



Universidade do Estado do Rio de Janeiro

Centro de Ciências Sociais

Faculdade de Administração e Finanças

Ana Clara de Melo Reis

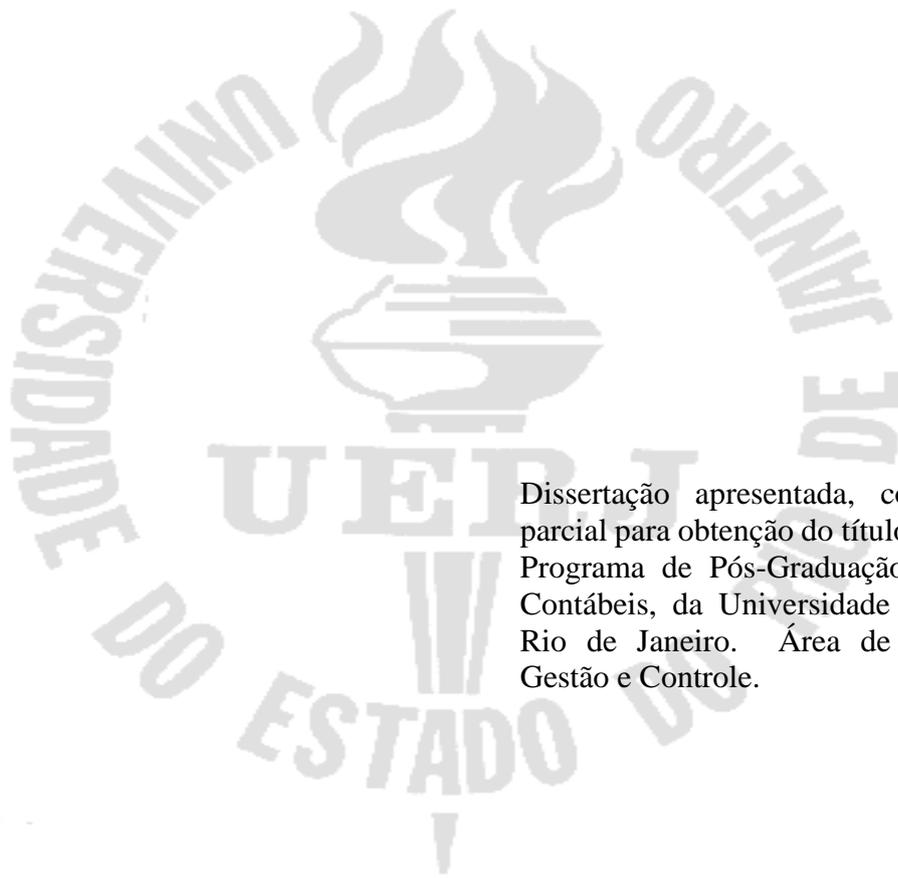
**A Inteligência Artificial é capaz de obter aprovação nos exames de
suficiência e qualificação técnica do CFC?**

Rio de Janeiro

2023

Ana Clara de Melo Reis

A Inteligência Artificial é capaz de obter aprovação nos exames de suficiência e qualificação técnica do CFC?



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Ciências Contábeis, da Universidade do Estado do Rio de Janeiro. Área de concentração: Gestão e Controle.

Orientador: Prof. Dr. José Francisco Moreira Pessanha

Coorientador: Prof. Dr. Ricardo Lopes Cardoso

Rio de Janeiro

2023

CATALOGAÇÃO NA FONTE
UERJ / REDE SIRIUS / BIBLIOTECA CCS/B

R375 Reis, Ana Clara de Melo.

A Inteligência Artificial é capaz de obter aprovação nos exames de suficiência e qualificação técnica do CFC? / Ana Clara de Melo Reis. – 2023. 67 f.

Orientador: Prof. Dr. José Francisco Moreira Pessanha.

Coorientador: Prof. Dr. Ricardo Lopes Cardoso.

Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro, Faculdade de Administração e Finanças.

Bibliografia: f. 50-51.

1. Inteligência artificial – Teses. 2. ChatGPT – Teses. 3. Conselho Federal de Contabilidade (Brasil) – Teses. 4. Qualificação técnica – Teses. I. Pessanha, José Francisco Moreira. II. Cardoso, Ricardo Lopes. III. Universidade do Estado do Rio de Janeiro. Faculdade de Administração e Finanças. IV. Título. CDU 657: 004.89

Bibliotecário: Fabiano Salgueiro CRB7/6974

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Ana Clara de Melo Reis

A Inteligência Artificial é capaz de obter aprovação nos exames de suficiência e qualificação técnica do CFC?

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Ciências Contábeis da Faculdade de Administração e Finanças, da Universidade do Estado do Rio de Janeiro. Área de concentração: Gestão e Controle.

Aprovada em 14 de dezembro de 2023.

Banca Examinadora:

Prof. Dr. José Francisco Moreira Pessanha (Orientador)
Faculdade de Administração e Finanças - UERJ

Prof. Dr. Ricardo Lopes Cardoso (Coorientador)
Faculdade de Administração e Finanças - UERJ

Prof. Dr. Carlos Eduardo Mello
Faculdade de Administração e Finanças - UERJ

Prof. Dr. Natan Szuster
Universidade Federal do Rio de Janeiro

Rio de Janeiro

2023

Prof. Dr. Felipe Buchbinder
FGV-EBAPE

Prof. Dr. Luiz Otavio Schmall dos Santos
Petrobras

Rio de Janeiro

2023

DEDICATÓRIA

Dedico este trabalho de mestrado ao meu filho, Leonardo Reis Padrone, motivador diário das minhas conquistas e avanços.

AGRADECIMENTOS

À Faculdade de Administração e Finanças da Universidade Estadual de Rio de Janeiro, seus docentes e funcionários, que desde 2022 quando ingressei no Curso de Mestrado em Ciências Contábeis, me acompanham neste meu percurso acadêmico e que contribuíram para conquista dessa titulação.

Ao meu Filho, pelo amor, carinho e motivação diária, me fez acreditar que era possível.

Agradeço aos meus pais, obrigada por se manterem ao meu lado o tempo todo.

Presto agradecimento ao Professor Ricardo Lopes Cardoso e ao Professor José Francisco Moreira Pessanha, pelo aprendizado e apoio na elaboração desta dissertação.

Aos meus colegas de curso pela troca e companhia ao longo do mestrado.

RESUMO

REIS, Ana Clara de Melo. *A Inteligência Artificial é capaz de obter aprovação nos exames de suficiência e qualificação técnica do CFC?*. 2023. 67 f. Dissertação (Mestrado em Ciências Contábeis) - Faculdade de Administração e Finanças, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

A relação entre Inteligência Artificial (IA) e contabilidade tem se tornado cada vez mais relevante para os estudos acadêmicos recentes e para sociedade em geral levando a reflexões sobre o impacto no mercado de trabalho, a segurança dos dados, transparência e a ética na sua utilização. O presente estudo tem por objetivo analisar o desempenho dos modelos ChatGPT, GPT-4 e Bard nos exames de suficiência do CFC (Conselho Federal de Contabilidade) e nos exames de qualificação técnica como QTG (Qualificação Técnica Geral), CVM (Comissão de Valores Mobiliários), BCB (Banco Central do Brasil), SUSEP (Superintendência de Seguros Privados), PREVIC (Superintendência Nacional de Previdência Complementar) e Perícia Contábil. Foi adotada uma abordagem metodológica quantitativa para examinar o desempenho das inteligências artificiais (IA) nos exames de suficiência e qualificação técnica ministrados pelo CFC em 2022. O modelo GPT-4 demonstrou um desempenho superior em relação a outros modelos, tendo alcançado aprovação em 100% dos exames de suficiência e em 83% nos exames técnicos de 2022, um resultado confirmado pelo teste exato de Fisher. Por meio de testes qui-quadrado de homogeneidade, aplicados às frequências de acertos e erros nos exames selecionados, foi possível avaliar a hipótese nula de desempenho consistente de um modelo de IA ao longo de disciplinas pertencentes a diversas áreas de conhecimento. No que tange aos exames de suficiência, os resultados obtidos indicam a falta de evidências significativas para rejeitar a hipótese nula nos casos do ChatGPT e GPT-4. Em contrapartida, no Bard a hipótese nula foi rejeitada, indicando um desempenho mais variável entre os grupos de disciplinas. Nos exames técnicos, os resultados refutam a hipótese nula, demonstrando que os modelos de IA exibiram variações de desempenho nas seis disciplinas analisadas. A presença de tabelas nas questões tende a prejudicar o desempenho dos modelos de IA. Porém, somente no GPT-4 observou-se um desempenho significativamente superior nas questões sem tabela. Ademais o índice de Tanimoto revelou uma afinidade moderada entre as respostas dos 3 modelos de IA. Por último, investigou-se a capacidade dos modelos responderem para questões discursivas das provas técnicas QGT e CVM e notou-se que os modelos ChatGPT e GPT-4 conseguiram desempenho suficientes nas questões discursivas para aprovação nas duas provas, já o Bard conseguiria ser aprovado apenas no exame QTG. Aderência à área de concentração (Gestão de Controle) e linha de pesquisa do PPGCC (Controladoria em Entidades Públicas e Privadas): Ao avaliar o desempenho de modelos de Inteligência Artificial (IA) generativa na obtenção de certificações profissionais na área contábil, esta dissertação impulsiona o desenvolvimento de novos conhecimentos e práticas no campo da contabilidade. A integração da IA na controladoria impulsiona avanços no conhecimento e nas práticas de controle de gestão, abrindo um amplo leque de possibilidades para inovação e maior eficiência na análise de dados financeiros, gestão de riscos e otimização de processos contábeis, tanto em âmbito nacional quanto internacional, e aplicáveis tanto em entidade públicas quanto privadas. A pesquisa também instiga reflexões sobre a revisão crítica das necessidades formativas dos estudantes de contabilidade, considerando o impacto da IA nas certificações profissionais. Abordando os desafios contemporâneos e emergentes na interseção entre IA e contabilidade, o estudo oferece insights valiosos para o futuro da profissão, redefinindo as competências essenciais para contabilistas na era digital. Como resultado, esta dissertação está alinhada ao objetivo do PPGCC de

promover pesquisas sistemáticas sobre problemas relevantes em Ciências Contábeis, com ênfase no fomento e criação de instrumentos inovadores, além de contribuir para a formação de quadros docentes em contabilidade.

Palavras-chave: Exame de Suficiência. Avaliação Conhecimento. Inteligência Artificial. ChatGPT. GPT-4. Bard.

ABSTRACT

REIS, Ana Clara de Melo. *Can Artificial Intelligence pass the CFC's proficiency and technical qualification exams?*. 2023. 67 f. Dissertação (Mestrado em Ciências Contábeis) - Faculdade de Administração e Finanças, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2023.

The relationship between Artificial Intelligence (AI) and accounting has become increasingly relevant in recent academic studies and in society as a whole, leading to reflections on its impact on the labor market, data security, transparency, and ethics in its use. This present study aims to analyze the performance of the ChatGPT, GPT-4, and Bard models in the exams conducted by the CFC (Federal Accounting Council) and in technical qualification exams such as QTG (General Technical Qualification), CVM (Securities and Exchange Commission), BCB (Central Bank of Brazil), SUSEP (Superintendence of Private Insurance), PREVIC (National Superintendence of Supplementary Pensions), and Forensic Accounting. We adopted a quantitative methodological approach to examine the performance of artificial intelligence (AI) in the proficiency and technical qualification exams administered by CFC in 2022. The GPT-4 model demonstrated remarkably superior performance compared to other models, achieving a 100% approval rate in proficiency exams and an 83% approval rate in technical exams in 2022, a result confirmed by the exact Fisher's test. Through homogeneity chi-square tests, applied to the frequencies of correct and incorrect responses in the selected exams, it was possible to evaluate the null hypothesis of consistent performance of an AI model across disciplines belonging to various knowledge areas. Regarding proficiency exams, the obtained results indicate a lack of significant evidence to reject the null hypothesis in the cases of ChatGPT and GPT-4. On the other hand, in the Bard model, the null hypothesis was rejected, indicating a more variable performance across disciplines. In technical exams, the results refute the null hypothesis, demonstrating that the AI models exhibited performance variations in the six analyzed disciplines. A presence of tables in the questions tends to impair the performance of AI models. However, only in GPT-4, a significantly superior performance was observed in questions without tables. Furthermore, the Tanimoto index revealed a moderate affinity among the responses of the three AI models. Finally, we investigated the models' ability to respond to open-ended questions in the QGT and CVM technical exams. It was observed that the ChatGPT and GPT-4 models demonstrated sufficient performance in the open-ended questions for approval in both exams, while the Bard model would only be approved in the QTG exam.

Alignment with the Concentration Area (Control Management) and Research Line of the PPGCC (Controllorship in Public and Private Entities): This dissertation advances the development of new knowledge and practices in the field of accounting by evaluating the performance of generative Artificial Intelligence (AI) models in acquiring professional certifications in the accounting sector. The integration of AI into controllorship fosters advancements in knowledge and management control practices, opening a broad spectrum of opportunities for innovation and enhanced efficiency in financial data analysis, risk management, and the optimization of accounting processes, both nationally and internationally, applicable to both public and private entities. Furthermore, this research stimulates critical reflections on the formative needs of accounting students, considering the impact of AI on professional certifications. By addressing contemporary and emerging challenges at the intersection of AI and accounting, the study offers valuable insights for the future of the profession, redefining the essential competencies for accountants in the digital age. Consequently, this dissertation aligns with the PPGCC's objective of promoting systematic

research on pertinent issues in Accounting Sciences, emphasizing the encouragement and creation of innovative tools, and contributing to the development of academic faculty in accounting.

Keywords: Proficiency Exam. Knowledge Assessment. Artificial Intelligence. ChatGPT. GPT-4. Bard.

LISTA DE ILUSTRAÇÕES

Tabela 1-	Resumo Resultado Exames de Suficiência 2022.....	29
Tabela 2-	Média nacional de humanos por região/estado e nota da IA, 1º exame suficiência 2022.....	30
Tabela 3-	Média nacional de humanos por região/estado e nota da IA, 2º exame suficiência 2022.....	31
Tabela 4-	Desempenho humano e IA por disciplina, 1º exame de suficiência 2022.....	32
Tabela 5-	Desempenho humano e IA por disciplina, 2º exame de suficiência 2022.....	33
Tabela 6-	Desempenho da IA no 1º exame suficiência CFC 2022, por disciplina e enunciados com ou sem Tabela.....	34
Tabela 7-	Desempenho da IA no 2º exame suficiência CFC 2022, por disciplina e enunciados com ou sem Tabela.....	35
Tabela 8-	Desempenho da IA no CFC 2022 com ou sem tabela.....	36
Tabela 9-	Índice de Tanimoto.....	36
Tabela 10-	Desempenho IA, exames de qualificação técnica 2022.....	37
Tabela 11-	Resultado consolidado nas 150 questões dos Exames Técnicos 2022.	38
Tabela 12-	Índice de Tanimoto.....	38
Tabela 13-	Desempenho IA por disciplina após agrupamento, 1º exame de suficiência 2022.....	39
Tabela 14-	Desempenho IA por disciplina após agrupamento, 2º exame de suficiência 2022.....	40
Tabela 15-	Desempenho IA por disciplina após agrupamento, exames de qualificação técnica 2022.....	41
Tabela 16-	Total de acertos e erros por disciplina nos exames de suficiência de 2022.....	42
Tabela 17-	Total de acertos e erros por disciplina nos exames técnicos de 2022.....	43
Tabela 18-	Questões discursivas, QGT e CVM 2022.....	44

LISTA DE ABREVIATURAS E SIGLAS

BEC	Business Environment and Concepts
BCB	Banco Central do Brasil
CFC	Conselho Federal de Contabilidade
CIA	<i>Certified Internal Auditor</i>
CMA	<i>Certified Management Accountant</i>
CPA	<i>Certified Public Accountant</i>
CRC	Conselho Regional de Contabilidade
CVM	Comissão de Valores Mobiliários
EA	<i>Enrolled Agent</i>
EQT	Exame de Qualificação Técnica
FAR	<i>Financial Accounting and Reporting</i>
IA	Inteligência Artificial
NSFN	Normas do Sistema Financeiro Nacional
PREVIC	Superintendência Nacional de Previdência Complementar
QTG	Qualificação Técnica Geral
SUSEP	Superintendência de Seguros Privados

SUMÁRIO

	INTRODUÇÃO.....	12
1	REVISÃO DE LITERATURA.....	16
1.1	Estudos sobre a certificação contábil.....	16
1.2	Estudos sobre aprovação e influência da IA em exames.....	19
2	CONTEXTO ANALISADO: EXAME DE SUFICIÊNCIA E EXAMES TÉCNICOS.....	22
3	ABORDAGEM METODOLÓGICA.....	25
3.1	Amostra e coleta de dados.....	26
3.2	Query para identificar o desempenho da IA.....	26
3.3	Metodologia de análise de dados.....	27
4	RESULTADOS.....	29
4.1	Análise dos Exames de Suficiência de 2022.....	29
4.2	Análises dos Exames de Qualificação Técnica 2022.....	37
4.3	Análise: agrupamento de disciplinas e distribuição de acerto alcançados em cada modelo de Inteligência Artificial.....	38
4.4	Análise: Questões discursivas QTG e CVM.....	43
	CONCLUSÕES.....	45
	REFERÊNCIAS.....	50
	APÊNDICE A – Links ChatGPT, GPT-4 e Bard: Exames Suficiência e Exames Técnicos 2022.....	52
	APÊNDICE B – Solicitação Abertura de dados Exames CFC 2022.....	54
	APÊNDICE C – Resposta da Solicitação de Abertura de dados Exames CFC 2022.....	55
	APÊNDICE D – Análise questões discursivas exames técnicos QGT e CVM 2022.....	56

INTRODUÇÃO

A relação entre Inteligência Artificial (IA) e contabilidade tem se tornado cada vez mais relevante para os estudos acadêmicos recentes e para sociedade em geral suscitando reflexões sobre o impacto no mercado de trabalho, a segurança dos dados, transparência e a ética na sua utilização. Com os recentes avanços em inteligência artificial e processamento de linguagem natural, geradores de texto de inteligência artificial (AITG – *Artificial Intelligence Text Generators*), como o ChatGPT (*Generative Pre-trained Transformer*), adquiriram a capacidade de produzir respostas sofisticadas para uma ampla variedade de consultas (Chen, 2023).

O ChatGPT é uma inteligência artificial desenvolvida pela OpenAI que interage de forma conversacional. Seu formato de diálogo possibilita que o ChatGPT responda as perguntas. Seu principal concorrente é o Bard, desenvolvido pela Google, recentemente lançado em fase experimental com o modelo PaLM 2 do Google em uma nova janela para a colaboração entre usuários e a IA generativa.

A aplicação da IA na área contábil abre oportunidades para aprimorar os processos, aumentar a qualidade das informações financeiras e fornecer *insights* valiosos para as organizações; contudo, essas vantagens são acompanhadas por alguns desafios e preocupações. Neste contexto, Eulerich *et al.* (2023) investigaram as possíveis implicações dos grandes modelos de linguagem (LLMs), como os GPTs, no mercado de trabalho dos Estados Unidos. Como resultado, identificaram que cerca de 80% da força de trabalho dos Estados Unidos poderia ter pelo menos 10% de suas tarefas afetadas pela introdução dos LLMs, enquanto aproximadamente 19% dos trabalhadores podem ter pelo menos metade de suas tarefas impactadas.

Outros estudos promoveram reflexões sobre a utilização dessa nova tecnologia no ensino. Chen *et al.* (2023) avaliaram as capacidades dos modelos ChatGPT e GPT-4 para fornecer soluções para oito casos de ensino em contabilidade. Como resultado, observaram que a capacidade do ChatGPT de fornecer soluções precisas varia dependendo do tipo de requisito do caso e da área temática da contabilidade, com melhor desempenho em tarefas que requerem elementos que exigem explicação, aplicação de regras e avaliação ética.

Eulerich *et al.* (2023) identificaram que o ChatGPT é capaz de obter aprovação nos principais exames de certificação contábil, especificamente os seguintes exames profissionais

que são mais populares nos Estados Unidos: *Certified Public Accountant* (CPA), *Certified Management Accountant* (CMA), *Certified Internal Auditor* (CIA) e *Enrolled Agent* (EA).

Similar à pesquisa de Eulerich *et al.* (2023), o presente estudo tem por objetivo analisar o desempenho dos modelos ChatGPT, GPT-4 e Bard nos exames de suficiência do Conselho Federal de Contabilidade (CFC) e nos exames de qualificação técnica como Qualificação Técnica Geral (QTG), Comissão de Valores Mobiliários (CVM), Banco Central do Brasil (BCB), Superintendência de Seguros Privados (SUSEP), Superintendência Nacional de Previdência Complementar (PREVIC) e Perícia Contábil.

O CFC é responsável pela aplicação dos dois exames aos profissionais da área contábil: o Exame de Suficiência e as provas do Exame de Qualificação Técnica (EQT). O Exame de Suficiência é executado pela Consultoria e Planejamento em Administração (Consulplan), sob sua inteira responsabilidade e controle. O Exame de Suficiência tem por objetivo comprovar conhecimentos médios, consoante os conteúdos programáticos desenvolvidos no curso de Bacharelado em Ciências Contábeis, com objetivo de promover a obtenção de registro na categoria Contador para os aprovados. Vale destacar que esse exame pode ser prestado exclusivamente pelos bacharéis e estudantes do último ano letivo do curso de Ciências Contábeis.

Conforme orientações do CFC, as provas do EQT avaliam o conhecimento e a competência técnico-profissional dos contadores que desejam atuar como auditores independentes em organizações reguladas pela CVM, BCB, SUSEP e PREVIC.

Além disso, o exame também é relevante para obter registro no Cadastro Nacional dos Auditores Independentes (CNAI) do CFC ou ingressar na área de perícia contábil. Os exames técnicos são destinados aos contadores devidamente registrados no Conselho Regional de Contabilidade (CRC) de sua jurisdição, que já atuam ou têm interesse em atuar como auditores independentes ou peritos contábeis, seja como funcionários, sócios de empresas de auditoria ou profissionais autônomos. Essas provas visam garantir que os profissionais da contabilidade que atuam nessas áreas tenham o conhecimento técnico necessário e estejam alinhados com as normas e legislações específicas de cada setor.

A questão central deste estudo é investigar se a tecnologia atual de inteligência artificial é capaz de obter sucesso nesses importantes exames profissionais. Compreender o desempenho dessas tecnologias nos exames de suficiência e qualificação técnica é crucial para avaliar seu potencial e impactos na formação e contratação de profissionais contábeis, bem como refletir

sobre estrutura e qualidade dos exames de suficiência e qualificação técnica atualmente aplicados.

Questões de pesquisa:

O presente estudo busca responder as seguintes questões:

1. A Inteligência Artificial (ChatGPT, GPT-4 e Bard) é capaz de obter aprovação no exame de suficiência do CFC e nos exames de qualificação técnica (QTG, CVM, BCB, SUSEP, PREVIC e Perícia Contábil)?
2. Em caso afirmativo, a Inteligência Artificial (ChatGPT, GPT-4 e Bard) apresenta desempenho superior à média dos humanos?
3. As características das questões podem impactar a taxa de acerto da Inteligência Artificial?
4. Após o agrupamento de disciplinas em temas específicos realizados nos exames de suficiência e técnicos de 2022, os três modelos de Inteligência Artificial (ChatGPT, GPT-4 e Bard) apresentaram distribuições de acertos e erros semelhantes ao longo dos grupos de disciplinas, i.e., um desempenho consistente?

Justificativa

Essa investigação é relevante para entender se os avanços recentes na inteligência artificial, representados pelos modelos ChatGPT, GPT-4 e Bard, são capazes de se equiparar às habilidades e conhecimentos exigidos pelo CFC nessas certificações profissionais.

Compreender o desempenho dessas tecnologias nos exames de suficiência e qualificação técnica é crucial para avaliar seu potencial de aplicação prática na área contábil e seus possíveis impactos na formação e contratação de profissionais.

Ao examinar o desempenho dos modelos ChatGPT, GPT-4 e Bard nas provas de suficiência e qualificação técnica, pode-se obter *insights* sobre a capacidade da inteligência

artificial em realizar tarefas específicas relacionadas à contabilidade. Essa análise ajuda a compreender o progresso alcançado e as limitações dessas tecnologias em campos altamente especializados.

Metodologia

Para responder às questões propostas, foi utilizada a metodologia quantitativa para coletar, analisar e interpretar dados resultantes da aplicação das provas dos exames de suficiência do CFC e exames de qualificação técnica (QTG, CVM, BCB, SUSEP, PREVIC e Perícia Contábil) de 2022 nos modelos de IA (ChatGPT, GPT-4 e Bard), em seguida, foi comparado o desempenho da IA com o desempenho de humanos que efetivamente realizaram essas provas em 2022.

Para o exame de suficiência, observaram-se os dados estatísticos dos candidatos que realizaram as duas edições de 2022 por meio do site do CFC. Utilizaram-se os critérios de aprovação definidos pelo CFC, ou seja, mínimo 50% de acerto no exame de suficiência, e mínimo 60% de acerto na prova objetiva para os exames de qualificação técnica, controlando-se pelas características das questões a partir da área temática, conforme taxonomia estabelecida no respectivo edital, a fim de avaliar o desempenho por disciplina versus IA.

A investigação sobre eventuais diferenças significativas entre as pontuações baseia-se no teste exato de Fisher (Upton e Cook, 2008; Toyama *et al.*, 2023) e no teste qui-quadrado de homogeneidade (Verzani, 2004; Taloni *et al.*, 2023) realizado com o *software* estatístico R (R Core Team, 2022). O grau de similaridade entre as respostas dos modelos de IA foi avaliado pelo índice de Tanimoto (O’Leary, 2023).

Por último, realizou-se uma análise qualitativa das questões discursivas das provas técnicas QGT e CVM nos modelos de IA para compreender a capacidade de desenvolvimento em questões discursivas, observando os conhecimentos técnicos e habilidade textual.

1. REVISÃO DE LITERATURA

1.1 Estudos sobre a certificação contábil

Este tópico aborda estudos que investigaram a percepção da sociedade acadêmica em relação aos principais exames de certificação contábil, como o Exame de Contador Público Certificado (CPA) nos Estados Unidos e o Exame de Suficiência no Brasil.

O Exame de Suficiência tem por objetivo verificar o domínio dos conhecimentos médios adquiridos durante o curso de Bacharelado em Ciências Contábeis, permitindo que os aprovados obtenham o registro na categoria de Contador. No Brasil, esse exame é exclusivo para bacharéis e estudantes do último ano do curso de Ciências Contábeis.

Por outro lado, as provas do Exame de Qualificação Técnica seguem as diretrizes do CFC e avaliam o conhecimento e a competência técnico-profissional dos contadores que desejam atuar como auditores independentes em organizações reguladas pela CVM, BCB, SUSEP e PREVIC.

Silva et al. (2020) investigaram a percepção dos estudantes em relação ao Exame de Suficiência e encontraram que o exame é considerado tanto importante quanto necessário para a profissão contábil. Eles acreditam que o exame desempenha um papel fundamental na garantia da qualidade dos cursos de contabilidade e na valorização dos profissionais da área. Para a realização dessa pesquisa, a equipe utilizou uma análise fatorial para criar construtos a partir das percepções dos alunos. Como resultado desse processo, foram identificados dois construtos principais: (i) valorização profissional que obteve uma pontuação significativa de 88,15%, refletindo a crença dos alunos de que o exame contribui para elevar o prestígio da profissão contábil; (ii) ambiente acadêmico com uma pontuação de 6,49%, indica que os alunos percebem o exame como um elemento que influencia o ambiente de aprendizado e ensino no contexto acadêmico.

Esses achados sugerem a importância do Exame de Suficiência na perspectiva dos estudantes e destacam seu papel tanto na valorização profissional quanto na dinâmica do ambiente acadêmico.

Miranda et al. (2017) analisaram a percepção de docentes e profissionais em relação ao Exame de Suficiência na profissão contábil, utilizando uma abordagem descritiva para analisar dados coletados *online* junto a 1.185 autores de artigos apresentados em congressos e

publicados em periódicos da área contábil. A amostra da pesquisa abrangeu diversos perfis: 41,5% dos participantes eram pesquisadores que desempenham o papel de docentes, 29,1% atuavam exclusivamente como profissionais do mercado e 29,4% desempenhavam papéis em ambas as áreas. Os resultados revelaram taxa significativa de concordância, com 90% dos participantes apoiando a exigência do Exame de Suficiência. Entretanto, 57,7% expressaram a opinião de que o exame deveria incluir revalidações periódicas. Além disso, 84,6% dos participantes acreditam que, além do exame, os profissionais da área contábil deveriam comprovar atualização continuada em seus conhecimentos. Pouco mais da metade (53%) dos entrevistados concordaram que o modelo de aplicação do exame é adequado. Em geral, as questões do exame foram avaliadas como medianas, mas cerca de 70% dos docentes relataram utilizá-las em maior ou menor grau em suas atividades didáticas. Eles também destacaram a importância de bons resultados nesses exames para a imagem das Instituições de Ensino Superior. Esses resultados fornecem uma visão abrangente da percepção dos docentes e profissionais sobre o Exame de Suficiência na profissão contábil e indicam áreas de potencial melhoria no exame e em suas práticas associadas.

Bunker e Cagle (2021) recomendam a revisão do Exame CPA da Becker, que insiste fortemente que os candidatos considerem iniciar sua jornada no exame CPA com a seção *Financial Accounting and Reporting* (FAR). A pesquisa em questão reuniu evidências relacionadas às taxas de aprovação nas diferentes seções do exame, comparando candidatos que o realizaram no período de até um ano após a graduação com aqueles que optaram por adiar o exame por um ano ou mais. Apesar das recomendações, a maioria dos candidatos opta por começar pela seção *Business Environment and Concepts* (BEC).

Vale ressaltar que as taxas de aprovação no BEC se mantêm consistentemente como as mais elevadas entre as quatro partes do exame, independentemente se os alunos realizarem o exame imediatamente após a graduação ou após um período de espera. Observou-se no estudo que as taxas de aprovação para a seção FAR diminuem da segunda posição mais alta entre as quatro seções para a posição mais baixa quando são comparados os candidatos que realizam o exame no prazo de um ano após a graduação com aqueles que aguardam ao menos um ano.

Segundo Bunker e Cagle (2021), os resultados observados oferecem evidências adicionais que sustentam a sugestão de que candidatos que desejam maximizar suas chances de sucesso devem considerar não apenas iniciar pelo exame da seção FAR, mas também fazê-lo o mais rapidamente possível após a graduação. Para os autores, é fundamental que educadores de contabilidade colaborem com os alunos na elaboração de um plano estratégico para concluir o

Exame CPA, maximizando suas perspectivas de êxito. Segundo os autores, encorajar os estudantes a realizar o Exame CPA no prazo de um ano após a graduação e a começar pela seção FAR parece ser a abordagem mais eficaz para que obtenham êxito o mais rapidamente possível nessa empreitada.

Cordis e Steven (2020) investigaram a relação entre os gastos per capita com ensino superior por parte dos governos estaduais e locais e o desempenho dos candidatos no Exame de CPA. Os autores destacaram a relevância dos governos estaduais e locais no fornecimento de acesso ao ensino superior em instituições públicas, dado que existe uma considerável variabilidade nos gastos per capita com ensino superior entre os estados, o que consequentemente resulta em disparidades na qualidade da educação. Os resultados revelaram que o nível de gastos per capita com ensino superior tem um impacto marginal positivo nas taxas de aprovação no Exame CPA para universidades públicas. Em contraste, esse mesmo indicador não demonstrou influência nas taxas de aprovação no Exame CPA para universidades privadas. Assim, a pesquisa destaca a importância do financiamento público no ensino superior e como ele pode afetar o desempenho dos candidatos no Exame CPA, particularmente nas universidades públicas.

Espahodi *et al.* (2023) destacaram que a certificação de CPA representa uma medida crucial de realização profissional e desempenha um papel essencial no progresso da carreira em contabilidade. Este exame é conhecido por sua rigidez, com taxas de aprovação geralmente inferiores em comparação com outros exames profissionais, sendo ainda mais baixas para grupos historicamente desfavorecidos.

Ademais, Espahodi *et al.* (2023) examinaram a influência de fatores socioambientais, como o status socioeconômico e as disparidades de educação e renda, juntamente com outros fatores que impactam o desempenho no exame CPA. Os resultados indicam que a oportunidade desempenha um papel significativo no desempenho no exame, com efeitos diferenciados para os grupos historicamente desfavorecidos; candidatos caucasianos e do sexo masculino parecem usufruir de mais oportunidades em comparação com grupos historicamente desfavorecidos. Além disso, constatou-se que todos os candidatos com melhores oportunidades, como aqueles que frequentaram universidades públicas, instituições mais seletivas, aquelas acreditadas pela AACSB e aquelas que têm recrutamento das "Big Four", tendem a apresentar um desempenho superior em relação aos seus pares. Desta forma, os resultados da análise reforçam a importância da oportunidade, evidenciando que ela desempenha um papel diferenciado no desempenho dos candidatos historicamente desfavorecidos.

Diante dessas constatações, Espahbodi *et al.* (2023) enfatizam a necessidade das partes interessadas no CPA reconhecerem o impacto dos fatores socioeconômicos no desenvolvimento, recrutamento e retenção de talentos diversos. Além disso, destacam a importância de abordar as lacunas de oportunidade desde as fases iniciais da educação, do ensino fundamental até a graduação.

Observou-se nos estudos apresentados anteriormente, a percepção da sociedade acadêmica em relação aos principais exames de certificação contábil é influenciada por uma série de fatores. Embora esses exames sejam amplamente reconhecidos como importantes para a profissão contábil, há uma crescente ênfase na necessidade de abordar questões relacionadas à igualdade de oportunidades, acessibilidade e eficácia na preparação para esses exames, a fim de promover uma profissão contábil mais diversificada e inclusiva.

1.2 Estudos acerca da aprovação e influência das IAs em exames

O tópico a seguir aborda estudos que examinaram a influência da IA na aprovação dos principais exames de certificação contábil, bem como sua aplicação na resolução de casos acadêmicos.

Eulerich *et al.* (2023a) examinaram se os modelos de IA são capazes de obter aprovação nos principais exames de certificação contábil, incluindo os exames CPA, CMA, CIA e EA. Como resultado, constatou-se que o ChatGPT não foi capaz de obter êxito em nenhum desses exames (com uma pontuação média em todas as avaliações de 53,1%). No entanto, o GPT-4 apresentou desempenho significativamente melhor. Além disso, ao fornecer treinamento com 10 exemplos, as pontuações foram aprimoradas em mais 6,6 pontos percentuais. Permitir que o modelo utilize raciocínio e ação, como permitir o uso de calculadora e outros recursos, melhorou as pontuações em mais 8,9 pontos percentuais. Após todas essas melhorias, o GPT-4 alcançou média 85,1% de acerto em todas as seções dos exames e obteve aprovação em todos eles. Segundo os autores, esse alto desempenho sugere que a IA possui potencial disruptivo nas indústrias de contabilidade e auditoria.

Eulerich *et al.* (2023b) investigaram as possíveis implicações dos grandes modelos de linguagem (LLMs), como os transformadores pré-treinados generativos, no mercado de trabalho dos Estados Unidos. Como resultado, identificaram que cerca de 80% da força de

trabalho daquele país poderia ter pelo menos 10% de suas tarefas afetadas pela introdução dos LLMs, enquanto aproximadamente 19% dos trabalhadores podem ter pelo menos metade de suas tarefas impactadas. Os resultados sugerem também que, com acesso a um LLM, cerca de 15% de todas as tarefas dos trabalhadores nos EUA poderiam ser concluídas significativamente mais rapidamente, mantendo o nível de qualidade. Ao incorporar *software* e ferramentas construídas com base nos LLMs, essa proporção aumenta para entre 47% e 56% de todas as tarefas.

Por outro lado, Chen *et al.* (2023) avaliaram as capacidades dos modelos ChatGPT e GPT-4 para fornecer soluções em oito casos de ensino em contabilidade. Como resultado, observaram que a capacidade do ChatGPT de fornecer soluções precisas varia dependendo do tipo de requisito do caso e da área temática da contabilidade, com melhor desempenho em tarefas que exigem explicação, aplicação de regras e avaliação ética. No entanto, o ChatGPT apresentou um desempenho ruim em tarefas que exigem elaboração de demonstrações financeiras, lançamentos contábeis ou uso de *software*. Além disso, constataram que as ferramentas de detecção fornecidas pelo ChatGPT e GPT-4 são ineficazes na identificação de texto criados pela IA. Concluem que desenvolver uma compreensão da capacidade do ChatGPT em fornecer soluções para casos de ensino em contabilidade é importante, dado que muitos materiais de casos são concluídos fora da sala de aula, e os educadores possuem menos controle sobre a capacidade dos alunos de utilizar ferramentas tecnológicas quando os casos de contabilidade são feitos como tarefas de casa ou projetos, o que pode ameaçar potencialmente a integridade acadêmica.

Wood *et al.* (2023) observaram que o ChatGPT geralmente tinha um desempenho inferior aos alunos, embora houvesse alguma variação no que diz respeito ao seu desempenho em áreas temáticas de contabilidade e tipos de questões do exame. Os autores sugerem que os instrutores não precisam se preocupar excessivamente com os alunos que utilizam sistemas de *chat* baseados em IA para responder às questões do exame. Essa conclusão leva em consideração tanto o desempenho dos alunos quanto as limitações práticas comumente encontradas durante os exames, como a realização em sala de aula ou a presença de um fiscal. No entanto, os resultados também sugerem que os sistemas de *chat* baseados em IA têm potencial para alcançar desempenhos superiores. A versão mais recente do ChatGPT se aproxima da qualidade média das respostas dos alunos em diversos requisitos contábeis do caso.

Borji e Mohammadian (2023) realizaram uma avaliação extensiva de LLMs e *bots* conversacionais usando o "conjunto de dados Wordsmiths", composto por 1002 perguntas em

27 categorias. Nessa avaliação, o GPT-4 se destacou com o melhor desempenho em todas as categorias, alcançando uma taxa de sucesso de 84,1%. Por outro lado, Bard enfrentou desafios e alcançou uma taxa de sucesso de 62,4%. O Bard é menos correlacionado com outros modelos, enquanto ChatGPT e GPT-4 são altamente correlacionados em termos de suas respostas. Segundo os autores, os *Chatbots* mostraram competência em linguagem e fatos, mas enfrentaram desafios em matemática e raciocínio. Os modelos também foram avaliados quanto ao viés e ética, demonstrando serem relativamente seguros.

Nunes *et al.* (2023) investigaram a eficácia dos modelos LLM, como GPT-3.5 e GPT-4, em testes como o ENEM utilizado nas universidades brasileiras. Ao analisar as questões de 2009-2017 e 2022, o estudo explorou várias estratégias de *prompt*, destacando-se a *Chain-of-Thought* (CoT) com o GPT-4 alcançando 87% de precisão em 2022, 11% a mais que o GPT-3.5. A pesquisa considerou a memorização de respostas pelos modelos, enfatizando a relevância dos LLM em educação e indicando direções para futuras inovações em aprendizado multidisciplinar.

2. CONTEXTO ANALISADO: EXAME DE SUFICIÊNCIA E EXAMES TÉCNICOS

O presente estudo avalia o desempenho dos modelos de IA nos exames de certificação profissional do CFC considerando a temática da questão conforme taxonomia estabelecida no respectivo edital. Sendo assim, buscou-se o entendimento com relação aos conteúdos e estrutura das provas para classificação das questões dos exames explorados.

O Exame de Suficiência é conduzido pela Consultoria e Planejamento em Administração (Consulplan), sob sua total responsabilidade e controle. Seu objetivo é comprovar o domínio dos conhecimentos médios adquiridos durante o curso de Bacharelado em Ciências Contábeis, permitindo que os aprovados obtenham o registro na categoria de Contador. Importante mencionar que este exame é destinado exclusivamente a bacharéis e estudantes do último ano do curso de Ciências Contábeis.

Por outro lado, as provas do Exame de Qualificação Técnica (EQT) são realizadas conforme as diretrizes do CFC. Essas provas avaliam o conhecimento e a competência técnico-profissional dos contadores que desejam atuar como auditores independentes em organizações reguladas pela Comissão de Valores Mobiliários (CVM), Banco Central do Brasil (BCB), Superintendência de Seguros Privados (SUSEP) e Superintendência Nacional de Previdência Complementar (PREVIC). O EQT também é relevante para obter registro no Cadastro Nacional dos Auditores Independentes (CNAI) do CFC ou ingressar na área de perícia contábil. Os exames técnicos são destinados aos contadores devidamente registrados no CRC de sua jurisdição, essas avaliações têm um papel fundamental para garantir a qualidade e a competência dos profissionais contábeis no exercício de suas atividades. As matérias abordadas nos exames, conforme as orientações do CFC, são listadas a seguir:

- I. Exame suficiência (CFC): A prova destinada ao exame de suficiência dos bacharéis em Ciências Contábeis inclui as seguintes disciplinas: Contabilidade Geral, Contabilidade de Custos, Contabilidade Aplicada ao Setor Público, Contabilidade Gerencial, Controladoria, Noções de Direito e Legislação Aplicada, Matemática Financeira e Estatística, Teoria da Contabilidade, Legislação e Ética Profissional, Princípios de Contabilidade e Normas Brasileiras de Contabilidade, Auditoria Contábil, Perícia Contábil e Língua Portuguesa Aplicada.

- II. Prova de Qualificação Técnica Geral (QTG): Legislação e Ética Profissional; Princípios Fundamentais de Contabilidade e Normas Brasileiras de Contabilidade e suas Interpretações Técnicas, editadas pelo CFC.
- III. Prova Específica para atuar em auditoria de instituições reguladas pela CVM: Legislação Societária; Legislação e normas aplicáveis às Instituições reguladas pela CVM; Língua Portuguesa Aplicada.
- IV. Prova Específica para Atuação na Área de Auditoria nas Instituições Reguladas pelo BCB: Legislação sobre a Profissão Contábil; Princípios Fundamentais de Contabilidade e Normas Brasileiras de Contabilidade e suas Interpretações Técnicas, editadas pelo CFC; Legislação e Normas do Sistema Financeiro Nacional (SFN); Conhecimento de Operações da Área de Instituições Reguladas pelo Banco Central do Brasil e Contabilidade das Instituições do Sistema Financeiro Nacional.
- V. Prova Específica para Atuação na Área de Auditoria nas Instituições Reguladas pela SUSEP: Legislação sobre a Profissão Contábil; Princípios Fundamentais de Contabilidade e Normas Brasileiras de Contabilidade e suas Interpretações Técnicas, editadas pelo CFC; Normas e Procedimentos Contábeis e de Auditoria Aplicáveis às Entidades Reguladas pela SUSEP; Conhecimento da Estrutura do Sistema Nacional de Seguros Privados e das Operações da Área de Entidades Reguladas pela SUSEP; e Contabilidade das Instituições Reguladas pela SUSEP.
- VI. Prova específica para atuação do contador em Auditoria Independente nas entidades supervisionadas pela PREVIC: Conhecimentos de (a) legislação e normas aplicáveis às entidades supervisionadas pela PPREVIC; (b) operações de entidades supervisionadas pela PREVIC; (c) contabilidade de entidades supervisionadas pela PPREVIC.
- VII. Prova de Qualificação Técnica para Perícia Contábil: Legislação Profissional; Ética Profissional; Normas Brasileiras de Contabilidade, Técnicas e Profissionais, editadas pelo CFC, inerentes à perícia; Legislação Processual Civil aplicada à perícia; Língua Portuguesa e Redação; Direito Constitucional, Civil e Processual Civil afetos à legislação profissional, à prova pericial e ao perito.

Com o objetivo de avaliar a homogeneidade das distribuições dos acertos por disciplina resultantes dos três modelos de Inteligência Artificial, foi necessário agrupar as questões do exame de suficiência nas seguintes disciplinas: (i) Contabilidade Geral; (ii) Contabilidade Gerencial; (iii) Contabilidade Aplicada ao Setor Público, (iv) Matemática Financeira e Estatísticas; (v) Noções de Direito e Legislação Aplicada; (vi) Prin. de Cont. de Normas Brasileiras de Contabilidade; (vii) Legislação e ética Profissional, (viii) Auditoria e Perícia Contábil. (ix) Língua Portuguesa Aplicada.

Já as questões dos exames técnicos foram classificadas nas seguintes disciplinas: (i) Normas Profissionais, Técnicas, Legislação e ética Profissional; (ii) Legislação, Normas do Mercado de Capitais e Bovespa; (iii) Normas Brasileiras de Contabilidade (iv) Conhecimento de Operações da área de Instituições Reguladas pelo Banco Central (BCB);(v) Legislação e Normas Aplicáveis as sociedades supervisionas pela SUSEP; (vi) Legislação e Normas Aplicáveis as entidades supervisionadas pela PREVIC.

3. ABORDAGEM METODOLÓGICA

Adotou-se abordagem quantitativa para coletar, analisar e interpretar os dados resultantes da aplicação dos modelos de IA (ChatGPT, GPT-4 e Bard) nas edições de 2022 dos exames de suficiência do CFC e dos exames de qualificação técnica (QTG, CVM, BCB, SUSEP, PREVIC e Perícia Contábil).

Para os exames de suficiência, foi possível obter no *site* oficial do CFC as estatísticas dos candidatos que participaram de ambas as edições em 2022, e os critérios de aprovação estabelecidos pelo CFC, que consistem em alcançar no mínimo 50% de acerto no exame de suficiência e no mínimo 60% de acerto na prova objetiva para os exames de qualificação técnica.

É importante ressaltar que no *site* do CFC não foram encontrados dados estatísticos relativos ao desempenho humano nos exames de qualificação técnica, o que não permitiu de avaliar o desempenho da IA em relação aos humanos por disciplina para essas provas. Além disso, a análise para esses exames concentrou-se apenas na taxa de aprovação (60% de acerto) dos modelos de IA nas questões objetivas dos exames.

No que tange aos dados para a realização do estudo, foram solicitadas ao CFC estatísticas fundamentais (Apêndice B) relacionadas às provas administradas pelo CFC no ano de 2022: (i) Quantidade de candidatos que realizaram cada prova; (ii) Quantidade de candidatos que acertaram cada questão (ou o percentual de acertos por questão); (iii) O tipo de conteúdo avaliado em cada questão (por exemplo: Auditoria Contábil, Contabilidade Aplicada ao Setor Público, Contabilidade de Custos, etc.); (iv) Quais são essas abordagens (teoria ou critério de taxonomia) para cada prova aplicada em 2022?; (v) Qual a classificação efetiva de cada questão das provas de 2022 (Suficiência 1a edição, Suficiência 2a edição, CVM, BACEN, SUSEP, PREVIC e Perícia)?

A resposta recebida sobre a solicitação de abertura de dados (Apêndice C) não atendeu de forma satisfatória às necessidades deste estudo. Quanto à primeira solicitação, que se referia ao total de candidatos que prestaram o exame, o CFC forneceu informações apenas para os exames técnicos. No entanto, em relação à segunda solicitação, sobre o percentual de acerto por questão, foi informado que tal dado não existe. Em relação ao terceiro questionamento, o CFC esclareceu que não possui relatórios segmentados por conteúdo específico.

Sobre as duas últimas solicitações, o CFC elucidou que o Exame de Suficiência visa comprovar a aquisição de conhecimentos medianos alinhados aos conteúdos programáticos do curso de Bacharelado em Ciências Contábeis. Portanto, sua dificuldade é considerada mediana. Em contrapartida, o Exame de Qualificação Técnica (EQT) para Auditor e Perito Contábil busca avaliar o nível de conhecimento e a competência técnico-profissional de contadores que desejam atuar em auditoria e perícia contábil, demandando deles um profundo entendimento técnico e científico na área. Foi ainda ressaltado que não existe previsão contratual para interação com os elaboradores dos exames, devido ao sigilo e integridade que o processo exige.

3.1 Amostra e coleta de dados

Como parte do escopo do trabalho, optou-se por aplicar todos os exames de suficiência e qualificação técnica realizados em 2022 nos três modelos de IA avaliados. Para isso, foi necessário acessar o *site* do CFC para obter as provas e seus respectivos gabaritos correspondentes a esse período. Além disso, verificou-se se os modelos obtiveram aprovação em todas as provas, conforme os critérios de aprovação disponibilizados no *site* do CFC.

É importante destacar que a OpenAI implementou a mais recente atualização em sua inteligência artificial em abril de 2023, enquanto a atualização final do Bard ocorreu em 1º de agosto de 2023.

3.2 Query para identificar o desempenho da IA

As questões das provas foram aplicadas integralmente nos modelos de IA e as alternativas escolhidas pela IA foram registradas. A documentação das questões aplicadas pode ser encontrada no Apêndice A desta dissertação. A alternativa de resposta selecionada pela IA foi registrada em planilha eletrônica.

É importante ressaltar que existem estratégias *in-context*, ou seja, a criação de *prompts* específicos que potencializam o desempenho do Modelo de Linguagem de Grande Escala (LLM). Contudo, a implementação dessas técnicas de orientação não foi abordada nesta

pesquisa como parte da estratégia de interação com a Inteligência Artificial. A rigor, conforme demonstrado no Apêndice A, cada questão das provas analisadas foi copiada e colada no *playground* do chat, sem qualquer contexto que pudesse direcionar a tecnologia a obter melhor desempenho.

3.3 Metodologia de análise de dados

Com base nos dados obtidos através da aplicação das questões nos modelos de IA, empregou-se uma abordagem quantitativa para analisar e interpretar os resultados das provas de exame. Foram analisadas as características das questões, considerando a área temática e seguindo a taxonomia estabelecida no respectivo edital, a fim de avaliar o desempenho disciplinar em comparação com a IA.

Primeiramente, realizou-se uma análise descritiva e comparativa das taxas de acerto, bem como das frequências absolutas/relativas dos acertos e erros dos modelos de IA, juntamente com o desempenho médio dos seres humanos nos exames. Na sequência, foi utilizado o *software* estatístico R (R Core Team, 2022) para realizar testes exatos de Fisher (Upton e Cook, 2008) objetivando-se investigar a existência de diferenças estatisticamente significativas nas distribuições dos acertos e erros entre os modelos de IA e o efeito da presença de questões com tabelas na performance de cada modelo avaliado. Ademais, foram aplicados testes qui-quadrado de homogeneidade (Verzani, 2004) visando avaliar a significância das diferenças nas distribuições dos acertos e erros entre os modelos de IA nos diferentes temas avaliados (grupos de disciplinas).

Além disso, empregou-se o índice de similaridade de Tanimoto para mensurar a afinidade na correção das respostas geradas por dois modelos distintos de inteligência artificial (O’Leary, 2023). Em um exame com N questões, as respostas de um modelo de IA podem ser representadas numericamente como 0 para incorretas e 1 para corretas, resultando na formação de um vetor com N elementos binários. Denotando por $a(j) \in \{0,1\}$ e $b(j) \in \{0,1\} \forall j=1,N$ as respostas de dois modelos A e B, o índice de similaridade de Tanimoto entre os modelos A e B é dado por:

$$T = \frac{C}{A + B - C} \quad (1)$$

em que

$A = \sum_{j=1}^N a(j)$ total de respostas corretas do modelo A

$B = \sum_{j=1}^N b(j)$ total de respostas do modelo B

$C = \sum_{j=1}^N a(j) \times b(j)$ total de questões respondidas corretamente pelos modelos A e B

O índice de Tanimoto varia entre 0 e 1, em que 0 indica nenhuma similaridade entre as respostas e 1 indica respostas idênticas. O’Leary (2023) sugere a seguinte correspondência entre o índice de Tanimoto e o grau de similaridade entre as respostas de dois modelos de IA:

- similaridade moderada: $0,50 < T \leq 0,70$;
- similaridade alta: $0,70 < T \leq 0,90$;
- respostas essencialmente idênticas: $T > 0,90$

Por último, foi promovido uma análise qualitativa das questões discursivas das provas técnicas QGT e CVM para investigar a capacidade da IA responder questões discursivas, observando-se tanto os conhecimentos técnicos quanto a habilidade textual.

4. RESULTADOS

Os resultados são apresentados por exame analisado.

4.1 Análise dos Exames de Suficiência de 2022

Conforme demonstrado na Tabela 1, o ChatGPT obteve aprovação apenas na primeira edição do exame de suficiência de 2022, e o GPT-4 foi aprovado em ambas as edições, apresentando aproveitamento médio de 70% para o conjunto das duas provas. Já o Bard foi aprovado apenas na segunda edição do exame de suficiência:

Tabela 1: Resumo Resultado Exames de Suficiência 2022

Exame	Questões	Modelo de IA	Nota ou Acertos	Aproveitamento Taxa de acertos	Taxa de erros	Resultado
1º exame suficiência CRC 2022	50	ChatGPT	28	56%	34%	Aprovado
	50	GPT-4	34	68%	32%	Aprovado
	50	Bard	22	44%	56%	Reprovado
2º exame suficiência CRC 2022	50	ChatGPT	19	38%	62%	Reprovado
	50	GPT-4	36	72%	28%	Aprovado
	50	Bard	25	50%	50%	Aprovado
Consolidado CFC 2022	100	ChatGPT	47	47%	53%	Reprovado
	100	GPT4	70	70%	30%	Aprovado
	100	Bard	47	47%	53%	Reprovado

Fonte: A autora, 2023.

Na Tabela 1, o resultado consolidado por modelo de IA revela que o GPT4 se destaca com uma elevada taxa de acerto de 70%, mais que o dobro da taxa de erros. Em contrapartida, tanto o ChatGPT quanto o Bard exibem as mesmas taxas de acertos e erros, indicando a capacidade comparativamente menor do ChatGPT e Bard. Vale destacar que o teste exato de Fisher confirma o desempenho significativamente superior do GPT-4 em relação ao ChatGPT e Bard (p-valor = 0,002).

Na primeira edição do exame de suficiência 2022, as notas alcançadas pelo Chatgpt e GPT4 superaram a média das notas de humanos em todas as regiões e estados do Brasil, conforme indicado na Tabela 2, O Bard alcançou apenas 22 acertos (taxa de acerto < 50%), um valor insuficiente para aprovação no exame, mas equivalente à média observada nos estados com as melhores performances, O ChatGPT com 28 acertos foi um pouco melhor que a média

dos humanos, alcançando uma taxa de acerto da ordem de 56%, valor suficiente para aprovação.

Já o GPT4 alcançou o melhor resultado com 34 acertos, valor acima das médias nos estados.

Tabela 2: Média nacional de humanos por região/estado e nota da IA, 1º exame de suficiência 2022

1º Exame Suficiência 2022				
Região	Média humanos	Desempenho ChatGPT (28) melhor que a média dos humanos em	Desempenho GPT-4 (34) melhor que a média dos humanos em	Desempenho Bard (22) melhor que a média dos humanos em
CENTRO-OESTE	21	33%	62%	5%
<i>DF</i>	22	27%	55%	0%
<i>GO</i>	21	33%	62%	5%
<i>MT</i>	20	40%	70%	10%
<i>MS</i>	20	40%	70%	10%
NORDESTE	21	33%	62%	5%
<i>AL</i>	20	40%	70%	10%
<i>BA</i>	20	40%	70%	10%
<i>CE</i>	22	27%	55%	0%
<i>MA</i>	20	40%	70%	10%
<i>PB</i>	21	33%	62%	5%
<i>PE</i>	21	33%	62%	5%
<i>PI</i>	18	56%	89%	22%
<i>RN</i>	18	56%	89%	22%
<i>SE</i>	19	47%	79%	16%
NORTE	19	47%	79%	16%
<i>AC</i>	18	56%	89%	22%
<i>AM</i>	20	40%	70%	10%
<i>AP</i>	19	47%	79%	16%
<i>PA</i>	19	47%	79%	16%
<i>RO</i>	19	47%	79%	16%
<i>RR</i>	20	40%	70%	10%
<i>TO</i>	19	47%	79%	16%
SUDESTE	22	27%	55%	0%
<i>ES</i>	22	27%	55%	0%
<i>MG</i>	22	27%	55%	0%
<i>RJ</i>	22	27%	55%	0%
<i>SP</i>	22	27%	55%	0%
SUL	22	27%	55%	0%
<i>PR</i>	21	33%	62%	5%
<i>RS</i>	22	27%	55%	0%
<i>SC</i>	22	27%	55%	0%

Fonte: A autora, 2023.

Conforme ilustrado na Tabela 3, na segunda edição do exame de suficiência 2022, o GPT-4 e o Bard apresentaram desempenhos ligeiramente melhores que os observados na

primeira edição e obtiveram notas superiores às médias dos humanos em todas as regiões e estados do Brasil. Contudo, o ChatGPT obteve apenas 19 acertos, uma nota igual a menor média dos estados e insuficiente para aprovação (taxa de acerto < 50%).

Tabela 3: Média nacional de humanos por região/estado e nota da IA, 2º exame de suficiência 2022

2º Exame Suficiência 2022				
Região	Média humanos	Desempenho ChatGPT (19) inferior, igual ou superior (apenas AP) que humanos em	Desempenho GPT-4 (36) melhor que humanos em	Desempenho Bard (25) melhor que humanos
CENTRO-OESTE	20	-5%	80%	25%
<i>DF</i>	22	-14%	64%	14%
<i>GO</i>	20	-5%	80%	25%
<i>MT</i>	20	-5%	80%	25%
<i>MS</i>	20	-5%	80%	25%
NORDESTE	21	-10%	71%	19%
<i>AL</i>	20	-5%	80%	25%
<i>BA</i>	21	-10%	71%	19%
<i>CE</i>	22	-14%	64%	14%
<i>MA</i>	19	0%	89%	32%
<i>PB</i>	20	-5%	80%	25%
<i>PE</i>	21	-10%	71%	19%
<i>PI</i>	20	-5%	80%	25%
<i>RN</i>	20	-5%	80%	25%
<i>SE</i>	21	-10%	71%	19%
NORTE	19	0%	89%	32%
<i>AC</i>	19	0%	89%	32%
<i>AM</i>	19	0%	89%	32%
<i>AP</i>	19	0%	89%	32%
<i>PA</i>	19	0%	89%	32%
<i>RO</i>	19	0%	89%	32%
<i>RR</i>	19	0%	89%	32%
<i>TO</i>	19	0%	89%	32%
SUDESTE	21	-10%	71%	19%
<i>ES</i>	21	-10%	71%	19%
<i>MG</i>	21	-10%	71%	19%
<i>RJ</i>	22	-14%	64%	14%
<i>SP</i>	21	-10%	71%	19%
SUL	22	-14%	64%	14%
<i>PR</i>	22	-14%	64%	14%
<i>RS</i>	22	-14%	64%	14%
<i>SC</i>	22	-14%	64%	14%

Fonte: A autora, 2023.

É importante destacar que, na primeira edição do exame de suficiência, a nota máxima alcançada foi de 46 acertos na região Centro-Oeste, enquanto a nota mínima foi de 2 acertos nas regiões Sudeste e Nordeste. Na segunda edição, a nota máxima registrada foi de 45 acertos na região Sul, com a mínima sendo de 3 acertos nas regiões Sudeste e Nordeste do país. Os dados divulgados pelo CFC indicam somente as notas máximas e mínimas por região e estado, sem detalhar quantos alunos atingiram essas pontuações.

Conforme ilustrado na Tabela 4, no que tange ao desempenho por disciplina, na primeira edição do exame de suficiência 2022, o ChatGPT obteve 100% de acerto em Auditoria Contábil, Controladoria e Matemática Financeira, mas não acertou nenhuma questão em Contabilidade Gerencial, Língua Portuguesa Aplicada e Princípios da Contabilidade e de Normas Brasileiras de Contabilidade; nas demais disciplinas o ChatGPT alcançou um desempenho igual ou superior a 50%. Por outro lado, o GPT-4 acertou todas as questões de Controladoria, Matemática Financeira e Noções de Direito e Legislação Aplicada, e errou todas de Contabilidade Gerencial; e obteve desempenho igual ou superior a 50% nas demais disciplinas. O Bard alcançou 100% de acerto para Contabilidade Aplicada ao Setor Público, Controladoria, Legislação e Ética profissional e Noções de Direito e Legislação Aplicada.

Tabela 4: Desempenho humano e IA por disciplina, 1º exame de suficiência 2022

1º Exame de Suficiência 2022: Disciplina (número de questões)	% Média de Acerto Nacional	Acertos ChatGPT	Acertos GPT-4	Acertos Bard
Auditoria Contábil (2)	84%	100%	50%	50%
Contabilidade Aplicada ao Setor Público (3)	46%	67%	67%	100%
Contabilidade de Custos (4)	29%	50%	50%	0%
Contabilidade Geral (17)	41%	59%	76%	29%
Contabilidade Gerencial (2)	18%	0%	0%	0%
Controladoria (1)	48%	100%	100%	100%
Legislação e Ética profissional (4)	60%	75%	75%	100%
Língua Portuguesa Aplicada (2)	23%	0%	50%	0%
Matemática Financeira e Estatística (2)	18%	100%	100%	0%
Noções de Direito e Legislação Aplicada (3)	46%	67%	100%	100%
Perícia Contábil (2)	32%	50%	50%	50%
Prin. de Cont. de Normas Brasileiras de cont. (4)	32%	0%	50%	25%
Teoria da Contabilidade (4)	66%	75%	75%	75%

Fonte: A autora, 2023.

Na segunda edição do exame de suficiência 2022, o ChatGPT obteve um desempenho de 100% apenas em Auditoria Contábil. Nas demais disciplinas, o ChatGPT obteve as seguintes taxas de acerto: Contabilidade Aplicada ao Setor Público (33%), Contabilidade de Custos (50%), Contabilidade Geral (29%), Legislação e Ética Profissional (25%), Língua Portuguesa (50%), Noções de Direito e Legislação Aplicada (67%), Princípios de Contabilidade e Normas Brasileiras de Contabilidade (50%), Perícia Contábil (50%) e Teoria da Contabilidade (50%).

Já o modelo GPT-4, obteve um desempenho de 100% em Auditoria Contábil, Contabilidade Gerencial, Controladoria, Legislação e Ética Profissional, Língua Portuguesa Aplicada, Matemática Financeira e Perícia Contábil. Para as demais disciplinas, como Contabilidade de Custos, Contabilidade Geral, Noções de Direito e Legislação Aplicada, Princípios de Contabilidade e Normas Brasileiras de Contabilidade, Teoria da Contabilidade, o GPT-4 obteve um desempenho igual ou superior a 50%. O modelo Bard apresentou 100% de acerto em Matemática Financeira e Estatística e Perícia Contábil; conforme apresentado na Tabela 5.

Tabela 5: Desempenho humano e IA por disciplina, 2º exame de suficiência 2022

2º Exame de Suficiência 2022: Disciplina (número de questões)	% Médio de Acerto Nacional	Acertos ChatGPT	Acertos GPT-4	Acertos Bard
Auditoria Contábil (2)	37,69%	100%	100%	50%
Contabilidade Aplicada ao Setor Público (3)	27,11%	33%	33%	67%
Contabilidade de Custos (4)	33,26%	50%	75%	25%
Contabilidade Geral (17)	42,96%	29%	59%	47%
Contabilidade Gerencial (2)	33,24%	0%	100%	0%
Controladoria (1)	16,76%	0%	100%	0%
Legislação e Ética profissional (4)	62,74%	25%	100%	25%
Língua Portuguesa Aplicada (2)	24,87%	50%	100%	50%
Matemática Financeira e Estatística (2)	47,44%	0%	100%	100%
Noções de Direito e Legislação Aplicada (3)	36,98%	67%	67%	67%
Perícia Contábil (2)	62,24%	50%	100%	100%
Prin. de Cont. de Normas Brasileiras de cont. (4)	38,35%	50%	50%	50%
Teoria da Contabilidade (4)	52,53%	50%	75%	75%

Fonte: A autora, 2023.

A seguir, apresentam-se os resultados da análise descritiva focada na presença de tabelas nas questões e sua influência no desempenho dos modelos de IA. É importante salientar que as observações aqui reportadas se referem especificamente ao desempenho desses modelos nos exames de suficiência de 2022, conforme investigado neste estudo.

No primeiro exame de suficiência de 2022 a presença de tabelas no enunciado das questões não prejudicou o desempenho do ChatGPT, no entanto, prejudicou o desempenho do GPT-4 e do Bard, conforme evidenciado na Tabela 6 pela frequência de erros maior que a frequência de acertos nas questões com tabela.

Tabela 6: Desempenho da IA no 1º exame suficiência CFC 2022, por disciplina e enunciados com ou sem tabela

1º exame suficiência CFC 2022	Total de questões	Chat-GPT		GPT-4		Bard	
		Acertos	Erros	Acertos	Erros	Acertos	Erros
Questões com tabela	10	6	4	4	6	2	8
Auditoria Contábil	1	1			1		1
Contabilidade Aplicada ao Setor Público	1	1			1	1	
Contabilidade de Custos	2	1	1		2		2
Contabilidade Geral	5	2	3	3	2	1	4
Contabilidade Gerencial	0						
Controladoria	0						
Legislação e Ética profissional	0						
Língua Portuguesa Aplicada	0						
Matemática Financeira e Estatística	1	1		1			1
Noções de Direito e Legislação Aplicada	0						
Perícia Contábil	0						
Prin. de Cont. de Normas Brasileiras de Cont.	0						
Teoria da Contabilidade	0						
Questões sem tabela	40	22	18	30	10	20	20
Auditoria Contábil	1	1		1		1	
Contabilidade Aplicada ao Setor Público	2	1	1	2		2	
Contabilidade de Custos	2	1	1	2			2
Contabilidade Geral	12	8	4	10	2	4	8
Contabilidade Gerencial	2		2		2		2
Controladoria	1	1		1		1	
Legislação e Ética profissional	4	3	1	3	1	4	
Língua Portuguesa Aplicada	2		2	1	1		2
Matemática Financeira e Estatística	1	1		1			1
Noções de Direito e Legislação Aplicada	3	2	1	3		3	
Perícia Contábil	2	1	1	1	1	1	1
Prin. de Cont. de Normas Brasileiras de Cont.	4		4	2	2	1	3
Teoria da Contabilidade	4	3	1	3	1	3	1
Total geral	50	28	22	34	16	22	28

Fonte: A autora, 2023.

No segundo exame de suficiência de 2022 a presença de tabelas no enunciado das questões não prejudicou o desempenho do GPT-4, dado que o número de acertos é o dobro do

número de erros. No entanto, prejudicou o desempenho do ChatGPT e o Bard, conforme demonstrado na Tabela 7.

Tabela 7: Desempenho da IA no 2º exame suficiência CFC 2022, por disciplina e enunciados com ou sem tabela

2º exame suficiência CFC 2022	Total de questões	Chat-GPT		GPT-4		Bard	
		Acertos	Erros	Acertos	Erros	Acertos	Erros
Questões com tabela	18	5	13	12	6	8	10
Auditoria Contábil	0						
Contabilidade Aplicada ao Setor Público	1		1		1	1	
Contabilidade de Custos	3	1	2	2	2		3
Contabilidade Geral	10	4	6	7	3	6	4
Contabilidade Gerencial	2		2	2			2
Controladoria	0						
Legislação e Ética profissional	0						
Língua Portuguesa Aplicada	0						
Matemática Financeira e Estatística	1		1	1		1	
Noções de Direito e Legislação Aplicada	1		1				1
Perícia Contábil	0						
Prin. de Cont. de Normas Brasileiras de Cont.	0						
Teoria da Contabilidade							
Questões sem tabela	32	14	18	24	8	17	15
Auditoria Contábil	2	2		2		1	1
Contabilidade Aplicada ao Setor Público	2	1	1	1	1	1	1
Contabilidade de Custos	1	1		1		1	
Contabilidade Geral	7	1	6	3	4	2	5
Contabilidade Gerencial							
Controladoria	1		1	1			1
Legislação e Ética profissional	4	1	3	4		1	3
Língua Portuguesa Aplicada	2	1	1	2		1	1
Matemática Financeira e Estatística	1		1	1		1	
Noções de Direito e Legislação Aplicada	2	2		2		2	
Perícia Contábil	2	1	1	2		2	
Prin. de Cont. de Normas Brasileiras de Cont.	4	2	2	2	2	2	2
Teoria da Contabilidade	4	2	2	3	1	3	1
Total	50	19	31	36	16	25	25

Fonte: A autora, 2023.

Na Tabela 8 apresentam-se os valores consolidados das taxas de acertos e de erros dos modelos de IA nas questões com e sem tabelas nas edições dos exames do CFC realizados em 2022. Os resultados mostram que a taxa de acertos é menor nas questões com tabela, sendo que no ChatGPT e Bard a presença de tabelas implica em taxas de erros acima de 60%, enquanto no GPT4 a taxa de acertos nas questões com tabela é da ordem de 57%, mostrando mais vez o

melhor desempenho do GPT4. Por outro lado, nas questões sem tabela o ChatGPT e o Bard apresentaram taxas de acertos e de erros equilibradas, indicando a distribuição uniforme de respostas corretas e erradas. Já o GPT4 destaca-se pela predominância de respostas corretas entre as questões sem tabela, alcançando uma taxa de acertos de 75%.

A avaliação da significância estatística das diferenças entre as distribuições de erros e acertos nas questões com e sem tabela foi realizada pelo teste exato de Fisher unilateral (Toyama et al., 2023):

H_0 proporção de acertos sem tabela \leq proporção de acertos com tabela

H_1 proporção de acertos sem tabela $>$ proporção de acertos com tabela

O teste apontou significância estatística marginal apenas no caso do GPT-4 (p-valor 0,068), i.e., o desempenho do GPT-4 nas questões sem tabela é significativamente superior ao desempenho nas questões com tabela.

Tabela 8: Desempenho da IA no CFC 2022 com ou sem tabela

	Número de Questões	Estatística	ChatGPT		GPT4		Bard	
			Acertos	Erros	Acertos	Erros	Acertos	Erros
Com tabela	28	Frequências	11	17	16	12	10	18
		Taxas	39,3%	60,7%	57,1%	42,9%	35,7%	64,3%
Sem tabela	72	Frequências	36	36	54	18	37	35
		Taxas	50,0%	50,0%	75,0%	25,0%	51,4%	48,6%
p-valor do teste exato de Fisher			0,23		0,068*		0,117	

Nota: * indica significância estatística no nível de 10%

Fonte: A autora, 2023.

Os valores para o índice de Tanimoto na Tabela 9 refletem a afinidade moderada entre as respostas dos três modelos de IA.

Tabela 9: Índice de Tanimoto

	ChatGPT	GPT4	Bard
ChatGPT	1	0,5195	0,5161
GPT4	0,5195	1	0,5811
Bard	0,5161	0,5811	1

Fonte: A autora, 2023.

4.2 Análises dos Exames de Qualificação Técnica 2022

O GPT-4 seria aprovado em cinco exames de qualificação técnica aplicados pelo CFC em 2022, reprovando apenas no SUSEP. Já o ChatGPT conseguiu aprovação em quatro exames técnicos, reprovando apenas nos exames BCB e SUSEP. O Bard conseguiu aprovação em apenas três, reprovando nos exames CMV, BCB e SUSEP, conforme detalhado na Tabela 10. Vale destacar que os três modelos de IA não foram aprovados no exame da SUSEP.

Tabela 10: Desempenho IA, exames de qualificação técnica 2022:

Exame	Questões	2022				
		Modelo	Nota ou Acertos	Aproveitamento Taxa de acertos	Taxa de erros	Resultado
QTG 2022	25	ChatGPT	18	72%	28%	Aprovado
		GPT-4	22	88%	12%	Aprovado
		Bard	18	72%	28%	Aprovado
CVM 2022	25	ChatGPT	15	60%	40%	Aprovado
		GPT-4	21	84%	16%	Aprovado
		Bard	13	52%	48%	Reprovado
BCB 2022	25	ChatGPT	14	56%	44%	Reprovado
		GPT-4	18	72%	28%	Aprovado
		Bard	11	44%	56%	Reprovado
SUSEP 2022	25	ChatGPT	10	40%	60%	Reprovado
		GPT-4	12	48%	52%	Reprovado
		Bard	10	40%	60%	Reprovado
PREVIC 2022	25	ChatGPT	16	64%	36%	Aprovado
		GPT-4	19	76%	24%	Aprovado
		Bard	19	76%	24%	Aprovado
Perito 2022	25	ChatGPT	16	64%	36%	Aprovado
		GPT-4	17	68%	32%	Aprovado
		Bard	17	68%	32%	Aprovado

Fonte: A autora, 2023.

A seguir, a Tabela 11 exibe a performance de cada modelo de IA no conjunto formado pelas 150 questões dos Exames Técnicos. Os resultados consolidados mostram, mais uma vez, o melhor desempenho do GPT-4 com uma notável taxa de acerto da ordem de 72%, superando em mais de 10% as taxas de acerto dos demais modelos. O teste exato de Fisher confirma o desempenho significativamente superior do GPT-4 em relação ao ChatGPT (p-valor=0,02) e Bard (p-valor = 0,015).

Tabela 11: Resultado consolidado nas 150 questões dos Exames Técnicos 2022

Modelo de IA	Número de acertos	Taxa de acertos	Taxa de erros
ChatGPT	89	59%	41%
GPT4	109	73%	27%
Bard	88	59%	41%

Fonte: A autora, 2023.

Como no caso do CFC 2022, os valores para o índice de Tanimoto na Tabela 12 refletem a afinidade moderada entre as respostas dos três modelos de IA.

Tabela 12: Índice de Tanimoto

	ChatGPT	GPT4	Bard
ChatGPT	1	0,6364	0,6239
GPT4	0,6364	1	0,6417
Bard	0,6239	0,6417	1

Fonte: A autora, 2023.

4.3 Análise: agrupamento de disciplinas e distribuição dos acertos alcançados em cada modelo de Inteligência Artificial.

Na sequência, utilizou-se o teste qui-quadrado para avaliar a homogeneidade das distribuições dos acertos/erros por disciplina dos três modelos de IA. Conforme Larson (1982), o teste qui-quadrado de homogeneidade é um método para testar se duas ou mais distribuições multinomiais são iguais, i.e., se as proporções de acertos/erros são semelhantes entre as disciplinas.

Com o intuito de evitar frequências esperadas reduzidas, as quais poderiam impactar negativamente no teste, a realização do teste qui-quadrado de homogeneidade foi precedida pela consolidação das disciplinas em um número reduzido de temas ou grupos.

Assim para o exame de suficiência classificou-se as questões nas seguintes disciplinas: (i) Contabilidade Geral; (ii) Contabilidade Gerencial, Custos e Controladoria; (iii) Contabilidade Aplicada ao Setor Público, (iv) Matemática Financeira e Estatísticas; (v) Noções de Direito e Legislação Aplicada; (vi) Prin. de Cont. de Normas Brasileiras de Contabilidade e Teoria da Contabilidade; (vii) Legislação e ética Profissional, (viii) Auditoria e Perícia Contábil e (ix) Língua Portuguesa Aplicada.

Conforme apresentado na Tabela 13, o ChatGPT obteve 100% acerto em Matemática Financeira, mas não acertou nenhuma questão em Língua Portuguesa Aplicada. Por outro lado, o GPT-4 acertou todas as questões de Matemática Financeira e Noções de Direito e Legislação.

O Bard alcançou 100% de acerto para Contabilidade Aplicada ao Setor Público, Legislação e Ética profissional e Noções de Direito e Legislação Aplicada, mas não acertou nenhuma questão em Língua Portuguesa Aplicada e Matemática Financeira e Estatística:

Tabela 13: Desempenho IA por disciplina após agrupamento, 1º exame de suficiência 2022

1º Exame Suficiência 2022	Total Geral	ChatGPT		GPT-4		Bard	
		Acertos	%	acertos	%	acertos	%
Auditoria e Perícia Contábil	4	3	75%	2	50%	2	50%
Contabilidade Aplicada ao Setor Público	3	2	67%	2	67%	3	100%
Contabilidade Geral	17	10	59%	13	76%	5	29%
Contabilidade Gerencial, Custos e Controladoria	7	3	43%	3	43%	1	14%
Legislação e Ética profissional	4	3	75%	3	75%	4	100%
Língua Portuguesa Aplicada	2	0	0%	1	50%	0	0%
Matemática Financeira e Estatística	2	2	100%	2	100%	0	0%
Noções de Direito e Legislação Aplicada	3	2	67%	3	100%	3	100%
Prin. de Cont. de Normas Brasileiras de Cont. e Teoria da Contabilidade	8	3	38%	5	63%	4	50%
Total Geral	50	28	56%	34	68%	22	44%

Fonte: A autora, 2023.

Na segunda edição do exame de suficiência, o ChatGPT não apresentou aproveitamento de 100% em nenhuma disciplina. Já o GPT-4 obteve desempenho de 100% nas seguintes disciplinas: Auditoria e Perícia Contábil, Legislação e Ética profissional e Matemática Financeira e Estatística. O Bard apresentou 100% de acerto em Matemática Financeira e Estatística e Perícia Contábil, conforme Tabela 14:

Tabela 14: Desempenho IA por disciplina após agrupamento, 2º exame de suficiência 2022

2º Exame Suficiência 2022	Total Geral	ChatGPT		GPT-4		Bard	
		acertos	%	acertos	%	acertos	%
Auditoria e Perícia Contábil	4	3	75%	4	100%	3	75%
Contabilidade Aplicada ao Setor Público	3	1	33%	1	33%	2	67%
Contabilidade Geral	17	5	29%	10	59%	8	47%
Contabilidade Gerencial, Custos e Controladoria	7	2	29%	6	86%	1	14%
Legislação e Ética profissional	4	1	25%	4	100%	1	25%
Língua Portuguesa Aplicada	2	1	50%	2	100%	1	50%
Matemática Financeira e Estatística	2	0	0%	2	100%	2	100%
Noções de Direito e Legislação Aplicada	3	2	67%	2	67%	2	67%
Prin. de Cont. de Normas Brasileiras de Cont. e Teoria da Contabilidade	8	4	50%	5	63%	5	63%
Total Geral	50	19	38%	36	72%	25	50%

Fonte: A autora, 2023.

Já as questões dos exames técnicos foram classificadas nos seguintes grupos de disciplinas: (i) Normas Profissionais, Técnicas, Legislação e ética Profissional; (ii) Legislação, Normas do Mercado de Capitais e Bovespa; (iii) Normas Brasileiras de Contabilidade (iv) Conhecimento de Operações da área de Instituições Reguladas pelo Banco Central (BCB); (v) Legislação e Normas Aplicáveis as sociedades supervisionas pela SUSEP; (vi) Legislação e Normas Aplicáveis as entidades supervisionadas pela PREVIC.

Pode-se notar na Tabela 15 que para os exames técnicos nenhum modelo alcançou 100% de acerto, bem como não se observou ausência de aproveitamento para nenhuma disciplina:

Tabela 15: Desempenho IA por disciplina após agrupamento, exames de qualificação técnica 2022

exames de qualificação técnica	Total Geral	ChatGPT		GPT-4		Bard	
		Acerto	%	certo	%	acerto	%
Conhecimento de Operações da Área de Instituições Reguladas pelo Banco Central do Brasil (BCB)	12	8	67%	8	67%	4	33%
Legislação e Normas Aplicáveis às Entidades Supervisionadas pela PREVIC	14	8	57%	10	71%	10	71%
Legislação e Normas Aplicáveis às Sociedades Supervisionadas pela SUSEP	18	6	33%	7	39%	7	39%
Legislação, Normas do Mercado de Capitais e BOVESPA	7	5	71%	5	71%	3	43%
Normas Brasileiras de Contabilidade	14	5	36%	9	64%	10	71%
Normas Profissionais, Técnicas, Legislação e Ética Profissional	85	57	67%	70	82%	54	64%
Total geral	150	89	59%	109	73%	88	59%

Fonte: A autora, 2023.

As Tabelas 13 e 14 apresentam muitas frequências menores que cinco, inclusive algumas frequências nulas. Com o intuito de evitar frequências esperadas reduzidas, as quais poderiam comprometer o teste qui-quadrado de homogeneidade, as disciplinas Legislação e Ética profissional, Língua Portuguesa Aplicada, Matemática Financeira e Estatística, Noções de Direito e Legislação Aplicada foram agrupadas na categoria outros, conforme apresentado na Tabela 16 contendo os totais consolidados de erros e acertos dos modelos de IA nos exames de suficiência em 2022. Ademais, em função da persistência de algumas frequências menores que 5, o p-valor do teste qui-quadrado foi calculado por simulação de Monte Carlo, ao invés da aproximação pela distribuição qui-quadrado (Verzani, 2004). Os totais consolidados de erros e acertos nos exames técnicos são apresentados na Tabela 17.

Tabela 16: Total de acertos e erros por disciplina nos exames de suficiência de 2022

Disciplinas	ChatGPT		GPT-4		Bard		Questões
	acertos	erros	Acertos	erros	acertos	erros	
Auditoria e Perícia Contábil	6	2	6	2	5	3	8
Contabilidade Aplicada ao Setor Público	3	3	3	3	5	1	6
Contabilidade Geral	15	19	23	11	13	21	34
Contabilidade Gerencial, Custos e Controladoria	5	9	9	5	2	12	14
Prin. de Cont. de Normas Brasileiras de Cont. e Teoria da Contabilidade	7	9	10	6	9	7	16
Outros	11	11	19	3	13	9	22
Total	47	53	70	30	47	53	100
Taxas de acertos/erros	47%	53%	70%	30%	47%	53%	
Estatística qui-quadrado χ_5^2	3,52		4,78		12,86		
p-valor	0,64		0,47		0,02**		

Nota: ** indica significância estatística no nível de 5%

Fonte: A autora, 2023.

As Tabelas 16 e 17 apresentam estatísticas qui-quadrado, acompanhadas dos p-valores associados. Estes resultados são obtidos a partir de testes qui-quadrado de homogeneidade aplicados às distribuições de frequências de acertos e erros de cada modelo de IA em seis grupos de disciplinas. Neste contexto, a hipótese nula (H_0) postula a consistência no desempenho global do modelo IA ao longo dos grupos de disciplinas, i.e., não se observam diferenças significativas das distribuições relativas de acertos/erros ao longo dos grupos de disciplinas. Em contrapartida, a hipótese alternativa (H_1) aponta para uma variação significativa das distribuições relativas de acertos/erros ao longo dos grupos de disciplinas (Talon et al., 2023).

Conforme indicado pelos p-valores na Tabela 16, tanto para o ChatGPT (p-valor 0,64) quanto para o GPT-4 (p-valor 0,47), os resultados indicam a ausência de evidências suficientes para rejeitar a hipótese nula de consistência nos resultados entre os grupos de disciplinas. Por outro lado, no caso do Bard (p-valor 0,02) o teste qui-quadrado aponta para a rejeição da hipótese nula, i.e., um desempenho significativamente variável do Bard ao longo dos 6 grupos de disciplinas, em especial no grupo formado por Contabilidade Gerencial, Custos e Controladoria, onde a taxa de acerto foi de apenas 14%, enquanto no conjunto de questões o Bard alcançou uma taxa de acerto de 47%.

Tabela 17: Total de acertos e erros por disciplina nos exames técnicos de 2022

Disciplinas	ChatGPT		GPT-4		Bard		Questões
	acertos	Erros	acertos	Erros	Acertos	erros	
Conhecimento de Operações da área de Instituições Reguladas pelo Banco Central (BCB)	8	4	8	4	4	8	12
Legislação e Normas Aplicáveis as entidades supervisionadas pela PREVIC	8	6	10	4	10	4	14
Legislação e Normas Aplicáveis às sociedades supervisionadas pela SUSEP	6	12	7	11	7	11	18
Legislação, Normas do Mercado de Capitais e Bovespa	5	2	5	2	3	4	7
Normas Brasileiras de Contabilidade	5	9	9	5	10	4	14
Normas Profissionais, Técnicas, Legislação e ética Profissional	57	28	70	15	54	31	85
Total	89	61	109	41	88	62	150
Taxas de acertos/erros	59%	41%	73%	27%	59%	41%	
Estatística qui-quadrado χ^2_5	11,10		15,08		9,51		
p-valor	0,049**		0,008***		0,089*		

Nota: ***, ** e * indicam significância estatística nos níveis de 1%, 5% e 10% respectivamente

Fonte: A autora, 2023.

No caso dos exames técnicos, os p-valores informados na Tabela 17 indicam a rejeição da hipótese nula nos níveis usuais de significância. Portanto, neste caso, os modelos de IA demonstraram um desempenho significativamente variável nos seis grupos de disciplinas, sobretudo em Legislação e Normas Aplicáveis às sociedades supervisionadas pela SUSEP, onde todos os modelos apresentaram taxas de erros maiores que as taxas de acertos, uma performance distinta do desempenho global dos modelos IA no conjunto de exames no qual a taxa de acertos supera a taxa de erros.

4.4 Análise: Questões discursivas QTG e CVM

Com o objetivo de avaliar a eficácia do ChatGPT, GPT-4 e Bard na resolução de questões discursivas, inicialmente foi desenvolvido um gabarito detalhado para as questões discursivas dos exames QTG e CVM, edição de 2022. Em seguida, essas questões foram aplicadas aos modelos de IA.

Conforme estipulado pelo edital dos exames, além do conhecimento técnico, a habilidade de elaboração e o desenvolvimento textual são critérios essenciais na atribuição de notas. Para avaliar esses aspectos de maneira objetiva e padronizada, foi utilizada a ferramenta Editor do Microsoft Word. Essa ferramenta analisa os textos sob várias perspectivas, incluindo

pontuação, formalidade da escrita, clareza, concisão, escolha vocabular, linguagem formal, ortografia, gramática, e similaridade com fontes online.

Todos os detalhes pertinentes, como a coleta de dados, o gabarito elaborado, as avaliações textuais realizadas pelo Microsoft Word, bem como as notas atribuídas e suas justificativas, estão documentados no Apêndice D. Conforme os resultados apresentados na Tabela 18, os modelos ChatGPT e GPT-4 conseguiram desempenho suficientes nas questões discursivas para aprovação nas duas provas, já o Bard conseguiria ser aprovado apenas no exame QGT.

Tabela 18: Questões discursivas, QGT e CVM 2022

	QGT	Resultado	CVM	Resultado
ChatGPT	60%	Aprovado	80%	Aprovado
GPT-4	60%	Aprovado	80%	Aprovado
Bard	60%	Aprovado	40%	Reprovado

Fonte: A autora, 2023.

CONCLUSÕES

O ChatGPT obteve aprovação em apenas uma das edições do exame de suficiência de 2022 e o GPT-4 foi aprovado em ambas as edições. Já o Bard foi aprovado apenas na 2ª edição do exame de suficiência.

A análise consolidada dos resultados por modelo de IA mostra que o GPT-4 se sobressai com uma taxa de acerto de 70%, mais que o dobro da sua taxa de erros. Por outro lado, o ChatGPT e o Bard apresentaram taxas de acertos e erros iguais, refletindo uma capacidade inferior em comparação ao GPT-4. Além disso, o teste exato de Fisher corroborou a superioridade significativa do desempenho do GPT-4 sobre o ChatGPT e o Bard, com um p-valor de 0,002.

Assim, observa-se que o modelo GPT-4, além de ser aprovado nas duas edições do exame de suficiência de 2022, teve uma média de desempenho superior a todas as regiões e estados do Brasil.

No exame técnico, o GPT-4 seria aprovado em cinco exames de qualificação técnica aplicados pelo CFC em 2022, reprovando apenas no SUSEP. Ao passo que o ChatGPT seria reprovado apenas no BCB e SUSEP, e o Bard seria reprovado nos exames de CMV, BCB e SUSEP.

Com resultados de aprovação, o modelo GPT-4 demonstrou um desempenho notavelmente superior em relação a outros modelos. Em 2022, alcançou 100% de aprovação nos exames de suficiência e 83% nos exames técnicos.

Com relação ao desempenho da IA versus desempenho humano, observou-se na primeira edição do exame de suficiência de 2022, uma performance notável dos modelos ChatGPT e GPT-4, que superaram as médias de notas humanas em todas as regiões e estados do Brasil. Enquanto o Bard atingiu apenas 22 acertos, abaixo do limiar de 50% necessário para a aprovação, este resultado foi equivalente à média das performances humanas nos estados. O ChatGPT obteve um desempenho ligeiramente superior, com 28 acertos e uma taxa de acerto de aproximadamente 56%, ultrapassando a média humana e atingindo o critério de aprovação. Destacando-se entre os modelos, o GPT-4 alcançou o melhor resultado com 34 acertos, excedendo as médias estaduais e demonstrando um desempenho excepcional.

Na segunda edição do exame de suficiência de 2022, o GPT-4 e o Bard melhoraram ligeiramente seus desempenhos em comparação com a primeira edição, conseguindo notas acima das médias humanas em todas as regiões e estados do Brasil. No entanto, o ChatGPT não teve o mesmo sucesso, alcançando apenas 19 acertos, o que corresponde à menor média estadual e não é suficiente para a aprovação, ficando abaixo do limiar de 50% de acerto.

Em 2022, o modelo GPT-4 destacou-se ao alcançar 100% de aprovação nos exames de suficiência e 83% nos exames técnicos, superando outros modelos e apresentando desempenho superior em todas as regiões e estados do Brasil. Esses resultados são consistentes com o estudo de Eulerich et al. (2023^a), que também evidenciou a superioridade do GPT-4, registrando uma média de acertos de 85,1% em todas as seções dos exames.

Vale ressaltar que não foi identificado no *site* do CFC dados estatísticos com desempenho humano para os exames de qualificação técnica, sendo assim não foi possível avaliar o desempenho comparativo de humanos *versus* IA por disciplina. Além disso, ressaltou-se que foi avaliada apenas a aprovação (60% de acerto) dos modelos de IA para os exames técnicos das questões objetivas dos exames, os gabaritos das questões discursivas não foram divulgados no *site*.

Sobre as características capazes de influenciar a taxa de acerto da IA, notou-se na análise descritiva que na primeira edição do exame de suficiência de 2022 a presença de tabelas no enunciado das questões não prejudicou o desempenho do ChatGPT, no entanto, prejudicou o desempenho do GPT-4 e do Bard. Na segunda edição de 2022, observou-se que a presença de tabelas no enunciado prejudicou o desempenho do ChatGPT e o Bard.

A análise dos valores consolidados das taxas de acertos e erros dos modelos de IA (ChatGPT, Bard e GPT-4) em questões com e sem tabelas nos exames do CFC revelou tendências distintas. Foi constatado que a presença de tabelas nas questões resulta em uma taxa de acertos menor para todos os modelos. Especificamente, tanto o ChatGPT quanto o Bard apresentaram taxas de erros acima de 60% em questões com tabelas, enquanto o GPT-4 mostrou um desempenho relativamente melhor com uma taxa de acertos de 57%.

Por outro lado, nas questões sem tabelas, o ChatGPT e o Bard exibiram taxas de acertos e de erros equilibradas, indicando uma distribuição uniforme de respostas corretas e erradas. Em contraste, o GPT-4 destacou-se com uma predominância de respostas corretas, atingindo uma taxa de acertos de 75% nas questões sem tabelas.

Além disso, a análise estatística revelou uma significância marginal para o GPT-4 (p-valor de 0,068), sugerindo que o desempenho deste modelo em questões sem tabelas é

significativamente superior ao seu desempenho em questões com tabelas. Este achado ressalta a habilidade diferenciada do GPT-4 em lidar com questões sem elementos visuais complexos, como tabelas, e enfatiza a importância de considerar esses elementos ao avaliar a eficácia dos modelos de IA em contextos de teste específicos.

As estatísticas resultantes do teste qui-quadrado de homogeneidade aplicados às distribuições de frequências de acertos e erros de cada modelo de IA em seis grupos de disciplinas dos exames de suficiência 2022, revelaram que tanto para o ChatGPT (p-valor de 0,64) quanto para o GPT-4 (p-valor de 0,47), não há evidências suficientes para refutar a hipótese nula de homogeneidade das distribuições relativas de acertos/erros ao longo dos grupos de disciplinas, o que aponta para a consistência dos desempenhos do ChatGPT e GPT-4 nas diversas áreas avaliadas nos exames de suficiência.

Por outro lado, para o modelo Bard (p-valor de 0,02), o teste qui-quadrado indicou uma rejeição da hipótese nula, demonstrando um desempenho significativamente variável entre os seis grupos de disciplinas. Essa variabilidade foi particularmente notável no grupo de Contabilidade Gerencial, Custos e Controladoria, onde o Bard apresentou uma taxa de acerto de apenas 14%, contrastando com uma taxa de acerto de 47% no conjunto geral de questões. Essa diferença sugere que o modelo Bard possui limitações específicas em determinadas áreas do conhecimento, destacando a necessidade de uma abordagem mais especializada ou aprimorada para esses temas.

No caso dos exames técnicos, os resultados apontam para a rejeição da hipótese nula nos níveis convencionais de significância, evidenciando que os modelos de IA exibiram variações significativas de desempenho nos seis grupos de disciplinas analisadas. Notavelmente, na área de Legislação e Normas Aplicáveis às sociedades supervisionadas pela SUSEP, todos os modelos de IA apresentaram taxas de erros superiores às taxas de acertos. Este achado destaca-se do desempenho global dos modelos de IA nos exames como um todo, onde, de forma geral, as taxas de acertos foram maiores do que as taxas de erros. Assim, a pesquisa demonstra que, embora os modelos de IA tenham mostrado eficácia no conjunto dos exames, existem variações significativas em seu desempenho quando analisadas disciplinas específicas, evidenciando a necessidade de aperfeiçoamentos focados em áreas de conhecimento particulares.

Foi solicitado ao CFC dados estatísticos fundamentais (Apêndice B) relacionados às provas administradas pelo CFC no ano de 2022: (i) Quantidade de candidatos que realizaram cada prova; (ii) Quantidade de candidatos que acertaram cada questão (ou o percentual de

acertos por questão); (iii) O tipo de conteúdo avaliado em cada questão (por exemplo: Auditoria Contábil, Contabilidade Aplicada ao Setor Público, Contabilidade de Custos, etc.); (iv) Quais são essas abordagens (teoria ou critério de taxonomia) para cada prova aplicada em 2022?; (v) Qual a classificação efetiva de cada questão das provas de 2022 (Suficiência 1ª edição, Suficiência 2ª edição, CVM, BACEN, SUSEP, PREVIC e Perícia)?

A resposta recebida sobre a solicitação de abertura de dados (Apêndice C) não atendeu de forma satisfatória às necessidades deste estudo. Quanto à primeira solicitação, que se referia ao total de candidatos que prestaram o exame, o CFC forneceu informações apenas para os exames técnicos. No entanto, em relação à segunda solicitação, sobre o percentual de acerto por questão, foi informado que tal dado não existe. Em relação ao terceiro questionamento, o CFC esclareceu que não possui relatórios segmentados por conteúdo específico.

Sobre as duas últimas solicitações, o CFC elucidou que o Exame de Suficiência visa comprovar a aquisição de conhecimentos medianos alinhados aos conteúdos programáticos do curso de Bacharelado em Ciências Contábeis. Portanto, sua dificuldade é considerada mediana. Em contrapartida, o Exame de Qualificação Técnica (EQT) para Auditor e Perito Contábil busca avaliar o nível de conhecimento e a competência técnico-profissional de contadores que desejam atuar em auditoria e perícia contábil, demandando deles um profundo entendimento técnico e científico na área. Foi ainda ressaltado que não existe previsão contratual para interação com os elaboradores dos exames, devido ao sigilo e integridade que o processo exige.

Após a exposição da resposta do CFC, ressalta-se como intenção a continuidade dessa investigação com objetivo de compreender se as características solicitadas influenciam a taxa de acerto da IA são as mesmas que influenciam a taxa de acerto por humanos, com a obtenção dos dados seria possível compreender se o CFC / Consulpan utilizam alguma teoria ou critério de taxonomia das questões, como, por exemplo, a categorização por nível de dificuldade (fácil / média / difícil) ou por nível de complexidade, ou ainda, se levam em conta as competências avaliadas; e se utilizam a “teoria de resposta ao item”.

A aprovação da IA em exames de certificação profissional da área contábil poderia gerar uma série de impactos e reflexões. Primeiramente vale mencionar uma possível desvalorização da certificação, considerando que se uma IA pode passar no exame, qual é o valor real da certificação para os profissionais humanos? Isso pode levar a uma reavaliação do que realmente significa ter o registro profissional do CFC e quais habilidades e competências são realmente avaliadas nos exames. No mercado de trabalho, poderia ocorrer uma possível redução na

demanda para alguns trabalhos contábeis, uma vez que algumas atividades da área poderiam ser automatizadas, ocasionando reduções significativas nos custos.

No que se refere à formação e educação dos profissionais, as instituições de ensino superior em contabilidade podem precisar adaptar seus currículos à nova realidade tecnológica. Além disso, se o humano delegar à IA determinadas tarefas inerentes ao profissional da contabilidade, surge a questão da responsabilidade caso uma IA cometa um erro ou uma infração ao Código de Ética da profissão.

Embora não existam leis específicas para o uso da IA no campo das ciências contábeis, cresce a necessidade de pesquisas que busquem a compreensão dos aspectos éticos, limitações e o potencial impacto dessas tecnologias no campo. Pesquisas futuras podem se concentrar em abordar esses desafios e explorar como a IA pode ser integrada de forma eficaz e ética nas operações de contabilidade.

Pesquisas futuras podem examinar o desempenho das IA no exame Nacional de Desempenho dos Estudantes (Enade) na graduação em ciências contábeis. O Enade tem como proposta avaliar o rendimento dos concluintes dos cursos de graduação em relação aos conteúdos previstos nas diretrizes curriculares dos cursos, o desenvolvimento de competências e habilidades necessárias ao aprofundamento da formação geral e profissional, sendo uma investigação necessária para gerar reflexões sobre a relação da qualidade do ensino contábil nas Universidades.

Além disso, é necessário que novos estudos investiguem o desempenho da IA nas questões discursivas propostas dos exames de qualificação técnica BCB (Banco Central do Brasil), SUSEP (Superintendência de Seguros Privados), PREVIC (Superintendência Nacional de Previdência Complementar) e Perícia Contábil.

REFERÊNCIAS

BORJI, Ali; MOHAMMADIAN, Mehrdad. **Battle of the Wordsmiths: Comparing ChatGPT, GPT-4, Claude, and Bard**, 2023. DOI: <http://dx.doi.org/10.2139/ssrn.4476855>. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4476855. Acesso em: 30 set. 2023.

BUNKER, R. B.; CAGLE, C. S. Does the Sequencing of Exam Sections Impact Candidate Success on the CPA Exam?. **Journal of Accounting and Finance**, [s. l.], v. 21, n. 3, 2021. DOI: <https://doi.org/10.33423/jaf.v21i3.4401>. Disponível em: <https://articlegateway.com/index.php/JAF/article/view/4401>. Acesso em: 10 ago. 2023.

CHEN, X. *et al.* **Artificial intelligence's capabilities, limitations, and impact on accounting education: Investigating ChatGPT's performance on educational accounting cases**, 2023. DOI: <http://dx.doi.org/10.2139/ssrn.4431202>. Disponível em: <https://ssrn.com/abstract=4431202>. Acesso em: 8 ago. 2023.

CONSELHO FEDERAL DE CONTABILIDADE (Brasil). **Resolução CFC n. 853**. Institui o Exame de Suficiência como requisito para obtenção de Registro Profissional em CRC. Brasília, DF: CFC, 28 jul. 1999.

CONSELHO FEDERAL DE CONTABILIDADE (Brasil). **Resolução CFC n. 1486**. Regulamenta o Exame de Suficiência como requisito para obtenção de Registro Profissional em Conselho Regional de Contabilidade (CRC). Brasília, DF: CFC, 15 maio 2015.

CORDIS, Adriana S.; MUZATKO, Steven. **Higher Education Spending and CPA Exam Performance**, 2020. DOI: <http://dx.doi.org/10.2139/ssrn.3701060>. Disponível em: <https://ssrn.com/abstract=3701060>. Acesso em: 8 ago. 2023.

ESPAHBODI, A. *et al.* Determinants of CPA exam performance, *Journal of Accounting Education*. **Journal of Accounting Education**, [s. l.], v. 64, 2023. DOI: <https://doi.org/10.1016/j.jaccedu.2023.100859>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0748575123000313>. Acesso em: 5 ago. 2023.

EULERICH, M. *et al.* **Can Artificial Intelligence Pass Accounting Certification Exams? ChatGPT: CPA, CMA, CIA, and EA?**, 2023. DOI: <http://dx.doi.org/10.2139/ssrn.4452175>. Disponível em: <https://ssrn.com/abstract=4452175>. Acesso em: 8 ago. 2023.

EULERICH, M. *et al.* **The Impact of Audit Technology on Audit Task Outcomes: Evidence for Technology-Based Audit Techniques**, 2022. DOI: <http://dx.doi.org/10.2139/ssrn.3444119>. Disponível em: <https://ssrn.com/abstract=3444119>. Acesso em: 5 ago. 2023.

LARSON, H.J. **Introduction to Probability Theory and Statistical Inference**. 3. ed. Toronto: John Wiley e Sons, 1982.

MELLO, M. P., PETERNELLI, L. A. **Conhecendo o R: uma visão mais que estatística**. Viçosa, MG: UFV, 2013.

MIRANDA, C. de S.; ARAÚJO, A. M. P. de; MIRANDA, R. A. de M. O exame de suficiência em contabilidade: uma avaliação sob a perspectiva dos pesquisadores. **Revista Ambiente Contábil** - Universidade Federal Do Rio Grande Do Norte - ISSN 2176-9036, [s. l.], v. 9, n. 2, p. 158–178, 2017. DOI: 10.21680/2176-9036.2017v9n2ID10760. Disponível em: <https://periodicos.ufrn.br/ambiente/article/view/10760>. Acesso em: 8 set. 2023.

O'LEARY, D.E.O. An analysis of Watson vs. BARD vs. ChatGPT: The Jeopardy! Challenge. **AI Magazine**, [s. l.], v. 44, n. 3, 2023. DOI: <https://doi.org/10.1002/aaai.12118>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/aaai.12118>. Acesso em: 8 dez. 2023.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2022. Disponível em: <https://www.R-project.org/>. Acesso em: 15 set. 2023.

SILVA, J. V. da; DURIGON, A. R.; MATTIELLO DA SILVA, J. V. V.; SANTOS, R. dos. O Exame de Suficiência na percepção dos alunos de Ciências Contábeis. **Revista Catarinense da Ciência Contábil**, [s. l.], v. 19, 2020. DOI: 10.16930/2237-766220202952. Disponível em: <https://revista.crcsc.org.br/index.php/CRCSC/article/view/2952>. Acesso em: 8 set. 2023.

TALONI, A. *et al.* Comparative performance of humans versus GPT-4.0 and GPT-3.5 in the self-assessment program of American Academy of Ophthalmology. **Scientific Reports**, [s. l.], v. 12, 18562, 2023. DOI: <https://doi.org/10.1038/s41598-023-45837-2>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37899405/>. Acesso em: 8 dez. 2023.

TOYAMA, Y. *et al.* Performance evaluation of ChatGPT, GPT-4, and Bard on the official board examination of the Japan Radiology Society. **Japanese Journal of Radiology**, 2023. DOI: <https://doi.org/10.1007/s11604-023-01491-2>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37792149/>. Acesso em: 8 dez. 2023.

UPTON G, COOK I. **A dictionary of statistics**. 2 ed. Oxford: Oxford University Press, 2008.

VERZANI, J. **Using R for introductory statistics**. [s. l.]: Chapman and Hall/CRC, 2004.

APÊNDICE A – Links ChatGPT, GPT-4 e Bard: Exames Suficiência 2022 e Exames Técnicos 2022

Exames	IA	Link
1º exame suficiência CRC 2022	ChatGPT	https://chat.openai.com/share/a3c3178c-2e37-4e74-b60a-9aa341c65271
1º exame suficiência CRC 2022	GPT-4	https://chat.openai.com/share/6708b904-13b4-45d0-8614-b4bffa2293ff
1º exame suficiência CRC 2022	Bard	https://g.co/bard/share/f1987b09acfe
2º exame suficiência CRC 2022	ChatGPT	https://chat.openai.com/share/f1aa8538-aecc-4e55-9309-ec48312ba1ea
2º exame suficiência CRC 2022	GPT-4	https://chat.openai.com/share/5f113812-eb9b-4a1a-8e59-0ee3d54af86e
2º exame suficiência CRC 2022	Bard	https://g.co/bard/share/82bcdb6a29a5
QTG 2022	ChatGPT	https://chat.openai.com/share/712b2e5b-ca98-4264-a0a5-801e4b3d469f
QTG 2022	GPT-4	https://chat.openai.com/share/a0817c0c-6380-4703-b4b2-54cf27d6759f
QTG 2022	Bard	https://g.co/bard/share/b3182dc5143b
CVM 2022	ChatGPT	https://chat.openai.com/share/207817c3-0fab-4310-9fae-64dee2ef2504
CVM 2022	GPT-4	https://chat.openai.com/share/66601163-c805-439f-bfcd-2330be2060cd
CVM 2022	Bard	https://g.co/bard/share/98b230574e2c
BCB 2022	ChatGPT	https://chat.openai.com/share/99ce43aa-36e4-4215-b287-c2c8deda82cc
BCB 2022	GPT-4	https://chat.openai.com/share/f9534ab2-a6ac-4827-9d0e-51ee277794b3
BCB 2022	Bard	https://g.co/bard/share/8ca39a7cd53b
SUSEP2022	ChatGPT	https://chat.openai.com/share/c23747ed-ac65-4ebc-92e1-c6eca20b51f3
SUSEP2022	GPT-4	https://chat.openai.com/share/f4e4911b-cff0-435c-bd80-e82e9c9bb749
SUSEP2022	Bard	https://g.co/bard/share/8ca39a7cd53b
PREVIC 2022	ChatGPT	https://chat.openai.com/share/34745a2e-97b7-40bc-925f-dab800578763

PREVIC 2022	GPT-4	https://chat.openai.com/share/4d5df619-48ad-4780-a871-4578837d69e0
PREVIC2022	Bard	https://g.co/bard/share/ad92d0876b2a
Perito 2022	ChatGPT	https://chat.openai.com/share/5db8506a-3f33-4c75-a648-45df6ce92bda
Perito 2022	GPT-4	https://chat.openai.com/share/6ca53056-46d7-4d33-b06e-29808a959941
Perito 2022	Bard	https://g.co/bard/share/21fe9c69b1fc

APÊNDICE B – Solicitação Abertura de dados Exames CFC 2022

<https://ouvidoria.cfc.org.br/visao/consulta.frm.php?codigo=1®istro=SZ1693841383S64f5f7e7be427>



Consulte o andamento da sua Manifestação	
Inscrição nº do registro aqui:	<input type="text" value="SZ1693841383S64f5f7e7be427"/> Consultar
Nome do manifestante:	Ana Clara de Melo Reis
Telefone:	21995521350
Email:	anaclaramreis@gmail.com
Data de envio da manifestação:	04/09/2023
Prazo para resposta:	04/10/2023
Data de envio da resposta:	Não informada.
Assunto:	Solicitação de abertura de dados - Pesquisa Acadêmica
Conteúdo da manifestação:	<p>Prezados, Atualmente, como aluna do Programa de Mestrado em Ciências Contábeis da UERJ, estou conduzindo uma pesquisa abrangente acerca dos exames de suficiência e qualificação técnica. Tal pesquisa está sendo desenvolvida sob a orientação do Prof. Ricardo Lopes, com o propósito de aprofundar nossa compreensão nesse campo crucial. Nesse contexto, estamos buscando obter dados fundamentais relacionados às provas administradas pelo CFC no ano de 2022. Especificamente, nossa atenção está voltada para as edições seguintes: Suficiência 1ª edição, Suficiência 2ª edição, bem como os exames da CVM, BACEN, SUSEP, PREVIC e Perícia. A obtenção dessas informações é de vital importância para a continuidade de nossa pesquisa. Os dados que precisamos são os seguintes: 1) Quantidade de candidatos que realizaram cada prova. 2) Quantidade de candidatos que acertaram cada questão (ou o percentual de acertos por questão). * 3) O tipo de conteúdo avaliado em cada questão (por exemplo: Auditoria Contábil, Contabilidade Aplicada ao Setor Público, Contabilidade de Custos, etc.). * 4) Observo que essas informações não estão disponibilizadas no site do CFC. Lá há informação sobre a taxa de acerto por conteúdo, mas preciso da taxa de acerto por questão (item 2 acima) e o conteúdo coberto em cada questão (item 3 acima). Além disso, tenho um grande interesse em conversar com os responsáveis pela elaboração dessas provas. O objetivo é compreender se o CFC / Consulpan utilizam alguma teoria ou critério de taxonomia das questões, como, por exemplo, a categorização por nível de dificuldade (fácil / média / difícil) ou por nível de complexidade, ou ainda, se levam em conta as competências avaliadas; e se utilizam a "teoria de resposta ao item". Caso o CFC / Consulpan adote alguma teoria ou critério de taxonomia, seria de imenso valor conhecer em detalhes: 4) Quais são essas abordagens (teoria ou critério de taxonomia) para cada prova aplicada em 2022? 5) Qual a classificação efetiva de cada questão das provas de 2022 (Suficiência 1ª edição, Suficiência 2ª edição, CVM, BACEN, SUSEP, PREVIC e Perícia)? A colaboração que você pode oferecer nesse assunto será fundamental para o progresso da pesquisa. Desde já, agradeço imensamente pela atenção dispensada.</p>
Setores por onde a manifestação passou:	CFC · Ouvidoria, CFC · DEPEX
Respostas/Resolução:	Sua manifestação ainda não teve uma resposta final.
Arquivos anexos:	Não existem arquivos anexos.

APÊNDICE C – Resposta da Solicitação de Abertura de dados Exames CFC 2022

Consulte o andamento da sua Manifestação

Insira nº do registro aqui:
S21693841383564597e7be427 Consultar

Nome do manifestante: Ana Clara de Melo Reis
Telefone: 21985521350
Email: anaclaramreis@gmail.com
Data de envio da manifestação: 04/09/2023
Prazo para resposta: ~~04/10/2023~~ 03/11/2023 (prazo prorrogado)
Data de envio da resposta: 29/09/2023
Assunto: Solicitação de abertura de dados - Pesquisa Acadêmica

Conteúdo da manifestação:
Prezados, Atualmente, como aluna do Programa de Mestrado em Ciências Contábeis da UERJ, estou conduzindo uma pesquisa abrangente acerca dos exames de suficiência e qualificação técnica. Tal pesquisa está sendo desenvolvida sob a orientação do Prof. Ricardo Lopes, com o propósito de aprofundar nossa compreensão nesse campo crucial. Nesse contexto, estamos buscando obter dados fundamentais relacionados às provas administradas pelo CFC no ano de 2022. Especificamente, nossa atenção está voltada para as edições seguintes: Suficiência 1ª edição, Suficiência 2ª edição, bem como os exames da CVM, BACEN, SUSEP, PREVIC e PERICIA. A obtenção dessas informações é de vital importância para a continuidade de nossa pesquisa. Os dados que precisamos são os seguintes: 1) Quantidade de candidatos que realizaram cada prova. 2) Quantidade de candidatos que acertaram cada questão (ou o percentual de acertos por questão). * 3) O tipo de conteúdo avaliado em cada questão (por exemplo: Auditoria Contábil, Contabilidade Aplicada ao Setor Público, Contabilidade de Custos, etc.). * * Observe que essas informações não estão disponibilizadas no site do CFC. Lá há informação sobre a taxa de acerto por conteúdo, mas preciso da taxa de acerto por questão (item 2 acima) e o conteúdo coberto em cada questão (item 3 acima). Além disso, tenho um grande interesse em conversar com os responsáveis pela elaboração dessas provas. O objetivo é compreender se o CFC / Consulplan utilizam alguma teoria ou critério de taxonomia das questões, como, por exemplo, a categorização por nível de dificuldade (fácil / média / difícil) ou por nível de complexidade, ou ainda, se levam em conta as competências avaliadas e se utilizam a "teoria de resposta ao item". Caso o CFC / Consulplan adote alguma teoria ou critério de taxonomia, seria de imenso valor conhecer em detalhes: 4) Quais são essas abordagens (teoria ou critério de taxonomia) para cada prova aplicada em 2022? 5) Qual a classificação efetiva de cada questão das provas de 2022 (Suficiência 1ª edição, Suficiência 2ª edição, CVM, BACEN, SUSEP, PREVIC e Perícia)? A colaboração que você pode oferecer nesse assunto será fundamental para o progresso da pesquisa. Desde já, agradeço imensamente pela atenção dispensada.

Setores por onde a manifestação passou: CFC - Ouvidoria, CFC - DEPEX, CFC - DEPEX, CFC - DEPEX 

Respostas/Resolução: Em virtude do volume de dados solicitados, o prazo para resposta será prorrogado, seguindo as diretrizes da Lei n.º 13.460/2017 e Resolução CFC n.º 1.544/2018.

Respostas/Resolução: Em virtude do volume de dados solicitados, o prazo para resposta será prorrogado, seguindo as diretrizes da Lei n.º 13.460/2017 e Resolução CFC n.º 1.544/2018.

Senhora Ana, bom dia!

Em atenção a sua manifestação, seguem respostas aos questionamentos:

1. Quantidade de candidatos que realizaram cada prova.

Exame de Suficiência: Informação disponível no site do CFC, nos relatórios estatísticos (<https://cfc.org.br/registro/exame-de-suficiencia/relatorios-estatisticos-do-exame-de-suficiencia/>).

Exame de Qualificação Técnica: Seguem os dados na tabela abaixo (Edição 2022)

PROVA	INSCRITOS	PRESENTES
CVM	951	805
BCB	156	122
SUSEP	81	67
PREVIC	102	77
PERÍCIA	1049	744

2. Quantidade de candidatos que acertaram cada questão (ou o percentual de acertos por questão):

Informamos que não temos relatório específico quanto à quantidade de acertos, por questão, apenas, o percentual de acertos e erros, por conteúdo, disponível em: <https://cfc.org.br/registro/exame-de-suficiencia/relatorios-estatisticos-do-exame-de-suficiencia/>.

3. O tipo de conteúdo avaliado em cada questão (por exemplo: Auditoria Contábil, Contabilidade Aplicada ao Setor Público, Contabilidade de Custos, etc.).

Informamos que não dispomos de relatório específico, por conteúdo. Todas as provas anteriores, do Exame de Suficiência e do Exame de Qualificação Técnica, podem ser acessadas no site do CFC, bem como nos sites das bancas Consulplan e Fundação Cesgranrio, respectivamente, podendo auxiliar no levantamento dos conteúdos abordados.

Lembramos que o conteúdo cobrado nas provas está previsto nas seguintes normas:

a) **Resolução CFC n.º 1.486/2015**, que regulamenta o Exame de Suficiência como requisito para obtenção de Registro Profissional em Conselho Regional de Contabilidade (CRC);

b) **Norma Brasileira de Contabilidade, NBC PP 02**, de 21 de outubro de 2016, que dispõe sobre o exame de qualificação técnica para perito contábil;

c) **Norma Brasileira de Contabilidade, NBC PA 13 (R3)**, de 13 de agosto de 2020, que dá nova redação à NBC PA 13 (R2), que dispõe sobre o Exame de Qualificação Técnica para Auditor.

Quanto aos demais questionamentos, ressaltamos que o Exame de Suficiência é a prova de equalização destinada a comprovar a obtenção de conhecimentos médios, consoante os conteúdos programáticos desenvolvidos no curso de Bacharelado em Ciências Contábeis. Diante do exposto, informamos que o nível de dificuldade ou complexidade da prova é considerado médio.

Já o Exame de Qualificação Técnica (EQT) para Auditor e perito contábil tem por objetivo aferir o nível de conhecimento e a competência técnico-profissional necessários aos contadores (profissionais registrados) que pretendem atuar na atividade de Auditoria e perícia contábil. Para estes profissionais, é esperado elevado conhecimento técnico e científico do contador na respectiva área.

Ademais, informamos que não há previsão contratual para contato com os responsáveis pela elaboração das provas, dado o sigilo e a lisura que o processo requer.

Atenciosamente,
Ouvidoria do CFC

Arquivos anexos: Não existem arquivos anexos.

APÊNDICE D – Análise questões discursivas exames técnicos QGT e CVM 2022

QUALIFICAÇÃO TÉCNICA GERAL (QTG) - Discursivas

Respostas dos modelos de AI Questão 26 - QTG:

ChatGPT	Valor Questão:25	Nota: 25	GPT-4	Valor Questão:25	Nota: 25	Bard	Valor Questão:25	Nota: 25
<p>https://chat.openai.com/c/e6ce97eb-9b24-4b34-9d27-f08c64c89daf</p> <p>Um planejamento adequado de auditoria, conforme estabelecido pela NBC TA 300 (R1)/2016, oferece diversos benefícios cruciais para a condução eficaz da auditoria das demonstrações contábeis. Dois desses benefícios fundamentais são:</p> <p>Eficiência na Alocação de Recursos: O planejamento permite que o auditor identifique os riscos significativos e as áreas de alto risco nas demonstrações contábeis da entidade auditada. Isso possibilita a alocação adequada de recursos, como tempo e equipe, para focar nas áreas mais críticas, maximizando assim a eficiência do trabalho de auditoria.</p> <p>Redução de Riscos e Erros: O planejamento adequado ajuda o auditor a antecipar problemas potenciais e a desenvolver estratégias para mitigar riscos. Isso inclui a identificação de controles internos insuficientes e a formulação de procedimentos de auditoria apropriados para abordar essas deficiências. Consequentemente, o planejamento contribui para reduzir a probabilidade de erros materiais não detectados nas demonstrações contábeis.</p> <p>Para definir a estratégia global de um planejamento de auditoria, o auditor deve realizar as seguintes quatro ações de acordo com a NBC TA 300 (R1)/2016:</p> <p>Entendimento da Entidade e de seu Ambiente: O auditor deve adquirir um profundo conhecimento da entidade auditada e do ambiente em que opera. Isso envolve entender a estrutura organizacional, a indústria em que atua, os principais processos de negócios e os riscos específicos associados. Essa</p>			<p>https://chat.openai.com/c/c19447f2-0f20-4fa6-bb03-fd793f92e0f5</p> <p>O planejamento adequado da auditoria, conforme estipulado pela NBC TA 300 (R1)/2016, é essencial para garantir a eficácia e eficiência dos procedimentos de auditoria e para assegurar que os objetivos da auditoria sejam alcançados. Dois benefícios decorrentes de um planejamento adequado de auditoria:</p> <p>Eficiência na condução dos trabalhos: Um planejamento adequado permite que o auditor identifique previamente os principais riscos de auditoria e áreas críticas. Ao concentrar esforços nessas áreas e ao estabelecer de antemão as técnicas e procedimentos a serem utilizados, o auditor otimiza o tempo e os recursos, tornando o processo de auditoria mais eficiente.</p> <p>Redução de surpresas indesejadas: Ao planejar de forma apropriada, o auditor pode antecipar possíveis problemas ou obstáculos que possam surgir durante a execução da auditoria. Esta antecipação ajuda na preparação e na pronta resposta às questões, minimizando a possibilidade de ocorrências não previstas que possam comprometer o resultado da auditoria.</p> <p>Quatro ações que devem estar presentes na definição da estratégia global de um planejamento de auditoria:</p> <p>Entendimento da entidade e seu ambiente: O auditor deve obter um entendimento completo da entidade, do seu setor de atuação, dos fatores regulatórios, operacionais e financeiros que a afetam. Este entendimento é crucial para identificar os riscos associados e planejar os procedimentos de auditoria.</p> <p>Definição dos recursos necessários: A estratégia global deve identificar os recursos (como tempo, mão-de-obra e</p>			<p>https://g.co/bard/share/c1edde3f4805</p> <p>O planejamento de auditoria é uma etapa fundamental do processo de auditoria, pois é responsável por orientar e direcionar todo o trabalho a ser realizado pelo auditor. Um planejamento adequado, realizado de acordo com a NBC TA 300 (R1)/2016, traz diversos benefícios, dentre os quais se destacam:</p> <p>Redução do risco de auditoria: O planejamento permite que o auditor identifique e avalie os riscos de distorções relevantes nas demonstrações contábeis, bem como que desenvolva procedimentos de auditoria que sejam apropriados para reduzir esses riscos a um nível aceitável.</p> <p>Eficiência e eficácia do trabalho: O planejamento ajuda o auditor a organizar e coordenar o trabalho da equipe de auditoria, evitando desperdícios de tempo e recursos.</p> <p>Comunicação com a administração e com os responsáveis pela governança: O planejamento permite que o auditor se comunique com a administração e com os responsáveis pela governança sobre o escopo do trabalho, os procedimentos a serem realizados e os resultados esperados.</p> <p>A estratégia global de um planejamento de auditoria é um documento que define as principais ações a serem realizadas pelo auditor ao longo do trabalho. As quatro ações que devem estar presentes na definição da estratégia global são:</p> <p>Definição do escopo do trabalho: O auditor deve definir o escopo do trabalho de auditoria, incluindo os objetivos da auditoria, os procedimentos a serem realizados e os prazos para a conclusão do trabalho.</p> <p>Identificação e avaliação dos riscos de distorções relevantes: O auditor deve identificar e avaliar os riscos de distorções relevantes nas demonstrações contábeis, tanto de natureza inerente quanto de controle.</p>		

<p>compreensão é essencial para identificar os riscos de distorção relevante nas demonstrações contábeis.</p> <p>Identificação e Avaliação de Riscos de Distorção Relevante: Com base no entendimento adquirido, o auditor deve identificar os riscos de distorção relevante nas demonstrações contábeis. Isso inclui a avaliação dos riscos inerentes e de controle associados a cada área significativa das demonstrações contábeis, bem como a estimativa da magnitude desses riscos.</p> <p>Estabelecimento de Objetivos Globais de Auditoria: Com base na identificação e avaliação dos riscos, o auditor deve estabelecer objetivos globais de auditoria que orientarão o trabalho de auditoria. Esses objetivos ajudarão o auditor a determinar o escopo e a natureza dos procedimentos de auditoria a serem aplicados em resposta aos riscos identificados.</p> <p>Desenvolvimento de um Plano de Auditoria: O auditor deve elaborar um plano de auditoria que detalhe as ações específicas a serem realizadas para atingir os objetivos globais de auditoria. Isso inclui a seleção de procedimentos de auditoria, a alocação de recursos, o cronograma de trabalho e a documentação necessária. O plano de auditoria serve como um guia detalhado para a execução do trabalho de auditoria.</p> <p>Em resumo, um planejamento adequado de auditoria oferece benefícios significativos, como eficiência na alocação de recursos e redução de riscos e erros. Para definir a estratégia global desse planejamento, o auditor deve realizar ações que envolvem o entendimento da entidade, a identificação e avaliação de riscos, o estabelecimento de objetivos globais e o desenvolvimento de um plano detalhado de auditoria. Isso garante uma abordagem sistemática e eficaz na condução da auditoria das demonstrações contábeis.</p>	<p>ferramentas técnicas) necessários para a auditoria, garantindo que sejam suficientes e estejam disponíveis quando necessários.</p> <p>Determinação da materialidade: A materialidade é um conceito chave na auditoria, ajudando o auditor a determinar a relevância e a importância de certas transações ou saldos. A estratégia deve estabelecer como a materialidade será determinada e aplicada ao longo da auditoria.</p> <p>Comunicação com a equipe de auditoria: A estratégia global deve também estabelecer os meios e a frequência de comunicação entre os membros da equipe de auditoria. Uma comunicação eficaz assegura que todos os membros estejam alinhados em relação aos objetivos, riscos identificados e procedimentos a serem realizados.</p> <p>Em resumo, o planejamento adequado de auditoria, em linha com a NBC TA 300 (R1)/2016, não apenas facilita a condução dos trabalhos de auditoria, mas também aumenta a probabilidade de se alcançar uma opinião de auditoria apropriada e fundamentada.</p>	<p>Desenvolvimento de procedimentos de auditoria: O auditor deve desenvolver procedimentos de auditoria apropriados para reduzir os riscos de distorções relevantes a um nível aceitável. Comunicação com a administração e com os responsáveis pela governança: O auditor deve se comunicar com a administração e com os responsáveis pela governança sobre o escopo do trabalho, os procedimentos a serem realizados e os resultados esperados. A elaboração de um planejamento de auditoria adequado é essencial para que o auditor possa realizar um trabalho de qualidade e cumprir seus objetivos.</p>
Avaliação textual editor Microsoft Word		
<p>Pontuação Editor: 96 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal.</p>	<p>Pontuação Editor: 96 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal.</p>	<p>Pontuação Editor: 93 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal.</p>

Ortografia: 0 observações Gramática: 2 observações uso de “)/” Similaridade: “6% do seu texto é semelhante a fontes online “	Ortografia: 0 observações Gramática: 2 observações uso de “)/” Similaridade: “0% do seu texto é semelhante a fontes online “	Ortografia: 1 observação – palavra questionada “Bard” Gramática: 1 observação o uso de “)/” Similaridade: “0% do seu texto é semelhante a fontes online
Justificativa Nota Atribuída		
Conhecimentos técnico e desenvolvimento textual com 100% de aproveitamento.	Conhecimento técnico e desenvolvimento textual com 100% de aproveitamento.	Conhecimentos técnico e desenvolvimento textual com 100% de aproveitamento.

Gabarito Proposto

A NBC TA 300 (R1)/2016 destaca os benefícios de um planejamento de auditoria bem conduzido das demonstrações:

Dedicar atenção apropriada às áreas cruciais da auditoria.

Identificar e resolver rapidamente problemas em potencial.

Organizar o trabalho de auditoria para otimizar eficácia e eficiência.

Selecionar membros da equipe baseando-se em competência e capacidade, e alocar tarefas adequadamente.

Facilitar a orientação, supervisão e revisão do trabalho da equipe.

Se necessário, coordenar o trabalho realizado por outros auditores e especialistas.

No processo de definição da estratégia global, o auditor é auxiliado a determinar, com base na conclusão dos procedimentos de avaliação de risco:

- a) A alocação de recursos para áreas específicas de auditoria - por exemplo, designação de membros com experiência adequada para áreas de alto risco ou envolvimento de especialistas em temas complexos.
- b) O número de recursos destinados a áreas específicas, como a quantidade de membros para observar contagens de estoque ou o orçamento de horas para áreas de alto risco.
- c) O momento de alocação desses recursos, seja em uma etapa intermediária ou em uma data-base específica.
- d) A gestão, direção e supervisão desses recursos, abordando tópicos como reuniões planejadas, revisões do sócio e gerente, e revisões de controle de qualidade.

Ao definir a estratégia global de planejamento de auditoria, quatro ações são essenciais:

Entendimento da entidade: O auditor deve ter um entendimento aprofundado da entidade, seu setor e influências externas para identificar riscos e planejar a auditoria.

Alocação de recursos: A estratégia precisa determinar os recursos (tempo, pessoal e ferramentas) e garantir que estejam disponíveis e sejam apropriados.

Determinação da materialidade: A estratégia deve estabelecer como o conceito de materialidade, que refere-se à relevância de transações ou saldos, será empregado na auditoria.

Comunicação com a equipe: Os métodos e frequência de comunicação entre os membros da equipe de auditoria devem ser claramente estabelecidos para garantir alinhamento de objetivos e procedimentos.

Respostas dos modelos de AI Questão 27 - QTG

ChatGPT	Valor Questão:25	Nota: 05	GPT-4	Valor Questão:25	Nota: 05	Bard	Valor Questão:25	Nota: 05
<p>https://chat.openai.com/c/e6ce97eb-9b24-4b34-9d27-f08c64c89daf</p> <p>Base para Opinião</p> <p>Realizamos nossa auditoria de acordo com as Normas Brasileiras de Auditoria do Conselho Federal de Contabilidade (NBC TAs). Tais normas requerem que a auditoria seja planejada e executada com o objetivo de obter segurança razoável de que as demonstrações contábeis estão livres de distorção relevante.</p> <p>No contexto do Patrimônio Líquido da Comercial SLA S/A em 31 de dezembro de 2021, conforme apresentado no Papel de Trabalho (PT), o saldo inicial em 01 de janeiro de 2021 totalizava R\$13.210.000,00. Durante o exercício, observou-se um lucro de R\$5.000.000,00 e movimentações específicas como a constituição de uma reserva legal de R\$250.000,00, retenção de lucro de R\$2.375.000,00 destinado à expansão das unidades e instalações comerciais e proposta de dividendos no valor de R\$2.375.000,00. O saldo em 31 de dezembro de 2021 foi de R\$15.835.000,00.</p> <p>Também notamos, conforme anotado no PT, que a reserva de retenção de lucros foi constituída em conformidade com o art. 196 da Lei nº 6.404/1976, baseando-se em um orçamento de capital submetido pela administração à assembleia geral. Entretanto, identificamos que os dividendos são omissos no estatuto da empresa, uma informação relevante que deverá ser considerada na formação da opinião do auditor.</p> <p>Considerando as informações acima e as verificações realizadas conforme os tiques de revisão do PT, concluímos que as demonstrações contábeis do Patrimônio Líquido da Comercial SLA S/A, no que diz respeito ao período auditado, foram preparadas, em todos os aspectos relevantes, em conformidade com a estrutura de relatório financeiro aplicável, exceto pela omissão mencionada relacionada aos dividendos no estatuto da empresa.</p>			<p>https://chat.openai.com/c/c19447f2-0f20-4fa6-bb03-fd793f92e0f5</p> <p>Base para Opinião</p> <p>Realizamos nossa auditoria de acordo com as Normas Brasileiras de Auditoria do Conselho Federal de Contabilidade (NBC TAs). Nossas responsabilidades, conforme estas normas, estão descritas na seção subsequente "Responsabilidades do Auditor para a Auditoria das Demonstrações Contábeis". Somos independentes em relação à Comercial SLA S/A nos termos da ética e das normas profissionais, e cumprimos com nossas demais responsabilidades éticas de acordo com estas normas.</p> <p>As anotações em nosso Papel de Trabalho (PT) para a auditoria do Patrimônio Líquido da Comercial SLA S/A indicam os seguintes pontos:</p> <p>O saldo inicial em 01/janeiro/2021 do Patrimônio Líquido apresentou um total de 13.210.000,00, distribuído entre o Capital, Reserva Legal, Retenção de Lucro e Lucro Acumulado.</p> <p>O lucro do exercício foi de 5.000.000,00. Deste, foi destinado à Reserva Legal o valor de 250.000,00 e à Retenção de Lucro 2.375.000,00. Notou-se que a retenção de lucros foi constituída para expansão das unidades e instalações comerciais, conforme o art. 196 da Lei no 6.404/1976, e respaldada por um orçamento de capital submetido pela administração à assembleia geral.</p> <p>O saldo em 31-dez-2021 do Patrimônio Líquido totalizou 15.835.000,00, evidenciando as movimentações e retenções ocorridas no período.</p> <p>Foi observado que, embora haja dividendos propostos no valor de (2.375.000,00), estes são omissos no estatuto da empresa.</p>			<p>https://g.co/bard/share/c1edde3f4805</p> <p>Base para opinião</p> <p>Com base nos procedimentos de auditoria realizados, o auditor concluiu que as demonstrações contábeis da Comercial SLA S/A, referentes ao exercício social encerrado em 31 de dezembro de 2021, apresentam distorção relevante no item de patrimônio líquido denominado "Retenção de Lucro".</p> <p>Essa distorção decorre da constituição de uma reserva de retenção de lucros para expansão das unidades e instalações comerciais, por orçamento de capital submetido pela administração a assembleia geral, nos termos do art. 196 da Lei no 6.404/1976.</p> <p>O auditor realizou procedimentos de auditoria para verificar a adequação da constituição da reserva de retenção de lucros, incluindo a análise do orçamento de capital, a consulta ao estatuto social e a obtenção de evidências de que a assembleia geral aprovou a constituição da reserva.</p> <p>Essa distorção afeta o saldo do patrimônio líquido em R\$ 2.375.000,00, o que representa 15,0% do total do patrimônio líquido.</p> <p>Em decorrência dessa distorção, o auditor expressa sua opinião com ressalvas sobre as demonstrações contábeis da Comercial SLA S/A. A ressalva é apresentada na seção "Opinião" do relatório do auditor.</p> <p>Observações:</p> <p>A base para opinião deve ser apresentada de forma clara, concisa e objetiva.</p> <p>A base para opinião deve ser fundamentada nos procedimentos de auditoria realizados e nas evidências obtidas.</p> <p>A base para opinião deve ser apresentada em conformidade com o formato determinado pela norma NBC TA 700/2016.</p>		

<p>Esta base para opinião foi elaborada de acordo com a NBC TA 700/2016 em vigor.</p>	<p>Com base nas informações contidas no PT e em nossa avaliação dos riscos de distorção relevante, concluímos que as demonstrações contábeis da Comercial SLA S/A representam adequadamente, em todos os aspectos relevantes, a posição patrimonial da empresa, em conformidade com a estrutura de relatório financeiro aplicável, exceto pela omissão dos dividendos no estatuto.</p>	<p>No caso em tela, a base para opinião é apresentada de forma clara e concisa, fundamentando-se nos procedimentos de auditoria realizados e nas evidências obtidas. A base para opinião também está em conformidade com o formato determinado pela norma NBC TA 700/2016.</p> <p>A ressalva expressa pelo auditor é apropriada, pois a distorção relevante no item de patrimônio líquido denominado "Retenção de Lucro" afeta o saldo do patrimônio líquido em um percentual significativo.</p>
Avaliação textual editor Microsoft Word		
<p>Pontuação Editor: 94 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 1 observação uso de "NBC TAs" Gramática: 0 observações Similaridade: "10% do seu texto é semelhante a fontes online"</p>	<p>Pontuação Editor: 94 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 1 observação uso de "NBC TAs" Gramática: 0 observações Similaridade: "0% do seu texto é semelhante a fontes online"</p>	<p>Pontuação Editor: 95 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 1 observação – palavra questionada "Bard" Gramática: 0 observações Similaridade: "0% do seu texto é semelhante a fontes online"</p>
Justificativa Nota Atribuída		
<p>Não apresentou conhecimento técnico correto para o desenvolvimento questão, não considerou todos os dados apresentados para apontamento da opinião sobre a empresa, ocasionando a perda de 20 pontos na nota.</p>	<p>Não apresentou conhecimento técnico correto para o desenvolvimento questão, não considerou todos os dados apresentados para apontamento da opinião sobre a empresa, ocasionando a perda de 20 pontos na nota.</p>	<p>Não apresentou conhecimento técnico correto para o desenvolvimento questão, não considerou todos os dados apresentados para apontamento da opinião sobre a empresa, ocasionando a perda de 20 pontos na nota.</p>

Gabarito Proposto

Após uma análise detalhada dos papéis de trabalho e das demais informações fornecidas, identificamos movimentações específicas relacionadas à constituição da Reserva Legal e à proposta de Dividendos:

Com base na análise do Patrimônio Líquido (PL), constatamos uma Reserva Legal de R\$ 250.000,00. Segundo a Lei 6.404, a constituição da reserva legal deve ser de 5% do Lucro Líquido do exercício, com um limite de até 20% do Capital Social. Considerando um Capital Social de R\$ 10.550.000,00, os 20% correspondem a R\$ 2.110.000,00. Observamos que esse limite de Reserva Legal já foi atingido em exercícios anteriores, portanto, não deveria haver nova constituição de reserva neste período.

Quanto à distribuição de dividendos, o estatuto não é claro. Assim, aplicamos a regra padrão de 50% do lucro líquido ajustado. O Lucro Líquido Ajustado, desconsiderando a Reserva Legal indevida, seria (5.000.000,00 – 250.000,00) = 4.750.000,00. Deste montante, seriam distribuídos 50%, ou seja, R\$ 2.375.000,00 como dividendos. No entanto, sem a constituição da Reserva Legal, a distribuição de dividendos deveria ser 50% de R\$ 5.000.000,00, portanto, deveria distribuir R\$ 2.500.000,00. Detectamos evidências de distorções significativas e não limitadas apenas a um item do NBTC 705 referente ao patrimônio líquido. Tais distorções impactam o balanço patrimonial, indicando um passivo subavaliado e inconsistências nas DLPA e DMPL.

Com relação a retenção de lucros refere-se a parte do lucro não distribuída como dividendos, sendo reinvestida na própria empresa, já o orçamento de capital é um processo que determina e avalia grandes investimentos ou despesas, como compra de equipamentos ou expansão de mercado. No Brasil, o Art. 196 da LSA estabelece que a decisão de reter lucros deve partir de uma proposta administrativa e ser aprovada pela assembleia-geral, estando alinhada a um orçamento de capital, conforme nota apresentada no papel de trabalho do auditor.

Concluímos que as informações analisadas conduzem a um parecer com opinião adversa.

ATUAÇÃO EM AUDITORIA INDEPENDENTE NAS INSTITUIÇÕES REGULADAS PELA COMISSÃO DE VALORES MOBILIÁRIOS (CVM) – Discursivas

Respostas dos modelos de AI Questão 26 - CVM:

ChatGPT	Valor Questão:25	Nota: 20	GPT-4	Valor Questão:25	Nota: 20	Bard	Valor Questão:25	Nota: 10
<p>CVM discursiva 3.5 (openai.com)</p> <p>As operações de fusão são processos complexos e que exigem uma compreensão detalhada das condições financeiras das entidades envolvidas, bem como dos impactos da fusão. Neste contexto, surgem as informações financeiras pro forma.</p> <p>1. Finalidade das informações financeiras pro forma: As informações financeiras pro forma têm por finalidade apresentar aos usuários das demonstrações financeiras, incluindo investidores, credores e demais partes interessadas, uma visão do que seria a situação financeira e os resultados da nova entidade, caso a operação de fusão tivesse ocorrido em uma data anterior. Assim, elas proporcionam uma visão simulada e ajudam na tomada de decisões ao oferecer uma perspectiva sobre o desempenho e posição financeira consolidados.</p> <p>2. Definição do nível de asseguarção requerido nas informações financeiras pro forma: O nível de asseguarção refere-se à extensão da confiança que os usuários podem depositar nas informações. Para as informações financeiras pro forma, é requerido um nível de asseguarção de revisão, ou seja, os profissionais devem realizar procedimentos analíticos e de inquirição para obter segurança moderada de que as informações estão livres de distorções relevantes.</p> <p>3. Requisitos para apresentação das demonstrações financeiras da Cia. Beta e da Gama Ind. & Com. Ltda.: Considerando que a Cia. Beta é uma sociedade anônima de capital fechado e a Gama Ind. & Com. Ltda. é uma sociedade limitada, suas demonstrações financeiras devem estar em conformidade com</p>			<p>CVM discursiva 4.0 (openai.com)</p> <p>As operações de fusão no mercado brasileiro são acompanhadas de uma série de requisitos de divulgação e conformidade. O Comunicado Técnico (CTO) 02 fornece orientações específicas sobre tais operações, incluindo as informações financeiras pro forma.</p> <p>1. Finalidade das Informações Financeiras Pro Forma: As informações financeiras pro forma têm o objetivo de fornecer aos stakeholders uma visão sobre como a entidade resultante teria se apresentado financeiramente se a fusão tivesse ocorrido em uma data anterior específica. Estas são elaboradas combinando os registros financeiros das entidades que estão sendo fundidas, considerando quaisquer ajustes que sejam necessários para refletir a operação como se fosse uma única entidade desde o início.</p> <p>2. Definição do Nível de Asseguarção Requerido nas Informações Financeiras Pro Forma: O CTO 02 exige que as informações financeiras pro forma sejam submetidas a procedimentos de revisão de asseguarção por parte de auditores independentes. Esta revisão visa fornecer um nível moderado de confiança de que as informações estão livres de distorções materiais. Não é tão detalhado quanto uma auditoria completa, mas fornece uma certa garantia aos usuários das demonstrações financeiras.</p> <p>3. Requisitos para Apresentação das Demonstrações Financeiras da Cia. Beta e da Gama Ind. & Com. Ltda.: Ambas as entidades, Cia. Beta e Gama Ind. & Com. Ltda., devem apresentar suas demonstrações financeiras de acordo</p>			<p>https://g.co/bard/share/2035b2cf6276</p> <p>A finalidade das informações financeiras pro forma</p> <p>As informações financeiras pro forma são elaboradas para demonstrar o impacto de uma transação ou evento hipotético nas demonstrações financeiras de uma entidade. Elas são úteis para investidores e credores para avaliar os possíveis resultados de uma transação ou evento, antes de sua concretização.</p> <p>A definição do nível de asseguarção requerido nas informações financeiras pro forma</p> <p>O nível de asseguarção requerido nas informações financeiras pro forma é de asseguarção razoável. Isso significa que o auditor independente deve expressar uma opinião sobre se as informações financeiras pro forma, em todos os aspectos relevantes, estão preparadas em conformidade com os critérios de apresentação aplicáveis, de forma a fornecer asseguarção razoável de que as informações financeiras pro forma não contêm distorções relevantes, quer sejam causadas intencionalmente ou não.</p> <p>Os requisitos para apresentação das demonstrações financeiras da Cia. Beta e da Gama Ind. & Com. Ltda.</p> <p>As demonstrações financeiras da Cia. Beta e da Gama Ind. & Com. Ltda. devem ser apresentadas de acordo com as normas brasileiras de contabilidade aplicáveis, de forma a refletir o impacto da transação de fusão.</p> <p>O significado da precisão das informações financeiras pro forma</p> <p>A precisão das informações financeiras pro forma é importante para que os investidores e credores possam tomar decisões de investimento e financiamento com base em informações confiáveis. O auditor independente deve aplicar</p>		

<p>as normas contábeis aplicáveis no Brasil. Adicionalmente, ao se envolverem em uma fusão que resultará em uma sociedade anônima de capital aberto, pode ser necessário adaptar e/ou revisar suas demonstrações para garantir conformidade com os padrões e exigências adicionais para empresas de capital aberto.</p> <p>4. Significado da perspectiva, considerada como exigência ética relevante: A perspectiva, neste contexto, refere-se à necessidade de os profissionais envolvidos no trabalho de asseguarção manterem um ponto de vista imparcial e objetivo, sem se deixarem influenciar por interesses pessoais ou pressões externas. Manter essa perspectiva é fundamental para garantir a integridade e a confiabilidade do trabalho de asseguarção, garantindo que as informações apresentadas são verídicas e apresentadas de forma justa.</p> <p>Portanto, o Comunicado Técnico (CTO) 02 estabelece diretrizes fundamentais que visam garantir a transparência, integridade e confiabilidade das informações financeiras no contexto de operações de fusão, como o descrito.</p>	<p>com as normas contábeis aplicáveis. Como a Cia. Beta é uma sociedade anônima de capital fechado e Gama é uma sociedade limitada, suas demonstrações financeiras não necessariamente têm o mesmo nível de detalhamento exigido de uma empresa de capital aberto. No entanto, para fins de fusão, é essencial que todas as demonstrações sejam transparentes e completas para proporcionar uma compreensão clara da posição financeira das empresas antes da fusão.</p> <p>4. Significado da Perspectiva Relacionada ao Trabalho de Asseguarção: A perspectiva, no contexto do trabalho de asseguarção, refere-se à integridade e objetividade com a qual o auditor ou profissional conduz sua revisão. É uma exigência ética relevante que garante que o trabalho seja feito sem viés, com um compromisso com a verdade e a justiça. Isso assegura que os stakeholders possam confiar nas informações financeiras apresentadas e nas conclusões dos profissionais envolvidos.</p> <p>Em suma, o CTO 02 garante que as operações de fusão sejam conduzidas com transparência e rigor, proporcionando confiança aos investidores e outras partes interessadas no mercado financeiro brasileiro.</p>	<p>procedimentos de auditoria para obter evidência de que as informações financeiras pro forma estão preparadas com precisão.</p> <p>Aplicação à operação hipotética No caso hipotético apresentado, as informações financeiras pro forma serão elaboradas para demonstrar o impacto da fusão das três entidades. O auditor independente deverá expressar uma opinião sobre se as informações financeiras pro forma, em todos os aspectos relevantes, estão preparadas em conformidade com os critérios de apresentação aplicáveis, de forma a fornecer asseguarção razoável de que as informações financeiras pro forma não contêm distorções relevantes, quer sejam causadas intencionalmente ou não.</p> <p>Para isso, o auditor independente deverá obter evidência de que as demonstrações financeiras da Cia. Beta e da Gama Ind. & Com. Ltda. foram preparadas de acordo com as normas brasileiras de contabilidade aplicáveis, de forma a refletir o impacto da transação de fusão. O auditor independente também deverá considerar os seguintes fatores na avaliação da precisão das informações financeiras pro forma:</p> <p>A complexidade da transação ou evento hipotético; A quantidade e a qualidade da evidência disponível; A subjetividade das estimativas e projeções utilizadas; O risco de distorções relevantes.</p> <p>A conclusão do auditor independente deve ser expressa em um relatório de asseguarção razoável.</p>
Avaliação textual editor Microsoft Word		
<p>Pontuação Editor: 90 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 0 observações Gramática: 5 observações – palavra questionada “proforma” Similaridade: “0% do seu texto é semelhante a fontes online “</p>	<p>Pontuação Editor: 91 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 0 observações Gramática: 5 observações – palavra questionada “proforma” Similaridade: “0% do seu texto é semelhante a fontes online “</p>	<p>Pontuação Editor: 82 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 11 observações – palavra questionada “proforma” Gramática: 0 observações Similaridade: “0% do seu texto é semelhante a fontes online “</p>
Justificativa Nota Atribuída		

Conhecimentos técnico apresentados com 100% de aproveitamento e desenvolvimento textual menor 95% ocasionando o desconto 5 pontos na pontuação.	Conhecimentos técnico apresentados com 100% de aproveitamento e desenvolvimento textual menor 95% ocasionando o desconto 5 pontos na pontuação.	Não apresentou conhecimento técnico suficiente para elaboração da questão e apresentou capacidade de desenvolvimento textual inferior aos outros modelos para essa questão, ocasionando a perda de 15 pontos na nota.
--	--	--

Gabarito Proposto

O Comunicado CTO 02, destaca que a finalidade das informações financeiras pro forma é a de exclusivamente ilustrar o impacto do evento ou da transação relevante sobre as informações financeiras históricas da entidade, como se o evento ou a transação tivesse ocorrido na data anterior selecionada para propósito ilustrativo. Sendo assim, em função de o requerimento de asseguarção razoável sobre informações financeiras pro forma ser proveniente de instrução emitida pelo órgão regulador do mercado de capitais (no caso, a CVM), a NBC TO 3420 é a norma aplicável para a realização desses trabalhos de asseguarção. A NBC TO 3420 destaca, entre vários assuntos, aspectos como: aceitação do trabalho; planejamento e execução; materialidade; obtenção de evidência sobre a adequação dos ajustes pro forma; representações formais; formação da opinião; e elaboração do relatório de asseguarção. Portanto, é fundamental sua completa leitura e entendimento antes da aceitação e realização do trabalho de asseguarção sobre informações financeiras pro forma. O auditor independente deve cumprir as exigências éticas relevantes, incluindo aquelas pertinentes à independência, relacionadas com trabalhos de asseguarção. Além disso, as demonstrações financeiras da empresa apresentada na questão devem:

Para os efeitos da operação, as sociedades envolvidas devem divulgar demonstrações financeiras, cuja data base: (i) seja a mesma para todas as sociedades envolvidas; (ii) não seja anterior a 180 (cento e oitenta) dias da data da assembleia que deliberará sobre a operação; (iii) as demonstrações financeiras referidas no caput devem ser elaboradas de acordo com a Lei n.º 6.404 e com as normas da CVM; (iv) Devem ser auditadas por auditor independente registrado na CVM; (v) As sociedades envolvidas na operação devem elaborar informações financeiras pro forma das sociedades que subsistirem ou que resultarem da operação, como se estas já existissem, referentes à data das demonstrações financeiras referidas. Além disso, temos a Orientação CPC 06 orienta sobre a elaboração e apresentação de demonstrações financeiras proforma, assegurando sua relevância, confiabilidade e comparabilidade. Ela define critérios de quando e como apresentar tais informações. Ao exibir dados proforma, é mandatório reconciliá-los com dados financeiros históricos e explicar os métodos e hipóteses adotados. A Orientação também destaca a necessidade de transparência, permitindo que os usuários avaliem a robustez das projeções e estejam cientes de eventos futuros que possam influenciar as informações apresentadas.

Questão nº 27

Em 01/01/X0, uma entidade adquiriu móveis para serem utilizados em seu escritório por R\$ 100.000,00. O frete para a entrega dos móveis era de R\$ 2.000,00, e a instalação, de R\$ 8.000,00.

A vida útil esperada dos móveis era de dez anos, e a entidade não considerava o valor residual. A entidade depreciava os seus ativos imobilizados pelo método de linha reta e realizava o teste de recuperabilidade de acordo com a NBC TG 01 (R4), de 2017, - Redução ao Valor Recuperável de Ativos.

No ano de X0, não houve indícios de perda por recuperabilidade.

Em 01/01/X1, um dos móveis precisou de reparo, o que custou à empresa R\$ 2.000,00.

Em 31/12/X1, a entidade estimou que o valor em uso dos móveis era de R\$ 85.000,00. Já o valor justo líquido de despesa de venda era de R\$ 90.000,00.

Em 31/12/X2, a entidade estimou que o valor em uso dos móveis era de R\$ 75.000,00. Já o valor justo líquido de despesa de venda era de R\$ 70.000,00.

Nesse contexto, elabore um texto, de 15 a 30 linhas, apresentando o valor contábil dos móveis em 31/12/X0, 31/12/X1 e 31/12/X2, explicitando o tratamento contábil aplicado e apresentando a memória de cálculo.

O texto deve desenvolver-se no sentido de:

- explicar como é calculada a depreciação anual dos móveis;
- explicar como é obtido o valor contábil em cada ano, considerando a depreciação e o teste de recuperabilidade;
- indicar se há perda por recuperabilidade e, em caso positivo, em qual(is) ano(s).

(Valor: 25,0 pontos)

Respostas dos modelos de AI Questão 27 - CVM:

ChatGPT	Valor Questão:25	Nota: 20	GPT-4	Valor Questão:25	Nota: 20	Bard	Valor Questão:25	Nota: 0
CVM discursiva 3.5 (openai.com)			CVM discursiva 4.0 (openai.com)			https://g.co/bard/share/2035b2cf6276		
Vamos começar pela determinação do valor contábil dos móveis em 01/01/X0 e calcular sua depreciação anual.			Vamos começar avaliando o custo inicial dos móveis, o cálculo da depreciação anual e, posteriormente, examinar o valor contábil nos anos indicados.			Valor contábil dos móveis		
1. Cálculo do valor contábil inicial: Valor de aquisição dos móveis: R\$ 100.000,00 Frete: R\$ 2.000,00 Instalação: R\$ 8.000,00 Total: R\$ 110.000,00			Custo Inicial dos Móveis: Valor de compra: R\$ 100.000,00 Frete: R\$ 2.000,00 Instalação: R\$ 8.000,00 Custo total: R\$ 110.000,00 Depreciação Anual:			Em 31 de dezembro de X0, o valor contábil dos móveis era de R\$ 100.000,00, sendo R\$ 100.000,00 o custo de aquisição e R\$ 0,00 o valor residual.		
						A depreciação anual dos móveis é calculada pela fórmula: Depreciação anual = (Custo - Valor residual) / Vida útil		

<p>A depreciação pelo método da linha reta é calculada dividindo o valor total do ativo pela sua vida útil. Neste caso, os móveis têm uma vida útil de 10 anos e não possuem valor residual.</p> <p>2. Cálculo da depreciação anual: Depreciação anual = Valor contábil inicial / Vida útil Depreciação anual = R\$ 110.000,00 / 10 = R\$ 11.000,00 por ano</p> <p>31/12/X0: Valor contábil = Valor contábil inicial - Depreciação anual Valor contábil = R\$ 110.000,00 - R\$ 11.000,00 = R\$ 99.000,00</p> <p>Não há indícios de perda por recuperabilidade neste ano, logo o valor contábil em 31/12/X0 é R\$ 99.000,00.</p> <p>31/12/X1: O reparo de R\$ 2.000,00 não altera o valor contábil dos móveis, pois é uma despesa de manutenção. Valor contábil = Valor contábil (31/12/X0) - Depreciação anual Valor contábil = R\$ 99.000,00 - R\$ 11.000,00 = R\$ 88.000,00</p> <p>Para o teste de recuperabilidade, usamos o maior valor entre o valor em uso (R\$ 85.000,00) e o valor justo líquido de despesas de venda (R\$ 90.000,00). Como o valor contábil (R\$ 88.000,00) é menor do que R\$ 90.000,00, não há perda por recuperabilidade em X1.</p> <p>31/12/X2: Valor contábil = Valor contábil (31/12/X1) - Depreciação anual Valor contábil = R\$ 88.000,00 - R\$ 11.000,00 = R\$ 77.000,00</p> <p>Para o teste de recuperabilidade em X2, o maior valor entre o valor em uso (R\$ 75.000,00) e o valor justo líquido de despesas de venda (R\$ 70.000,00) é R\$ 75.000,00. Como o valor contábil (R\$ 77.000,00) é maior do que R\$ 75.000,00, há uma perda por recuperabilidade de R\$ 2.000,00.</p>	<p>Como a vida útil esperada dos móveis é de dez anos e não há valor residual considerado, a depreciação anual pelo método de linha reta é calculada dividindo o custo total pelo período de vida útil. Depreciação anual = R\$ 110.000,00 / 10 anos = R\$ 11.000,00 por ano.</p> <p>31/12/X0: Custo inicial: R\$ 110.000,00 Depreciação (1 ano): R\$ 11.000,00 Valor contábil em 31/12/X0 = R\$ 110.000,00 - R\$ 11.000,00 = R\$ 99.000,00</p> <p>31/12/X1: Valor contábil inicial de X1: R\$ 99.000,00 (valor contábil de 31/12/X0) Depreciação (1 ano): R\$ 11.000,00 Reparo: R\$ 2.000,00 (Este valor não é deduzido do valor contábil, pois é uma despesa e não afeta o valor contábil dos móveis). Valor contábil em 31/12/X1 antes do teste de recuperabilidade = R\$ 88.000,00 Teste de Recuperabilidade em 31/12/X1: Os valores a serem comparados são: Valor contábil: R\$ 88.000,00 Maior valor entre o valor em uso (R\$ 85.000,00) e o valor justo líquido de despesa de venda (R\$ 90.000,00): R\$ 90.000,00 Não há perda por recuperabilidade em 31/12/X1, pois o maior valor recuperável (R\$ 90.000,00) é superior ao valor contábil (R\$ 88.000,00).</p> <p>31/12/X2: Valor contábil inicial de X2: R\$ 88.000,00 Depreciação (1 ano): R\$ 11.000,00 Valor contábil em 31/12/X2 antes do teste de recuperabilidade = R\$ 77.000,00 Teste de Recuperabilidade em 31/12/X2:</p>	<p>No caso dos móveis, a depreciação anual é de R\$ 10.000,00, sendo:</p> <p>Depreciação anual = (100.000,00 - 0,00) / 10 = 10.000,00 Em 31 de dezembro de X1, o valor contábil dos móveis era de R\$ 90.000,00, sendo R\$ 80.000,00 o valor contábil após a depreciação e R\$ 10.000,00 a reserva para redução ao valor recuperável.</p> <p>O valor contábil após a depreciação foi obtido da seguinte forma:</p> <p>Valor contábil após depreciação = Valor contábil anterior + Depreciação anual Valor contábil após depreciação = 100.000,00 + 10.000,00 = 110.000,00 A reserva para redução ao valor recuperável foi criada para reconhecer a perda por recuperabilidade, que foi estimada em R\$ 20.000,00, sendo:</p> <p>Reserva para redução ao valor recuperável = Valor em uso - Valor justo líquido de despesa de venda Reserva para redução ao valor recuperável = 85.000,00 - 90.000,00 = -5.000,00 Em 31 de dezembro de X2, o valor contábil dos móveis era de R\$ 75.000,00, sendo R\$ 65.000,00 o valor contábil após a depreciação e R\$ 10.000,00 a reserva para redução ao valor recuperável.</p> <p>O valor contábil após a depreciação foi obtido da seguinte forma:</p> <p>Valor contábil após depreciação = Valor contábil anterior + Depreciação anual Valor contábil após depreciação = 110.000,00 + 10.000,00 = 120.000,00 A reserva para redução ao valor recuperável foi acrescida em R\$ 5.000,00, sendo:</p>
--	--	---

<p>Em resumo:</p> <p>Valor contábil em 31/12/X0 = R\$ 99.000,00 Valor contábil em 31/12/X1 = R\$ 88.000,00 Valor contábil em 31/12/X2 = R\$ 75.000,00 após ajuste por perda por recuperabilidade.</p>	<p>Os valores a serem comparados são:</p> <p>Valor contábil: R\$ 77.000,00 Maior valor entre o valor em uso (R\$ 75.000,00) e o valor justo líquido de despesa de venda (R\$ 70.000,00): R\$ 75.000,00 Há uma perda por recuperabilidade em 31/12/X2 de R\$ 2.000,00 (R\$ 77.000,00 - R\$ 75.000,00). Portanto, o valor contábil dos móveis em 31/12/X2 após o ajuste é de R\$ 75.000,00.</p> <p>Em resumo, o valor contábil dos móveis é R\$ 99.000,00 em 31/12/X0, R\$ 88.000,00 em 31/12/X1 e R\$ 75.000,00 em 31/12/X2. Houve perda por recuperabilidade apenas em 31/12/X2.</p>	<p>Reserva para redução ao valor recuperável = $-5.000,00 + 5.000,00 = 0,00$ Conclusão Em 31 de dezembro de X2, houve perda por recuperabilidade, pois o valor em uso dos móveis era inferior ao valor justo líquido de despesa de venda. A perda por recuperabilidade foi reconhecida pela criação da reserva para redução ao valor recuperável. A memória de cálculo dos valores contábeis dos móveis está apresentada a seguir:</p> <table border="0"> <tr> <td>Data</td> <td>Valor contábil</td> </tr> <tr> <td>31/12/X0</td> <td>R\$ 100.000,00</td> </tr> <tr> <td>31/12/X1</td> <td>R\$ 90.000,00</td> </tr> <tr> <td>31/12/X2</td> <td>R\$ 75.000,00</td> </tr> </table> <p>É importante ressaltar que a perda por recuperabilidade pode ser revertida no futuro, caso o valor em uso dos móveis aumente.</p>	Data	Valor contábil	31/12/X0	R\$ 100.000,00	31/12/X1	R\$ 90.000,00	31/12/X2	R\$ 75.000,00
Data	Valor contábil									
31/12/X0	R\$ 100.000,00									
31/12/X1	R\$ 90.000,00									
31/12/X2	R\$ 75.000,00									

Avaliação textual editor Microsoft Word

<p>Pontuação Editor: 100% Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 0 observações Gramática: 0 observações Similaridade: “0% do seu texto é semelhante a fontes online “</p>	<p>Pontuação Editor: 100% Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 0 observações Gramática: 0 observações Similaridade: “13% do seu texto é semelhante a fontes online “</p>	<p>Pontuação Editor: 100 % Escrita formal - Aperfeiçoamentos 100%: Clareza; Concisão; Escolha Vocabular; Linguagem Formal. Ortografia: 0 observações Gramática: 0 observações Similaridade: “5% do seu texto é semelhante a fontes online “</p>
---	--	--

Justificativa Nota Atribuída

<p>Conhecimentos técnico apresentado para o cálculo correto, porém poderia ter mais detalhamento no desenvolvimento textual como solicitado no enunciado, ocasionando o desconto 5 pontos na pontuação.</p>	<p>Conhecimentos técnico apresentado para o cálculo correto, porém poderia ter mais detalhamento no desenvolvimento textual como solicitado no enunciado, ocasionando o desconto 5 pontos na pontuação.</p>	<p>Não apresentou conhecimento técnico correto para o desenvolvimento questão, não sendo possível atribuir uma nota.</p>
--	--	---

Gabarito Proposto

Para determinar a depreciação anual dos móveis, é necessário considerar o custo total de aquisição, que compreende o preço de compra e os gastos adicionais relacionados ao frete e instalação, conforme orientação do CPC 27. Portanto, ao somar R\$ 100.000,00 (valor de aquisição), R\$ 2.000,00 (frete) e R\$ 8.000,00 (instalação), obtemos um valor inicial de R\$ 110.000,00. A depreciação anual é então calculada dividindo esse valor inicial pelo período de vida útil dos móveis. Assim, a taxa de depreciação dos móveis é 10% a.a - $R\$ 110.000,00 \times 10\% \text{ a.a} = R\$ 11.000,00$ (valor da depreciação anual).

Em 31/12/X0, ao subtrairmos a depreciação do valor inicial, chegamos ao valor contábil de R\$ 99.000,00 ($R\$ 110.000,00 - R\$ 11.000,00$).

No ano seguinte, 31/12/X1, após deduzir a depreciação do valor contábil do ano anterior, obtemos R\$ 88.000,00.

Em X2, após a depreciação, o valor contábil dos móveis seria R\$ 77.000,00. Ao realizar o teste de recuperabilidade, identifica-se uma perda por recuperabilidade de R\$ 2.000,00 (R\$ 77.000,00 - R\$ 75.000,00), já que no momento do teste predomina o maior valor entre o valor de liquidação e o valor do valor em uso. Portanto o *impairment* é de R\$2.000 apenas e o valor contábil em X2 será de R\$75.000.