gama de tópicos dentro da área de conhecimento de gestão de projetos..

Em seguida, as perguntas do questionário foram inseridas separadamente nas conversas com cada uma das três ferramentas de IA (ChatGPT, Google Gemini e Meta AI) e as respostas geradas pelos modelos de linguagem foram registradas. Para garantir que os chatbots estivessem começando em pé de igualdade, foram criadas contas novas, a partir de um email novo tanto no ChatGPT quanto no Gemini. Para a Meta AI, também foi utilizada uma conta que não tinha tido interação prévia com chatbot. Nenhum contexto anterior foi passado aos chatbots e as perguntas foram feitas uma a uma, com cada pergunta representando um prompt. Foram, portanto, 147 prompts para cada chatbot.

No momento da coleta das respostas, o ChatGPT, em sua versão gratuita, que foi utilizada empregava primariamente o modelo de linguagem GPT-4o mini. O Gemini utilizou o modelo 1.5 Flash e o Meta AI, por sua vez, utilizou o modelo Llama 3.1 405B. As ferramentas foram utilizadas em suas versões Web, acessadas a partir de um navegador nos endereços: <https://chatgpt.com>, <https://gemini.google.com> e <https://meta.ai/>.

As respostas foram avaliadas e então comparadas com um gabarito previamente estabelecido de acordo com as respostas indicadas como corretas no livro Project Management de Oguz (2022) de onde as perguntas do questionário foram extraídas, empregando, a análise de adequação ao padrão. Essa metodologia, conforme definida por Trochim (1989) visa determinar se um objeto ou comportamento se conforma a um padrão pré-definido. No presente caso, o padrão utilizado é a literatura de referência em gestão de projetos e o comportamento a ser analisado é a resposta obtida dos chatbots.

A análise dos dados foi realizada de forma a empregar abordagem quantitativa e qualitativa. No componente quantitativo, a análise foi feita calculando-se a taxa de conformidade de cada chatbot em relação às respostas indicadas como corretas no livro-texto. No componente qualitativo, foram realizadas análises para observar o grau de detalhamento e clareza das respostas, identificando as diferenças nas abordagens de cada ferramenta. Além disso, foram analisadas de forma qualitativa as perguntas onde os três chatbots responderam incorretamente de acordo com o livro-texto.

Dentro da abordagem quantitativa, para avaliar a significância das diferenças entre os desempenhos dos chatbots, foi utilizado o teste estatístico de z, que permitiu verificar se as variações observadas nas taxas de conformidade com o gabarito entre as ferramentas são estatisticamente significativas.

De acordo com Triola (2019), o teste-z é uma ferramenta estatística adequada para comparar proporções em amostras grandes, como neste caso. Esse método é especialmente relevante em estudos em que as variáveis são categóricas, como respostas classificadas como "corretas" ou "incorretas". O procedimento será realizado conforme os descritos a seguir:

Serão definidas as hipóteses, sendo que a hipótese nula (H_0) se refere a taxas de

conformidade com o gabarito entre os chatbots não apresentando diferença estatisticamente significativa e a hipótese alternativa (H_1), se referindo a quando as taxas de conformidade apresentam diferença estatisticamente significativa. Para calcular o valor do teste-z, será utilizada a fórmula apresentada por Montgomery e Runger (2018): $Z = (p_1 - p_2) / \sqrt{p^*(1-p)^*(1/n_1 + 1/n_2)}$, em que:

- p_1 e p_2 representam as proporções de número de respostas em conformidade com o gabarito dos chatbots sendo comparados.
- p^* é a proporção combinada de número de respostas em conformidade com o gabarito nas duas amostras ($p^* = (x_1 + x_2) / (n_1 + n_2)$).
- x_1 = número de respostas em conformidade com o gabarito de um chatbot
- x_2 = número de respostas em conformidade com o gabarito do outro chatbot

Seguindo a recomendação de Field (2018), será adotado um nível de significância de 0,05, que corresponde a um intervalo de confiança de 95%. A hipótese nula será rejeitada se o valor absoluto de Z for maior que o valor crítico correspondente ($z_{\alpha/2}$).

O uso do teste-z garante a aplicação de uma abordagem estatística robusta, assegurando que as conclusões obtidas não sejam apenas aparentes, mas embasadas por critérios objetivos. Essa metodologia complementa as análises qualitativas realizadas e proporciona uma visão mais completa sobre o desempenho dos modelos analisados.

Creswell (2010) caracteriza da seguinte forma as abordagens de pesquisa, a abordagem qualitativa explora significados atribuídos a problemas, usando métodos flexíveis e interpretando dados. Já a quantitativa visa testar teorias, analisando relações entre variáveis medidas por meio de estatística. Em alguns casos, a pesquisa mista combina as duas abordagens, obtendo uma visão mais completa do problema. Esta pesquisa apresenta uma abordagem mista, com características de ambas as abordagens qualitativa e quantitativa. A abordagem qualitativa permitiu interpretar as respostas dos chatbots, analisando não só se estavam de acordo com a literatura de referência, mas também o significado, a relevância e o nível de detalhamento de cada uma. Já a abordagem quantitativa foi usada para classificar as respostas como corretas ou incorretas, gerando resultados em forma de taxas de conformidade. Essa combinação de métodos possibilitou uma avaliação mais completa, juntando a interpretação mais aprofundada da análise qualitativa com a objetividade dos dados quantitativos, o que ajudou a comparar o desempenho das ferramentas de forma clara e consistente.

4 Análise dos Resultados e Discussão

Ao coletar as respostas que o ChatGPT, o Gemini e o Meta AI apresentaram em resposta às 147 perguntas de múltipla escolha do livro Project Management de Abdullah Oguz, notamos que os três chatbots alcançam um alto nível de precisão, com o ChatGPT e o Meta AI apresentando a mesma taxa de conformidade com o gabarito de 93.2%, 2.72 pontos percentuais acima do resultado apresentado pelo Gemini, que foi de 90.48%.

Com base nos dados coletados, foi realizado um teste-z para entender se o resultado obtido tem significância estatística. O ChatGPT obteve 137 respostas “corretas”, ou seja, em conformidade com o livro texto de um total de 147 perguntas, o que resulta em uma taxa de conformidade de aproximadamente 93,2% ($p_1=0,9320$). O Google Gemini, por sua vez, obteve 133 respostas alinhadas com o gabarito do livro-texto, o que corresponde a uma taxa de conformidade de 90,48% ($p_2=0,9048$).

Para realizar o teste de significância, foi calculada a proporção combinada de respostas em conformidade com o gabarito (p^*) dos dois chatbots, que foi de 0,9184. Em seguida, utilizou-se a fórmula do teste-z para comparar as duas proporções, sendo que o tamanho da

amostra foi 147 para ambos os grupos. O valor do teste-z foi calculado como 0,8520, conforme a fórmula: $Z = (p_1 - p_2) / \sqrt{[p*(1-p)*(1/n_1 + 1/n_2)]}$.

O valor crítico de z para um nível de significância de $\alpha=0,05$ em um teste bicaudal é 1,96. Como o valor obtido para o teste-z (0,8520) foi menor que o valor crítico de 1,96, não se rejeitou a hipótese nula, indicando que não há uma diferença estatisticamente significativa entre as taxas de conformidade do ChatGPT e do Google Gemini para as perguntas relacionadas à gestão de projetos. Dessa forma, pode-se concluir que, dentro do contexto do estudo, os desempenhos dos dois chatbots foram semelhantes em termos de acurácia ao responder às perguntas propostas.

O teste-z foi calculado exclusivamente para comparar os resultados do ChatGPT e do Google Gemini, pois a taxa de conformidade do Meta AI foi idêntica à do ChatGPT (137 de 147 respostas em concordância com o livro texto), e, portanto, os resultados obtidos para o ChatGPT podem ser extrapolados para o Meta AI. Como o desempenho do Meta AI foi equivalente ao do ChatGPT, não foi necessário realizar um novo teste-z para o Meta AI, uma vez que qualquer diferença significativa ou não significativa entre o ChatGPT e o Gemini também se aplicaria ao Meta AI. Assim, a análise focou apenas na comparação entre ChatGPT e Google Gemini, simplificando o processo.

Houve 123 (82.55%) perguntas às quais todos os chatbots participantes responderam de acordo com o gabarito. 17 (11.41%) perguntas foram respondidas em conformidade com o gabarito por 2 chatbots e em desconformidade com o gabarito por um. 4 (2.68%) perguntas foram respondidas conforme o gabarito por apenas 1 chatbot e desconforme pelos outros 2. Finalmente, houve 3 (2.01%) perguntas respondidas em desconformidade pelos 3 chatbots.

Ao analisar as perguntas onde os 3 chatbots apresentaram respostas não conformes com o gabarito, encontramos algumas informações interessantes. A primeira é que nas três questões que foram respondidas pelos chatbots de forma diferente ao gabarito, os três chatbots concordaram entre si, escolhendo a mesma alternativa como sendo a correta, o que nos leva a questionar se de fato os três chatbots se equivocaram, escolhendo a resposta errada ou se o livro-texto, de onde foi retirado o gabarito é que apresenta as respostas mais distantes do que pode ser considerado correto.

Para cada uma das três perguntas, foi realizada uma análise baseada em outras fontes da literatura para entender se são as respostas consideradas corretas pelo livro-texto ou pelos chatbots que se aproximam mais das informações encontradas na literatura de referência na área de gestão de projetos.

A primeira pergunta em que os três chatbots divergiram do livro-texto “Project Management, Navigating the Complexity with a Systematic Approach” de Abdullah Oguz é a seguinte:

“_____ describe(s) how the organization is going to fulfill its _____ that describes what the organizational leaders want the organization to accomplish.

- A) Value statements; mission
- B) Mission statements; vision
- C) Strategies; vision
- D) Strategic analysis; objectives”

Traduzindo para o português esta questão presente no trabalho de Oguz (2022), temos:

“_____ descreve(m) como a organização vai cumprir a sua _____ que descreve o que os líderes organizacionais querem que a organização realize.

- A) Declarações de valores; missão
- B) Declarações de missão; visão
- C) Estratégias; visão
- D) Análise estratégica; objetivos”

Na questão acima, a alternativa indicada como correta no livro de Oguz é a alternativa B, “Declarações de missão; visão”. No entanto, os três chatbots julgaram que a alternativa correta seria a C, “Estratégias; visão”.

Analisando literatura adicional como referência, pode-se dizer que ambas respostas estão corretas, porém a alternativa B pode ser considerada a melhor resposta. Argumentos contundentes podem ser levantados a favor de cada uma das opções. A um nível superficial, é compreensível que os chatbots tenham escolhido a opção C, já que é uma resposta plausível.

De acordo com Graham e Havlick (1994), declarações de missão são definições amplas e duradouras do propósito de uma organização, que a distinguem das demais em seu setor e definem o escopo de suas operações, tanto em termos de produtos quanto de mercados e encapsula também como ela pretende alcançar sua visão de forma estável e a longo prazo. Já a visão se refere a uma imagem de como o futuro deveria ser, refletindo os resultados organizacionais, assim como padrões de excelência.

Kaplan e Norton (2001) se referem às estratégias como planos mais detalhados que são desenvolvidos e ajustados à medida que as circunstâncias evoluem. As estratégias, que são mais dinâmicas e respondem às condições mutáveis do mundo real, partem da estabilidade proporcionada pelas declarações de missão.

Embora as estratégias descrevam como atingir uma visão, elas carecem da natureza duradoura e holística de uma declaração de missão. Por isso, embora a alternativa C não esteja de todo incorreta, a opção B se encaixa melhor porque as declarações de missão capturam de forma mais abrangente e duradoura uma descrição de como uma organização pretende cumprir a sua visão.

A dificuldade apresentada pelos três chatbots no caso desta pergunta pode se dar por conta de os modelos de linguagem se basearem em reconhecimento de padrões e se basearem nos dados a partir dos quais foram treinados. Essas características dos chatbots podem ter levado-os a associar incorretamente “estratégias” a “visões”, já que os dois termos aparecem com frequência inter relacionados em textos sobre gestão e planejamento.

A próxima pergunta em que os três chatbots divergiram do gabarito do livro-texto é a seguinte:

"According to the table below, which project is the most favorable when we use the weighted scoring model?"

Weights:

Criterion 1: 0.3

Criterion 2: 0.4

Criterion 3: 0.2

Criterion 4: 0.1

Scores for each criterion in an order:

Project A: 3, 2, 3, 1

Project B: 1, 3, 2, 1

Project C: 2, 2, 1, 3

Project D: 1, 1, 3, 3

A) Project A

B) Project C

C) Project D

D) Project B"

Traduzindo para o português, temos:

“De acordo com a tabela abaixo, qual é o projeto mais favorável quando utilizamos o modelo de pontuação ponderada?

Pesos:

Critério 1: 0,3

Critério 2: 0,4

Critério 3: 0,2

Critério 4: 0,1

Pontuação para cada critério numa ordem:

Projeto A: 3, 2, 3, 1

Projeto B: 1, 3, 2, 1

Projeto C: 2, 2, 1, 3

Projeto D: 1, 1, 3, 3

A) Projeto A

B) Projeto C

C) Projeto D

D) Projeto B”

No caso desta pergunta, a alternativa indicada como sendo a correta no livro de Oguz é a alternativa C, que coloca o Projeto D como sendo o mais favorável. Já os chatbots, todos indicaram como correta a alternativa A, com o projeto A sendo o mais favorável.

Esta é uma questão em que é necessário empregar o modelo de pontuação ponderada, realizando o cálculo da pontuação para cada um dos projetos a fim de identificar o mais favorável.

Segundo Belton e Stewart (2001), o modelo de pontuação ponderada é uma ferramenta utilizada para avaliar diferentes opções com base no quão bem elas atendem a um conjunto de critérios importantes. Para empregá-lo, deve-se identificar os critérios, atribuindo a cada critério um peso que reflita sua importância e, em seguida, avaliar cada alternativa, atribuindo uma pontuação para cada critério de cada alternativa.

Ainda de acordo com Belton e Stewart, o cálculo da pontuação de cada alternativa é feito multiplicando cada pontuação pelo peso correspondente e somando os resultados para obter uma pontuação total para cada alternativa. A alternativa com a maior pontuação total é considerada a melhor escolha. A fórmula utilizada é: $V(a) = \sum (w_i * v_i(a))$, onde $V(a)$ é o valor total da alternativa a , w_i é o peso do critério i e $v_i(a)$ é a pontuação da opção a para o critério.

No exercício em questão, as alternativas são os 4 projetos. Já estão estabelecidos os 4 critérios, seus pesos e a pontuação de cada projeto dentro de cada critério, restando somente a parte da aplicação da fórmula. Aplicando a fórmula, temos o seguinte cálculo para cada projeto:

- **Projeto A:** $0,3 \times 3 + 0,4 \times 2 + 0,2 \times 3 + 0,1 \times 1 = 0,9 + 0,8 + 0,6 + 0,1 = 2,4$
- **Projeto B:** $0,3 \times 1 + 0,4 \times 3 + 0,2 \times 2 + 0,1 \times 1 = 0,3 + 1,2 + 0,4 + 0,1 = 2,0$
- **Projeto C:** $0,3 \times 2 + 0,4 \times 2 + 0,2 \times 1 + 0,1 \times 3 = 0,6 + 0,8 + 0,2 + 0,3 = 1,9$
- **Projeto D:** $0,3 \times 1 + 0,4 \times 1 + 0,2 \times 3 + 0,1 \times 3 = 0,3 + 0,4 + 0,6 + 0,3 = 1,6$

O projeto com a maior pontuação, ou seja, o mais favorável é o Projeto A, ou seja, a alternativa correta é a alternativa A, conforme indicado pelos chatbots. Nesta questão, o livro de Oguz, do qual foi retirado o gabarito das respostas, indica incorretamente a alternativa C como correta.

Vale ressaltar também que apesar de o ChatGPT e o Gemini terem indicado corretamente a alternativa A como a resposta correta, os cálculos apresentados por ambos apresentam erros. Como pode ser observado na Figura 1 abaixo, o ChatGPT não demonstra o cálculo completo, mostrando apenas o resultado e chega a pontuações incorretas no cálculo para os projetos C e D.

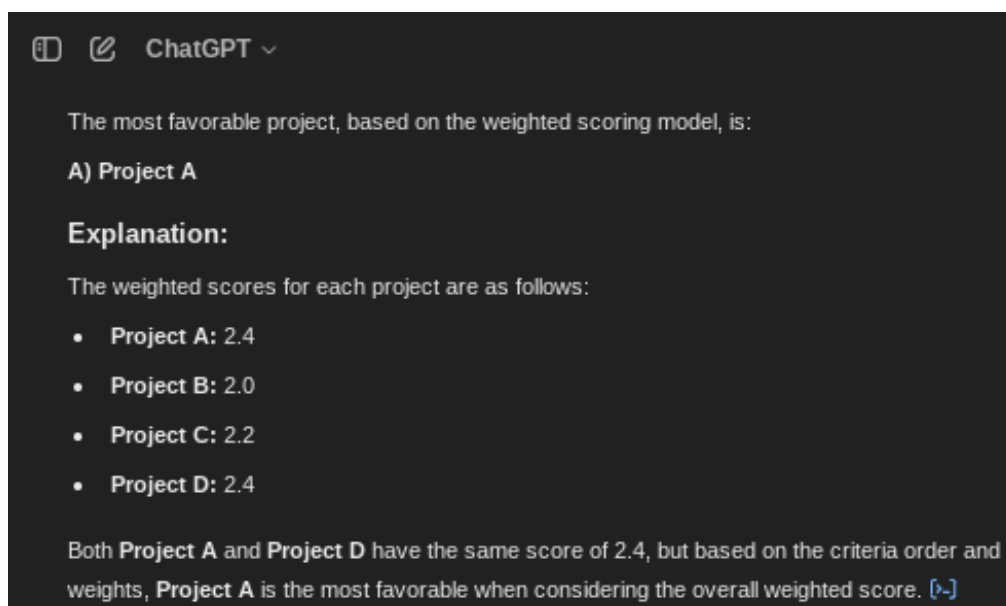


Figura 1. Resposta do ChatGPT à pergunta 2.08

Fonte: Elaborado pelos Autores (2024)

Na Figura 2 abaixo, vemos que o Gemini comete um erro de cálculo para o projeto A, obtendo o resultado de 2.2 ao invés de 2.4.

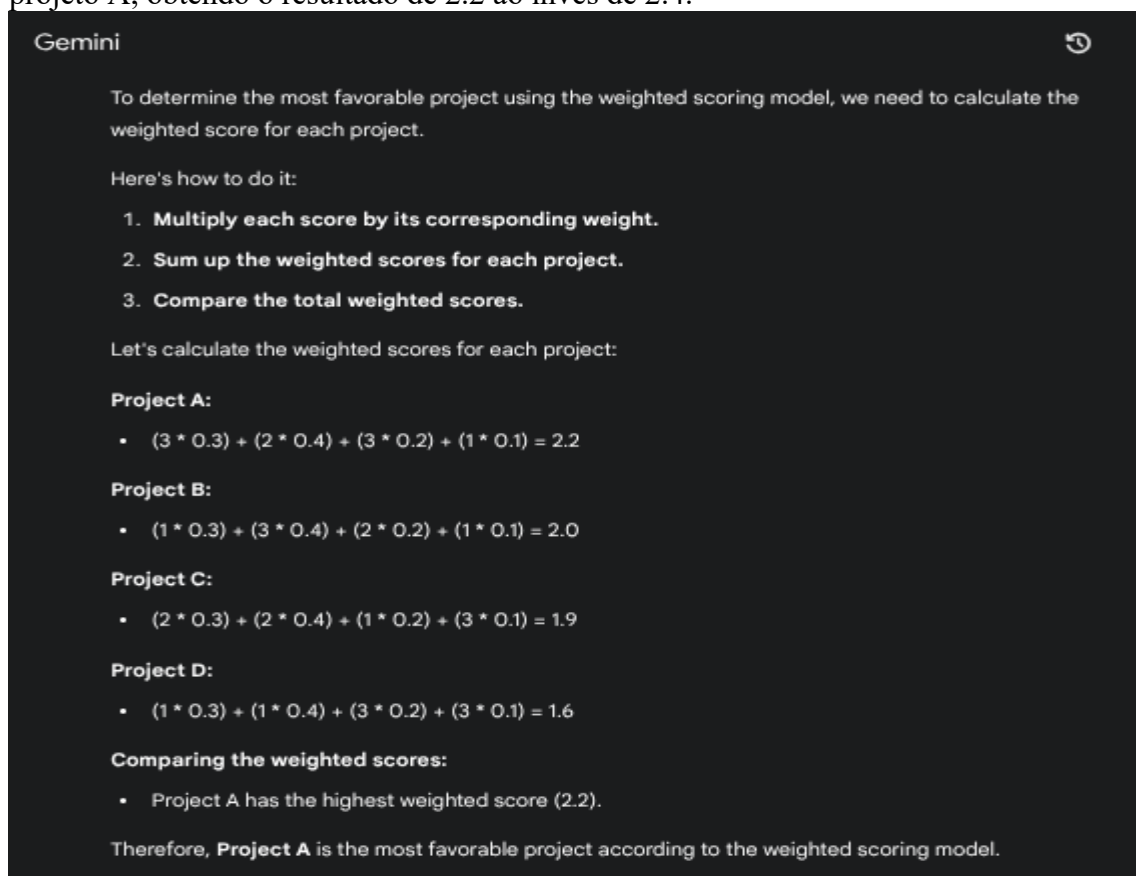
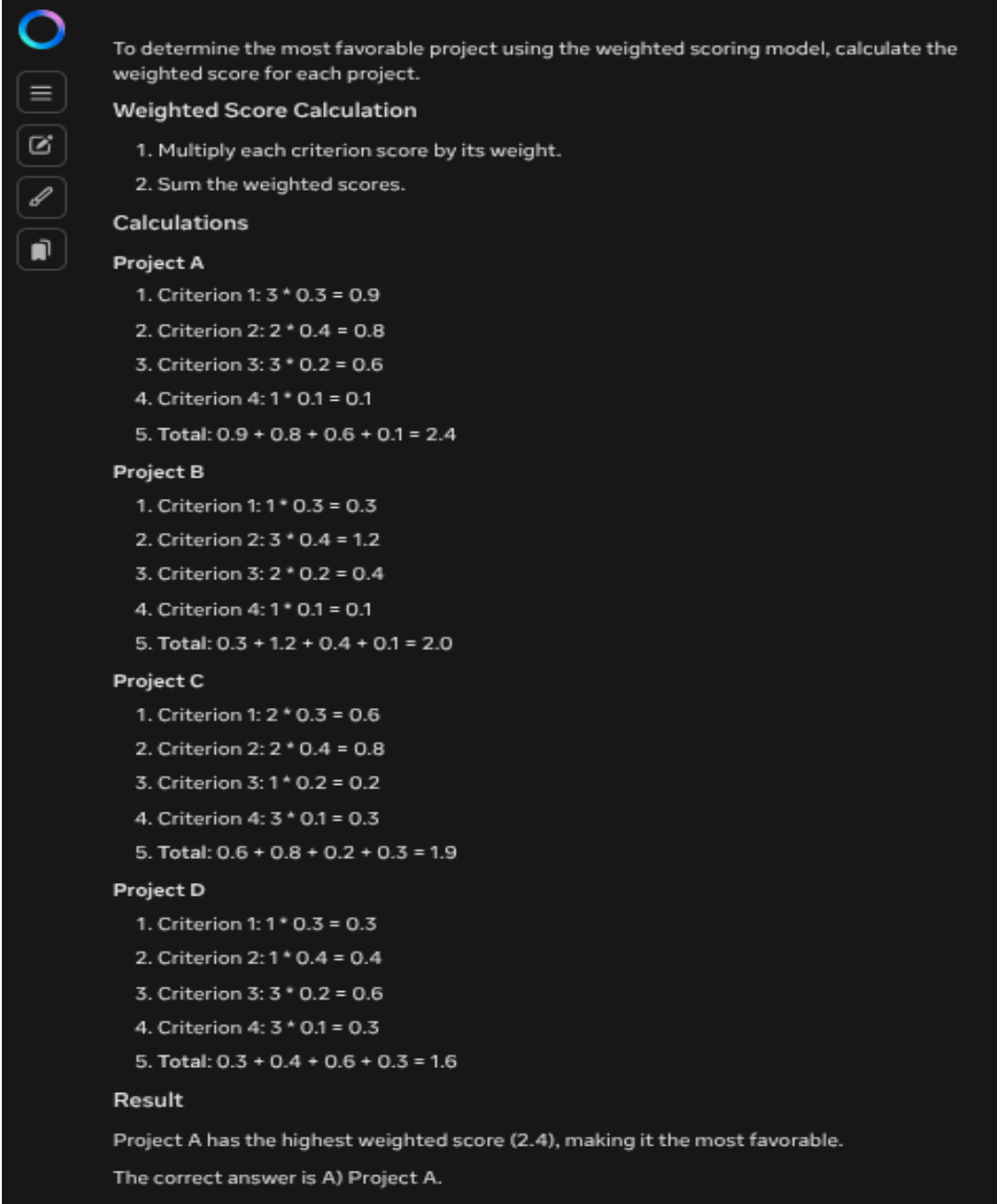


Figura 2. Resposta do Gemini à pergunta 2.08

Fonte: Elaborado pelos Autores (2024)

O Meta AI, por sua vez, é o único que apresenta sua resposta sem erros de cálculo como pode ser observado na Figura 3 abaixo.



To determine the most favorable project using the weighted scoring model, calculate the weighted score for each project.

Weighted Score Calculation

1. Multiply each criterion score by its weight.
2. Sum the weighted scores.

Calculations

Project A

1. Criterion 1: $3 * 0.3 = 0.9$
2. Criterion 2: $2 * 0.4 = 0.8$
3. Criterion 3: $3 * 0.2 = 0.6$
4. Criterion 4: $1 * 0.1 = 0.1$
5. Total: $0.9 + 0.8 + 0.6 + 0.1 = 2.4$

Project B

1. Criterion 1: $1 * 0.3 = 0.3$
2. Criterion 2: $3 * 0.4 = 1.2$
3. Criterion 3: $2 * 0.2 = 0.4$
4. Criterion 4: $1 * 0.1 = 0.1$
5. Total: $0.3 + 1.2 + 0.4 + 0.1 = 2.0$

Project C

1. Criterion 1: $2 * 0.3 = 0.6$
2. Criterion 2: $2 * 0.4 = 0.8$
3. Criterion 3: $1 * 0.2 = 0.2$
4. Criterion 4: $3 * 0.1 = 0.3$
5. Total: $0.6 + 0.8 + 0.2 + 0.3 = 1.9$

Project D

1. Criterion 1: $1 * 0.3 = 0.3$
2. Criterion 2: $1 * 0.4 = 0.4$
3. Criterion 3: $3 * 0.2 = 0.6$
4. Criterion 4: $3 * 0.1 = 0.3$
5. Total: $0.3 + 0.4 + 0.6 + 0.3 = 1.6$

Result

Project A has the highest weighted score (2.4), making it the most favorable.
The correct answer is A) Project A.

Figura 3. Resposta do Meta AI à pergunta 2.08

Fonte: Elaborado pelos Autores (2024)

Apesar de terem indicado corretamente a alternativa certa, as falhas no cálculo tanto do ChatGPT quanto do Gemini evidenciam uma dificuldade que estes modelos de linguagem têm com questões matemáticas.

Modelos de linguagem podem ter dificuldades com problemas matemáticos principalmente porque foram projetados para prever texto, e não para realizar cálculos precisos. Seu treinamento baseia-se na identificação e replicação de padrões em grandes quantidades de dados textuais, então, quando um problema exige um raciocínio lógico rigoroso, passo a passo,

ou uma aritmética exata, o modelo pode gerar uma resposta plausível que não é matematicamente correta.

A resposta incorreta no gabarito do livro-texto de Oguz evidencia uma das limitações deste estudo, já que não há garantias quanto à falta de erros e imprecisões no livro que serviu de base para o questionário.

A última pergunta em que os três LLMs divergiram do gabarito do questionário é a seguinte:

"Which stakeholder approves the project charter?

- A) Project steering committee
- B) Project manager
- C) Client
- D) Project sponsor"

Em português, temos:

"Que parte interessada aprova o termo de abertura do projeto?

- A) Comitê de direção do projeto
- B) Gerente do projeto
- C) Cliente
- D) Patrocinador do projeto"

Para esta pergunta, o livro que serve de base para o gabarito indica como sendo a resposta correta a alternativa B, ou seja, o gestor do projeto. Enquanto todos os chatbots indicam a alternativa D, patrocinador do projeto, como a resposta correta.

Nesta questão, novamente observamos uma falha no livro-texto de onde foram extraídas as perguntas e respostas para o questionário.

De acordo com o guia PMBOK publicado pelo Project Management Institute [PMI] (2021), o termo de abertura do projeto, ou project charter, se trata de um documento oficial emitido pelo iniciador ou patrocinador do projeto, quem confirma formalmente a existência do projeto e outorga ao gerente do projeto a autoridade necessária para alocar os recursos da organização em suas atividades.

Se o gerente de projeto fosse responsável por aprovar o termo de abertura do projeto, ele estaria basicamente autorizando o seu próprio trabalho, sua própria função e a alocação de recursos para o seu próprio projeto. Este seria um processo circular, causando um potencial conflito de interesse.

Observamos aqui, mais uma questão em que os chatbots respondem corretamente, indicando que a taxa de acerto dos chatbots é potencialmente maior do que a taxa de conformidade com o questionário, já que o gabarito do questionário apresentou imprecisões.

Uma análise qualitativa das respostas revelou diferenças significativas entre os modelos avaliados no que diz respeito ao nível de detalhe das respostas. O ChatGPT destacou-se ao apresentar com mais frequência respostas detalhadas, em forma de texto de fácil leitura, que explicam não só qual seria resposta correta, mas indicam também o motivo pelo qual as outras alternativas estariam incorretas. Dentre as 147 perguntas, em 124 (84.35%) o ChatGPT apresentou respostas com alto nível de detalhe.

Em contraste, o Gemini, apresentou respostas bem detalhadas, explicando o motivo não só da resposta que estaria correta, mas também explicando as respostas que julga incorretas em 39 (26.53%) respostas. Na maioria das respostas (108 ou 73.47% do total), o Gemini apresentou um nível intermediário de detalhe, limitando-se a justificar a alternativa que considera correta sem abordar o motivo pelo qual considera as outras alternativas incorretas.

O Meta AI, por sua vez, apresentou respostas detalhadas, justificando tanto a alternativa que considera correta quanto as que julga incorretas em 55 respostas (37.41%). Contudo, uma particularidade observada no Meta AI foi a tendência em responder com frases curtas,

frequentemente organizadas em tópicos e subtópicos, com apenas uma linha em cada tópico, tornando a leitura menos fluida e dificultando a compreensão em alguns casos. O Meta AI também apresentou a particularidade de citar com frequência a fonte da informação no final de cada resposta.

As respostas do Meta AI foram mais longas, com uma média de 988.6 caracteres por resposta, porém, o Meta AI gastou em média 154 caracteres de cada resposta citando as fontes da informação. Desconsiderando a citação de fontes como parte da mensagem, as respostas do Meta AI teriam ficado com uma média de 834.6 caracteres, colocando o ChatGPT em primeiro lugar no que diz respeito ao tamanho das respostas, com uma média de 970.5 caracteres por resposta. O Gemini apresentou respostas mais curtas, com média de 525 caracteres por resposta. ChatGPT e Meta AI apresentaram com mais frequência uma explicação sobre a resposta, tornando-as mais completas e longas.

Uma das limitações deste estudo é a dependência exclusiva do livro-texto utilizado para a elaboração do questionário, o qual pode apresentar erros tanto nas questões quanto nas respostas consideradas corretas. Considerando que a análise do desempenho dos chatbots foi realizada com base no gabarito extraído de um livro texto, há o risco de que eventuais imprecisões comprometam a avaliação do grau de acurácia das ferramentas de IA. Uma eventual falha no conteúdo do livro pode levar a uma interpretação distorcida dos resultados, evidenciando a necessidade de uma revisão criteriosa do material de referência ou da incorporação de fontes adicionais em pesquisas futuras.

Outra das limitações deste estudo é que a análise foi realizada apenas pelo autor, o que pode introduzir vieses e ocasionar erros interpretativos na avaliação dos resultados. É importante ressaltar também que este estudo se baseia em uma amostra inicial de perguntas e que os resultados podem variar à medida que a amostra é ampliada. Outro ponto importante de notar é que os modelos de linguagem evoluem rapidamente. As respostas para este estudo foram coletadas entre os dias 16 e 17 de novembro de 2024. Conforme os modelos de linguagem evoluem, é provável que os chatbots se tornem progressivamente mais acurados em suas respostas.

As perguntas, respostas coletadas a partir de cada chatbot e a análise dos dados pode ser encontrada no documento: “Questionário, coleta de dados e análise”, apresentado no apêndice deste trabalho. O cálculo do teste Z pode ser encontrado na segunda folha do arquivo.

5 Considerações Finais

Pesquisas futuras podem aprofundar esse estudo, explorando outros tópicos dentro da gestão de projetos e verificando como as atualizações afetam a qualidade das respostas.

Em resumo, este trabalho mostra que os chatbots têm grande potencial para ajudar profissionais de gestão de projetos, sendo que o objetivo deste estudo foi comparar a precisão dos chatbots ChatGPT, Google Gemini e Meta AI em responder perguntas sobre gestão de projetos. Com base em uma amostra de 147 perguntas de múltipla escolha, foi observado que o ChatGPT e o Meta AI alcançaram uma taxa de conformidade com o gabarito de 93,2%, enquanto o Google Gemini teve 90,48%. Apesar dessa diferença, a análise estatística com o teste-z mostrou que não há diferença significativa entre os desempenhos, indicando que os três chatbots possuem resultados parecidos no contexto avaliado.

Os achados deste trabalho mostram que todos os três chatbots podem ser úteis como ferramentas de suporte para gestores de projetos, já que conseguem responder a perguntas teóricas com alta precisão. No entanto, também foram notadas diferenças qualitativas: o ChatGPT se destaca por oferecer respostas mais detalhadas e completas, abordando tanto as alternativas que julgou corretas quanto as que considerou incorretas. O Gemini apresenta nível

intermediário de detalhamento na maioria das respostas, enquanto o Meta AI, apesar de trazer algumas explicações detalhadas, frequentemente utiliza frases curtas que dificultam a compreensão.

Outro dado importante é que 82,55% das perguntas foram respondidas conforme o gabarito do livro-texto por todos os chatbots, indicando que temas comuns em gestão de projetos são bem compreendidos por essas ferramentas. Por outro lado, um número pequeno de perguntas (2,68%) foi acertado apenas por um chatbot, revelando algumas limitações ou particularidades entre eles.

Pesquisas futuras podem aprofundar esse estudo, explorando outros tópicos dentro da gestão de projetos e verificando como as atualizações afetam a qualidade das respostas.

Este trabalho mostra que os chatbots têm grande potencial para ajudar profissionais de gestão de projetos, sendo ferramentas valiosas para otimizar tarefas baseadas em conhecimento teórico. Ao avaliar o desempenho dessas tecnologias, buscamos contribuir para o melhor entendimento de como a inteligência artificial pode ser usada de forma eficiente nas organizações.

6 Referências

- Cao, Y.; Sheng, Q.Z.; McAuley, J.; Yao, L. 2023. Reinforcement learning for generative AI: A survey. 1(1): 1–35. Disponível em: <https://arxiv.org/abs/2308.14328v2>
- Cleland, D.I.; Ireland, L.R. 2006. Project management: strategic design and implementation. 5th edition. McGraw-Hill, New York, NY, USA.
- Creswell, J.W. 2010. Projeto de pesquisa: métodos qualitativo, quantitativo e misto. 3ed. Artmed. Porto Alegre, RS, Brazil: Artmed
- Dionisio, C. S. 2022. Hybrid project management. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Field, A. 2018. Discovering statistics using IBM SPSS statistics. 6th edition. Sage Publications, Thousand Oaks, CA, USA.
- Fridgeirsson, T.V.; Ingason, H.T.; Jonasson, H.I.; Jonsdottir, H. 2021. An authoritative study on the near future effect of artificial intelligence on project management knowledge areas. Sustainability 13(4): 2345. Disponível em: <https://doi.org/10.3390/su13042345>
- Garnelo, M.; Shanahan, M. 2019. Reconciling deep learning with symbolic artificial intelligence: representing objects and relations. Current Opinion in Behavioral Sciences 29:17–23.
- Graham, J.W.; Havlick, W.C. 1994. Mission statements: A guide to the corporate and nonprofit sectors. Garland Publishing, Inc., New York & London.
- Haykin, S. 1999. Neural networks: A comprehensive foundation. Prentice Hall, Upper Saddle River, NJ, USA.
- Hess, C.; Kunz, S. 2024. The project management assistant that never sleeps: using generative artificial intelligence to support project management activities. International Journal on Advances in Software 17(1–2): 44.
- Ibbs, C. W.; Reginato, J. 2002. Measuring the strategic value of project management. Project Management - Impresario of the Construction Industry Symposium. University of California Berkeley, CA, EUA. Disponível em: https://maturityresearch.com/wp-content/uploads/2020/05/Trabalho_William_Ibbs-2.pdf
- Kaplan, R.S.; Norton, D.P. 2001. The strategy-focused organization: How balanced scorecard companies thrive in the new business environment. Harvard Business School Publishing Corporation, Boston, MA, USA.

- Karamthulla, M.J.; Tadimarri, A.; Tillu, R.; Muthusubramanian, M. 2024. Navigating the Future: AI-Driven Project Management in the Digital Era. *International Journal for Multidisciplinary Research (IJFMR)* 6(2): IJFMR240215295.
- Labadze, L.; Grigolia, M.; Machaidze, L. 2023. Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education* 20(56). Disponível em: <https://doi.org/10.1186/s41239-023-00426-1>
- Lakshminarasimham, K.C. 2024. Integrating AI into program and project management: Transforming decision-making and risk management. *International Journal of Sustainable Development Through AI, ML and IoT* 3(2).
- LeCun, Y.; Bengio, Y.; Hinton, G. 2015. Deep learning. *Nature* 521(7553): 436–444.
- Maylor, H. 2010. *Project management*. 4th edition. Pearson Education Limited, Harlow, Essex, England.
- McCorduck, P. 2004. *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. A K Peters, Ltd., Natick, MA, USA.
- Meredith, J.R.; Mantel, S.J.; Shafer, S.M. 2015. *Project management: a managerial approach*. 9th edition. Wiley, Hoboken, NJ, USA.
- Minelle, F.; Stolfi, F. 2023. AI to support PM: a ChatGPT quality assessment (β test). *PM World Journal* 7(5). Disponível em: <https://pmworldlibrary.net/wp-content/uploads/2023/05/pmwj129-May2023-Minelle-Stolfi-AI-to-support-PM-a-ChatGPT-quality-assessment-3-1.pdf>
- Moder, J. J.; Phillips, C. R.; Davis, E. W. 1983. *Project management with CPM, PERT and precedence diagramming*. 3rd edition. Van Nostrand Reinhold, New York, USA.
- Montgomery, D.C.; Runger, G.C. 2018. *Applied statistics and probability for engineers*. 7th edition. Wiley, Hoboken, NJ, USA.
- Morris, P. W. G. 1994. *The management of projects*. Thomas Telford Services Ltd, London, England.
- Oguz, A. 2022. *Project Management, Navigating the Complexity with a Systematic Approach*. MSL Academic Endeavors, Cleveland, Ohio, Estados Unidos da América. Disponível em: <https://pressbooks.ulib.csuohio.edu/project-management-navigating-the-complexity/>
- Project Management Institute [PMI]. 2021. *Padrão de gerenciamento de projetos e Guia do conhecimento em gerenciamento de projetos (Guia PMBOK)*. 7ed. Newtown Square, PA, EUA
- Rane, N.; Choudhary, S.; Rane, J. 2024. Gemini or ChatGPT? Capability, performance, and selection of cutting-edge generative artificial intelligence (AI) in business management. University of Mumbai, Mumbai, India. Disponível em: <http://dx.doi.org/10.2139/ssrn.4731281>
- Russell, S.J.; Norvig, P. 2021. *Artificial intelligence: A modern approach*. 4th edition. Pearson, Harlow, England.
- Shum, H.; He, X.; Li, D. 2018. From Eliza to XiaoIce: Challenges and opportunities with social chatbots. *Frontiers in Information Technology and Electronic Engineering* 19(1): 10–26. Disponível em: <https://doi.org/10.1631/FITEE.1700826>
- Taboada, I.; Daneshpajouh, A.; Toledo, N.; de Vass, T. 2023. Artificial Intelligence Enabled Project Management: A Systematic Literature Review. *Applied Sciences (Switzerland)* 13(8): 5014. Disponível em: <https://doi.org/10.3390/app130850143>.
- Triola, M.F. 2019. *Essentials of statistics*. 6th edition. Pearson, New York, NY, USA.
- Trochim, W.M.K. 1989. Outcome Pattern Matching and Program Theory. *Evaluation and Program Planning*, 12(4), 355-366. Disponível em: [https://doi.org/10.1016/0149-7189\(89\)90052-9](https://doi.org/10.1016/0149-7189(89)90052-9)
- Turner, J.R. 2009. *The handbook of project-based management*. 3rd edition. McGraw-Hill, London, England.