

Avaliação de LLMs em Relação a Questões Básicas de Redes de Computadores

Luca Vianna Martins Silveira¹, Luciano Bernardes de Paula¹

¹Instituto Federal de Educação, Ciência e Tecnologia de São Paulo (IFSP)
Bragança Paulista - SP - Brasil

luca.silveira@aluno.ifsp.edu.br, lbernardes@ifsp.edu.br

Abstract. *The popularization of LLMs (Large Language Models) in people's daily lives is becoming increasingly evident. Technology students with basic knowledge and internet access can consult them to support their learning. In Computer Science, computer networks are a fundamental topic, as they underpin all the infrastructure of today's digital world. The aim of this study is to analyze responses from LLMs regarding introductory and fundamental topics in computer networks. To this end, questions covering these topics were submitted to four different LLMs. Original multiple-choice questions, a rewritten multiple-choice version, and an essay-type version were used. The results showed that the overall accuracy rate of the LLMs was equal to or higher than 97.5%.*

Resumo. *A popularização dos LLMs (Large Language Models) no cotidiano das pessoas se mostra cada vez maior. Estudantes de tecnologia com nível básico de conhecimento e acesso a internet podem consultá-los para auxiliar em seu aprendizado. Na Computação, as redes de computadores são um assunto fundamental, pois alicerçam toda a infraestrutura do mundo digital atual. O objetivo deste trabalho é analisar respostas obtidas de LLMs em relação a assuntos introdutórios e fundamentais das redes de computadores. Para isso, foram submetidas a quatro diferentes LLMs questões que abordam esses assuntos. Foram utilizadas questões objetivas originais, uma versão objetiva reescrita e uma versão dissertativa. Os resultados mostraram que a taxa de acertos geral dos LLMs ficou igual ou superior a 97,5%.*

1. Introdução

Os LLMs (*Large Language Models*), desde o lançamento do ChatGPT ao público em geral, podem ser utilizados para diversos fins. Devido à facilidade de uso, uma vez que podem ser utilizados por meio da linguagem natural, não há barreira para que estudantes de tecnologia possam consultá-los sobre qualquer assunto.

Entre os fins para os quais os LLMs estão sendo utilizados, está a Educação [Marques and Morandini 2024] e, atualmente, é discutido e analisado o impacto dos LLMs no processo educativo, em diferentes níveis [Soares et al. 2023]. O uso dos LLMs para aprender é especialmente delicado em relação àqueles estudantes iniciantes no assunto alvo da aprendizagem, pois é sabido que essas ferramentas podem cometer erros e gerar informações equivocadas, levando pessoas que não possuem experiência nos assuntos consultados a considerarem as respostas como corretas, sem questionamentos.

Dentre os assuntos fundamentais da área de tecnologia, pode-se citar as redes de computadores [Kurose and Ross 2021], as quais são a base para toda a infraestrutura tecnológica que existe atualmente. O seu entendimento, tanto da parte teórica quanto da prática, é considerado um tópico básico na trajetória formativa de estudantes em cursos de tecnologia, os quais, no geral, possuem disciplinas sobre esse assunto em suas grades curriculares. Há, inclusive, cursos em diferentes níveis de educação nos quais o foco principal é as redes de computadores, tais como cursos de Tecnologia em Redes de Computadores e similares.

Como qualquer outro assunto atual, estudantes iniciantes na área de redes de computadores naturalmente utilizam e utilizarão os LLMs para sanarem dúvidas, resolverem questões e aprenderem sobre esse assunto. Portanto, é importante que seja testado o repertório dos LLMs em relação a esse conteúdo. Desta forma, este trabalho tem como objetivo analisar as respostas geradas por quatro dos principais LLMs atuais a respeito de tópicos considerados básicos da área de redes de computadores. Para os testes, realizados no período de maio a junho de 2025, foram utilizadas as versões gratuitas disponíveis para o público dos seguintes LLMs: ChatGPT (da OpenAI), Copilot (da Microsoft), DeepSeek (da própria DeepSeek) e o Gemini (da Google).

Inicialmente, foram definidos alguns tópicos da área de redes de computadores considerados básicos e, a partir desses tópicos, foram definidas questões relacionadas a cada um deles, e essas foram submetidas aos LLMs. As questões utilizadas foram obtidas de provas de concursos públicos para área de tecnologia, sendo consideradas apenas aquelas que envolviam a área de redes de computadores. As questões utilizadas eram todas objetivas e, para análise a respeito do desempenho dos LLMs, foram utilizadas de três formas: em sua forma original, assim como obtidas em suas fontes; de forma objetiva, porém reescritas, sem alteração de sentido e conteúdo; e reescritas para uma versão que exige uma resposta dissertativa. Essa variação foi feita para analisar se o comportamento dos LLMs é alterado, caso as questões também sejam alteradas.

Portanto, o objetivo principal é analisar a qualidade das respostas dos diferentes LLMs em relação aos tópicos básicos das redes de computadores, principalmente considerando que esses podem ser usados por alunos iniciantes, que não possuem experiência em relação a esse conteúdo.

Este artigo está organizado da seguinte forma: na Seção 2 é apresentada a fundamentação teórica do trabalho; na Seção 2.1 são apresentados trabalhos encontrados na literatura que são relacionados com este; na Seção 3 toda a metodologia utilizada no trabalho é explicada, inclusive todos os passos em relação aos testes realizados; na Seção 4 são descritos os testes realizados, assim como os resultados obtidos; na Seção 4.1 é apresentada uma discussão a respeito dos resultados; na Seção 5 são apresentadas as limitações da pesquisa e os trabalhos futuros e na Seção 6 o trabalho é concluído.

2. Fundamentação Teórica

Os LLMs são modelos estatísticos de grande escala que, após o lançamento do ChatGPT ao público em geral, têm sido amplamente utilizados por usuários para consultas de diversos tipos [Minaee et al. 2025]. O diferencial crucial dos LLMs é a possibilidade de utilizá-los por meio de linguagem natural, o que permite que, mesmo usuários sem conhecimento técnico, possam utilizá-los.

Uma vez que os LLMs são considerados inteligências artificiais generativas, ou seja, não são especializados em nenhum assunto específico, é possível utilizá-los para consultar sobre qualquer assunto e qualquer motivo, inclusive, para aprender. Entretanto, o uso dos LLMs na Educação é um assunto discutido atualmente, por vários motivos, tais como a geração de informações erradas, que devem ser checadas, questões sobre plágio e integridade acadêmica, entre outros [Lo 2023].

Atualmente, existem diversos LLMs disponíveis para uso, de forma gratuita. Neste trabalho foram utilizados quatro deles: ChatGPT, Copilot, DeepSeek e Gemini, que são apresentados a seguir.

O ChatGPT, segundo a sua desenvolvedora OpenAI, é um serviço *online* de inteligência artificial que pode ser utilizado para diversas tarefas. Esta ferramenta é capaz de processar e responder os *prompts* fornecidos pelo usuário, devido a um processo de aprendizagem, no qual o ChatGPT é submetido a um grande volume de dados. O modelo é capaz de reconhecer padrões nos textos analisados e completar frases, baseando-se na previsão da próxima palavra mais provável [OpenAI 2025]. O ChatGPT foi escolhido para ser usado neste trabalho, uma vez que foi o primeiro a se tornar público para uso.

O Copilot, segundo a sua desenvolvedora Microsoft, é um assistente conversacional que faz uso de inteligência artificial para aumentar a produtividade e automatizar tarefas. Esse assistente é projetado para adaptar-se às necessidades do usuário, fornecendo respostas mais adequadas ao seu contexto. Além disso, o Copilot possui integração com aplicativos do Microsoft 365, permitindo que o modelo auxilie em projetos no Word, Excel, PowerPoint, Teams, entre outros [Microsoft 2025]. Por ser desenvolvido pela Microsoft, uma empresa de tecnologia mundialmente conhecida, o Copilot foi escolhido para ser testado neste trabalho.

O DeepSeek é um modelo de aprendizagem de máquina que possui duas características únicas: possui código aberto e foi treinado utilizando componentes computacionais menos potentes que seus concorrentes. Dessa forma, o DeepSeek não necessita de muito poder computacional para responder as solicitações, implementa previsão de *tokens* múltiplos, ou seja, prevê os dois próximos *tokens* durante uma execução, e utiliza cadeia de pensamento [Tenable Security Response Team et al. 2025][DeepSeek-AI et al. 2024]. O DeepSeek foi escolhido pelo impacto que causou em seu lançamento, se posicionando como uma opção menos custosa que seus concorrentes.

O Gemini, segundo a sua desenvolvedora Google, é um modelo de inteligência artificial multimodal, que pode operar com diferentes tipos de dados, como texto, imagem, áudio, vídeo e código de programação. Este modelo é altamente flexível e, por isso, pode ser utilizado com eficiência em praticamente qualquer ambiente, seja em um *data center* ou um dispositivo móvel [Google et al. 2023]. O critério de escolha do Gemini neste trabalho se dá pelo fato de ser desenvolvido pelo Google, empresa de relevância mundial.

Uma vez que o uso dos LLMs na Educação é um assunto em discussão e, partindo-se dos modelos selecionados, este trabalho tem como objetivo principal testar e analisar respostas dadas por esses LLMs em relação ao conteúdo de redes de computadores. A intenção é entender o quanto os LLMs atuais podem ajudar estudantes iniciantes em sua formação neste assunto, considerado fundamental da área de tecnologia da informação.

Na próxima seção são apresentados trabalhos encontrados na literatura que possuem relação com este.

2.1. Trabalhos Relacionados

Esta seção apresenta uma revisão da literatura que aborda a utilização dos LLMs no ensino de assuntos relacionados com a Ciência da Computação, como a programação e as redes de computadores. Este conjunto de estudos, que se relaciona com o presente trabalho, fornece um panorama do cenário atual e evidencia as análises propostas. Para alcançar a bibliografia desejada, foram utilizados os sistemas de busca Google Scholar e acervos da ACM Digital Library e IEEE Xplore. Além disso, alguns dos termos utilizados nos *prompts* de pesquisa nessas bases continham a combinação dos seguintes termos: “LLMs”, “redes de computadores”, “educação”, “ensino”, entre outros correlatos. A seguir, os trabalhos mais relevantes obtidos são apresentados.

Em [Donadel et al. 2024], os autores avaliam a capacidade de alguns LLMs de compreenderem conteúdos de redes de computadores. Durante a pesquisa, foram feitas perguntas relacionadas à topologia de redes, com o intuito de explorar o potencial dos modelos em atuarem como assistentes dos administradores de sistemas. Um diferencial que pode ser citado em relação ao presente trabalho é que este analisa conteúdos introdutórios das redes de computadores e o trabalho citado analisa o uso dos LLMs como assistentes de administradores, o que exige um conhecimento mais especializado.

Em [Filho et al. 2023], os autores investigaram a qualidade das respostas dadas pelo ChatGPT em relação a conteúdos básicos da programação de computadores, de forma quantitativa e qualitativa. Em comparação, o atual trabalho tem como foco conteúdos básicos de redes de computadores e são utilizados quatro LLMs diferentes. Porém, o trabalho se relaciona com esse pela forma de avaliação dos LLMs.

Os autores de [Wang et al. 2024] abordam a dificuldade dos LLMs em responder a questões específicas da Ciência da Computação. Para experimentar uma possível solução, é construído um *dataset*, fundamentado em perguntas retiradas de bancos de dados universitários, e, em seguida, é feito o *fine-tuning* dos modelos estudados, visando aprimorar seu desempenho. O trabalho considera assuntos mais abrangentes da área, diferente deste trabalho, no qual tem o foco nas redes de computadores. O trabalho também utiliza questões propostas aos LLMs, assim como o presente trabalho.

Em [Lyu et al. 2024], os autores desenvolvem um LLM próprio, chamado *Code-Tutor*, para que os alunos de disciplinas introdutórias do curso de Ciência da Computação possam utilizá-lo. Posterior ao uso dos estudantes, é realizada uma medição da eficácia do modelo, baseando-se na experiência de cada grupo de controle. O trabalho considera a programação de computadores, contexto diferente do presente artigo, mas também analisa o potencial de tutor dos LLMs.

Em [Raihan et al. 2025], os autores analisam mais de 100 artigos científicos para compreender o impacto do uso dos LLMs no ensino da Ciência da Computação. Durante a análise, são apresentados dados sobre as áreas de concentração desses estudos, bem como as linguagens de programação, metodologias de pesquisa e modelos de linguagem mais utilizados. O foco do trabalho é mais amplo que este, que tem como objetivo analisar um conteúdo específico.

Em [Dakshit 2024], o autor desenvolve aplicações RAG (*Retrieval-Augmented Generation*), com o intuito de capacitar o LLM com fontes de conhecimento externas. Posteriormente, é feita uma medição do potencial do modelo na educação superior de Ciência da Computação, buscando responder se ele pode ser utilizado como assistente virtual para alunos e auxiliar o corpo docente, o que difere do presente trabalho.

Os autores de [Arora et al. 2025] focam em examinar o uso de LLMs por estudantes de disciplinas avançadas de programação (graduação e pós-graduação). De forma mais específica, são identificados os padrões de uso, além de avaliar a eficácia das respostas e compreender as percepções dos estudantes sobre a influência dos modelos de linguagem. O atual trabalho se difere do citado, uma vez que considera somente o conteúdo relacionado às redes de computadores.

Por fim, em [Ma and Wang 2024], os autores desenvolvem um *framework* que utiliza agentes de inteligência artificial para criar materiais educacionais de redes de computadores. A criação desses materiais de estudo aprimorados é fundamentada em experiências de aprendizagem, em que ciclos iterativos de interações entre os agentes de IA simulam interações de salas de aula reais. Diferentemente do presente trabalho, este utiliza os LLMs de forma controlada e supervisionada para a geração dos materiais de estudo. O presente trabalho analisa como são as respostas provenientes de LLMs, sem supervisão ou determinação de *prompts* que definiriam o comportamento e tipo de resposta gerada.

O presente trabalho considera conceitos pouco explorados pela maioria de seus relacionados, uma vez que tem como foco principal analisar se os LLMs utilizados têm condições de responder a questões sobre tópicos básicos de redes de computadores.

3. Metodologia

A metodologia utilizada neste trabalho configura-se como quantitativa experimental [Rangel et al. 2018], que consiste em responder perguntas de pesquisa sobre um contexto por meio de testes e experimentos. Neste trabalho, os testes consistiram em submeter questões aos LLMs, relacionadas com tópicos introdutórios de redes de computadores e obter suas respostas que foram contabilizadas em relação a erros e acertos.

De forma a orientar os objetivos deste trabalho, foram consideradas as seguintes questões de pesquisa:

- QP1 - os LLMs testados são aptos a responderem questões sobre tópicos básicos de redes de computadores?
- QP2 - os LLMs testados podem ajudar estudantes iniciantes em redes de computadores em seus estudos?

Para a definição das questões que seriam usadas, inicialmente foram definidos tópicos de redes de computadores considerados básicos ou essenciais. Esses tópicos são alguns daqueles apresentados em disciplinas introdutórias sobre o assunto e podem ser encontrados nos capítulos iniciais de referências clássicas, como em [Tanenbaum and Wetherall 2011]. Os tópicos definidos foram:

- Domínios e dispositivos: neste tópico são definidos o que são domínios de colisão, de *broadcast*, entre outros, assim como a forma em que os dispositivos que compõem uma rede interagem entre si;

- Topologias de redes: são as diferentes formas de organizar os dispositivos em uma rede e de estabelecer a comunicação entre eles. Algumas topologias são mais comuns que outras e, inclusive, possuem pontos fortes e pontos fracos;
- Modelo OSI (*Open Systems Interconnection*): é um modelo de referência de sete camadas que padroniza as funcionalidades necessárias para a comunicação em uma rede de computadores;
- Arquitetura TCP/IP (*Transmission Control Protocol/Internet Protocol*): trata-se de um conjunto de protocolos implementados e utilizados na Internet e, assim como o modelo OSI, é estruturado em camadas, cada qual com sua funcionalidade;
- Protocolos e endereçamento IP: como definido na arquitetura TCP/IP, as camadas possuem diferentes protocolos, cada qual com suas características e propósito, e o endereçamento IP diz respeito à forma como os dispositivos em uma rede são endereçados e localizados.

Após a definição dos tópicos considerados fundamentais sobre redes de computadores, foram realizadas buscas por fontes de questões objetivas relacionadas a esses tópicos. Nesse processo, optou-se por utilizar questões obtidas em provas de concursos públicos, relacionadas com os conteúdos abordados. A escolha justifica-se por se tratarem de fontes públicas de questões e possuem gabarito oficial, validado por especialistas. Foram usadas questões de provas de concursos para diversos órgãos e instituições públicas, tais como Institutos Federais, Dataprev, universidades estaduais, prefeituras, entre outros. As provas eram de concursos para vagas em diversos cargos tais como analistas de tecnologia da informação (TI), analistas de sistemas, suporte e infraestrutura, administrador de redes e professor do ensino básico técnico e tecnológico, realizadas de 2012 até 2025.

As questões objetivas selecionadas foram utilizadas de três formas: em sua forma original; em uma versão reescrita, também objetiva; e em uma versão reescrita, porém de forma dissertativa¹. Essa variação foi feita para que fosse averiguado o desempenho dos LLMs com variações de uma mesma questão, uma vez que as questões originais foram obtidas de provas de concursos públicos, que se encontram disponíveis *online*, inclusive com gabarito oficial e que podem ter sido usadas em treinamentos dos próprios LLMs, uma vez que esses utilizam dados da *web* pública em seus *datasets*.

A reescrita das questões foi feita utilizando também um LLM para otimizar a tarefa, no caso o Copilot. A escolha deste LLM para essa tarefa não possui motivo específico, podendo ser executada também com outros. Para a reescrita das questões objetivas, foram dados os seguintes passos:

- Passo 1: foi utilizado o seguinte *prompt* “*reescreva a pergunta objetiva abaixo, de modo que a resposta continue sendo a mesma e que o conteúdo desta não seja alterado*”;
- Passo 2: cada resposta gerada era analisada, verificando se a pergunta sofreu alteração em relação ao conteúdo e/ou à resposta. Se a resposta estivesse adequada, a reescrita da questão era aceita. Caso contrário, seguia-se para o próximo passo;

¹Todas as questões utilizadas, tanto as originais, quanto suas versões reescritas, podem ser encontradas em <http://bit.ly/4ID4Tf5>.

- Passo 3: era usado o *prompt* “gere uma nova versão da questão objetiva submetida, de forma que não haja mudança no conteúdo da questão e a resposta desta questão deve ser a mesma da original”. Então, voltava-se ao Passo 2.

Para a reescrita das questões objetivas em formato de questão dissertativa, foram dados os seguintes passos:

- Passo 1: foi utilizado o seguinte *prompt* “reescreva a pergunta objetiva abaixo em formato de pergunta dissertativa, de modo que a resposta continue sendo a mesma e que o conteúdo desta não seja alterado”.
- Passo 2: a resposta obtida era analisada, verificando se a pergunta sofreu alteração em relação ao conteúdo e/ou à resposta. Se a reescrita da questão estivesse adequada, esta era aceita. Caso contrário, seguia-se para o Passo 3.
- Passo 3: era usado o *prompt* “gere uma nova versão dissertativa da questão objetiva submetida, de forma que não haja mudança no conteúdo da questão e a resposta desta questão deve ter o mesmo conteúdo da original”. Voltava-se ao Passo 2.

A seguir é mostrado um exemplo de uma questão objetiva e suas versões reescritas, tanto objetiva quanto discursiva.

Questão original: (*IFTO, vaga para Técnico de Laboratório - Área Informática, concurso realizado em 2024*). *Um determinado técnico foi chamado para auxiliar na construção do documento de Estudo Técnico Preliminar. Durante esse processo, surgiu uma questão sobre a escolha do uso de switches em redes locais, em vez de hubs. Qual é a função principal de um switch em uma rede local e como ele difere de um hub? Alternativas: a) Um switch converte sinais analógicos em digitais, enquanto um hub amplifica o sinal de rede. b) Um switch divide a conexão de internet entre vários dispositivos, enquanto um hub gerencia endereços IP na rede. c) Um switch encaminha pacotes de dados para todos os dispositivos da rede, enquanto um hub encaminha apenas para o dispositivo correto. d) Um switch fornece conexões sem fio para dispositivos móveis, enquanto um hub conecta dispositivos com fio. e) Um switch encaminha pacotes de dados apenas para o dispositivo de destino correto, enquanto um hub envia os dados para todos os dispositivos na rede.* Neste exemplo, a alternativa correta é a letra *e*.

Questão objetiva reescrita: *Durante a elaboração de um Estudo Técnico Preliminar, um técnico foi chamado para colaborar e, nesse processo, surgiu uma dúvida sobre a escolha entre o uso de switches e hubs em redes locais. Considerando essa situação, qual é a função principal de um switch em uma rede local e de que forma ele se diferencia de um hub?* As alternativas foram mantidas iguais às da pergunta original, de forma a manter válido o mesmo gabarito oficial da questão.

Questão reescrita de forma discursiva: *Explique como um switch se diferencia de um hub. Em sua resposta, analise o impacto que cada dispositivo possui na eficiência da comunicação.*

Cada um dos grupos de questões (objetivas originais, objetivas reescritas e reescritas discursivas) foram submetidas a cada um dos LLMs selecionados, sem nenhuma orientação adicional em relação ao tipo de resposta que deveria ser gerada, e essas eram registradas, sendo que cada pergunta foi submetida uma vez para cada LLM. Em relação às questões objetivas, tanto as originais quanto as reescritas, foi analisado se a alternativa

correta tinha sido apontada pelo LLM e registrado se estava correta ou errada. Para as questões discursivas, a resposta era analisada em seu conteúdo, verificando se esta estava correta, errada ou parcialmente correta (quando a resposta não estava totalmente correta). A descrição detalhada dos testes e resultados obtidos são apresentados na próxima seção.

4. Testes e Resultados

Para os testes, foram obtidas 40 questões de concursos públicos, divididas entre os tópicos da seguinte forma: domínios e dispositivos (5 questões), topologias de redes (4 questões), modelo OSI (9 questões), arquitetura TCP/IP (6 questões) e protocolos e endereçamento (16 questões).

Cada uma das 40 questões foram usadas em sua forma objetiva original, objetiva reescrita e reescrita de forma discursiva, compondo 3 grupos de perguntas diferentes. Portanto, para cada LLM, foram submetidas as 40 questões originais objetivas, as 40 questões reescritas objetivas e as 40 questões reescritas como dissertativas, totalizando 120 submissões para cada LLM.

Para as questões objetivas, tanto as originais quanto as reescritas, era contabilizado se o LLM respondeu com a opção correta, assim como suas explicações. Para as respostas dissertativas, era analisado se a resposta dada pelo LLM continha o conteúdo da alternativa correta em sua versão original objetiva, assim como todas as outras informações contidas em sua resposta.

Os resultados obtidos das questões objetivas originais são apresentados na Tabela 1. É possível verificar que tanto o ChatGPT, o Copilot quanto o Gemini acertaram 100% das questões, somente o DeepSeek errou uma questão. O erro cometido pelo DeepSeek foi em uma questão sobre o modelo OSI, que perguntava a respeito da funcionalidade da camada de rede. Entre as alternativas, a correta era aquela que cita que esta camada provê serviço de entrega de segmentos à camada de transporte. O DeepSeek, em sua explicação, atribuiu essa funcionalidade à camada de enlace, o que não está correto, e apontou como correta a alternativa que cita que a camada de rede proporciona garantia de entrega entre transmissor e receptor. Porém, a camada de rede não oferece essa garantia, ficando a cargo dos protocolos da camada de transporte como, por exemplo, o TCP, utilizado na Internet.

LLM	Corretas/Total	Porcentagem de corretas (%)
ChatGPT	40/40	100%
Copilot	40/40	100%
DeepSeek	39/40	97,5%
Gemini	40/40	100%

Tabela 1. Resultado das questões objetivas originais. Fonte: os autores.

Em relação aos resultados das questões objetivas reescritas, apresentados na Tabela 2, todos os LLMs cometeram erros. O ChatGPT, o CoPilot e o DeepSeek erraram 1 questão cada, enquanto que o Gemini errou 2 questões.

A questão errada do ChatGPT era sobre topologias de redes, na qual são apresentados nomes de topologias de forma numerada (1 - topologia em anel, 2 - em estrela,

3 - em malha, e 4 - barramento), e então são apresentadas frases que apresentam características das topologias. É preciso escolher a alternativa que apresenta corretamente a correspondência entre o nome da topologia com a sua descrição. Na resposta dada pelo ChatGPT, as explicações estavam corretas e, inclusive, é citada a ordem correta, porém, ao indicar a alternativa, foi apresentada uma ordem errada.

A questão errada pelo Copilot era sobre o tópico protocolos e endereçamento, na qual é pedida um cálculo de máscara de rede do protocolo IPv4. A explicação dada pelo Copilot era adequada, porém finaliza indicando a resposta errada.

A questão errada pelo DeepSeek para as reescritas foi a mesma errada nas originais (Tabela 1), seguindo o mesmo padrão e, inclusive, apresentando uma explicação parecida.

O Gemini errou duas questões: uma sobre os protocolos utilizados na infraestrutura da Internet, na qual somente uma alternativa era correta, porém o LLM citou duas como corretas, se equivocando em relação a uma que versava sobre endereçamento IP. Na outra questão errada, era pedido um cálculo de sub-rede e o LLM apresentou um cálculo binário errado, portanto indicando uma alternativa errada.

LLM	Corretas/Total	Porcentagem de corretas (%)
ChatGPT	39/40	97,5%
Copilot	39/40	97,5%
DeepSeek	39/40	97,5%
Gemini	38/40	95%

Tabela 2. Resultado das questões objetivas reescritas. Fonte: os autores.

Em relação aos resultados obtidos das questões reescritas como questões dissertativas, essas foram consideradas como certas, erradas ou parcialmente certas, que era quando a resposta estava errada de acordo com o enunciado, mas possuía elementos corretos, como pode ser visto na Tabela 3. Para essas questões, o ChatGPT gerou uma resposta parcialmente correta, o Copilot errou uma questão, o DeepSeek não cometeu erros e o Gemini teve uma resposta considerada parcialmente correta.

Na resposta considerada parcialmente correta do ChatGPT, era preciso fazer um comparativo entre as camadas do modelo OSI com as camadas da arquitetura TCP/IP, apresentando as funcionalidades destas. A resposta do ChatGPT não estava errada, porém, não fez um comparativo em relação às funcionalidades das camadas especificamente, somente em termos gerais.

O Copilot respondeu de forma errada uma questão que apresentava 4 afirmações e era preciso explicar cada uma delas, informando se é verdadeira ou falsa. O Copilot informou erroneamente que uma afirmação correta sobre a representação de endereços IPv4 em IPv6 estava errada.

O Gemini teve uma resposta considerada parcialmente correta. A questão pedia para apresentar a principal vantagem de se utilizar a arquitetura TCP/IP. Na sua versão original objetiva, a opção correta remete ao fato desses protocolos permitir transmissões eficientes e confiáveis. Na resposta do Gemini, apesar de estar correta, baseou-se em

motivos mais gerais, não especificando a eficiência e confiabilidade, que podem ser consideradas características principais do TCP/IP.

LLM	Corretas/Total	Porcentagem de corretas (%)
ChatGPT	39/40 (1 parcialmente)	97,5%
Copilot	39/40	97,5%
DeepSeek	40/40	100%
Gemini	39/40 (1 parcialmente)	97,5%

Tabela 3. Resultado da questões reescritas de forma dissertativa. Fonte: os autores.

Em relação às Tabelas 1, 2 e 3, é mostrado que não houve grande alteração no desempenho dos LLMs em relação à variação das questões em suas formas (original, reescrita e dissertativa). Para a versão original das questões, o número de erros foi menor, sendo cometido somente um erro por um dos LLMs, e para as outras versões, outros erros aconteceram, porém esses não foram significativamente diferentes.

No total, cada LLM respondeu 3 grupos de questões, sendo um grupo com as questões objetivas originais (40 questões), um grupo de questões objetivas reescritas (40 questões) e um grupo de questões reescritas em forma dissertativa (40 questões), totalizando 120 questões. É possível verificar que todos os LLMs tiveram resultados parecidos, sendo que o ChatGPT, Copilot e DeepSeek erraram 2 questões no total geral e o Gemini errou 3, com resultados muito próximos se considerada a porcentagem de acertos.

Na Tabela 4 são apresentados os resultados agregados de todas as questões submetidas aos LLMs, divididos pelos tópicos de redes de computadores definidos neste trabalho. Em relação aos tópicos “domínios e dispositivos”, todos os LLMs tiveram 100% de acerto; em “topologias de redes”, somente o ChatGPT errou uma das questões; sobre “modelo OSI”, somente o DeepSeek errou duas questões; a respeito da “arquitetura TCP/IP”, o ChatGPT e o Gemini erraram uma questão cada; sobre “protocolos e endereçamentos”, tanto o Copilot quanto o Gemini erraram duas questões cada. Dessa forma, não houve destaque entre os LLMs em relação a algum tópico específico, uma vez que os resultados ficaram bem próximos entre eles, considerando o número de questões testadas e a porcentagem de acerto.

Tópico/LLM	ChatGPT	Copilot	DeepSeek	Gemini
Domínios e Dispositivos	15/15 (100%)	15/15 (100%)	15/15 (100%)	15/15 (100%)
Topologias de Redes	11/12 (91,67%)	12/12 (100%)	12/12 (100%)	12/12 (100%)
Modelo OSI	27/27 (100%)	27/27 (100%)	25/27 (92,59%)	27/27 (100%)
Arquitetura TCP/IP	17/18 (94,44%)	18/18 (100%)	18/18 (100%)	17/18 (94,44%)
Protocolos e Endereçamento	48/48 (100%)	46/48 (95,83%)	48/48 (100%)	46/48 (95,83%)
Total	118/120 (98,33%)	118/120 (98,33%)	118/120 (98,33%)	117/120 (97,50%)

Tabela 4. Desempenho dos LLMs por tópico. Fonte: os autores.

4.1. Discussão dos Resultados

Como comentário geral, é importante citar que, todas as vezes que os LLMs geravam respostas corretas, essas eram completas e possuíam explicações de qualidade. Em relação às questões objetivas, é interessante pontuar que, além de indicar a alternativa correta, os LLMs explicam a razão daquela ser a certa e as outras serem erradas. Considerando um estudante iniciante, esse comportamento é positivo, possibilitando o aprendizado e mostrando o potencial pedagógico da ferramenta, pois esta não se restringe a apenas apresentar a resposta correta. Esse comportamento demonstra o potencial dos LLMs como auxiliares em atividades propostas aos alunos, demandando uma análise das respostas obtidas, o que pode gerar impactos positivos em relação ao processo de aprendizagem.

Entretanto, como foi possível observar, os LLMs podem gerar respostas erradas e informações imprecisas, demandando a checagem daquilo que é apresentado, assim como a supervisão e orientação docente. Questões éticas também devem ser consideradas, uma vez que estudantes podem usar as respostas geradas pelos LLMs sem questionamento crítico, gerando questões sobre autoria e possivelmente causando falha de aprendizagem.

Na seção anterior, foram apresentados os resultados obtidos ao submeter questões de tópicos introdutórios de redes de computadores aos LLMs. Os resultados obtidos com cada LLM são parecidos, sendo que de um total de 120 questões (somando-se as questões dos 3 grupos), a variação estatística entre os LLMs é baixa, sendo uma diferença de apenas 0,83% entre aqueles que acertaram 118 das 120 questões para o LLM que acertou 117 questões.

Em relação aos tópicos de redes, não houve nenhum que se destacasse de forma negativa. O máximo de questões erradas por tópico foi duas, sendo que o DeepSeek errou duas questões sobre o modelo OSI, e o Copilot e o Gemini erraram duas questões sobre protocolos e endereçamento. Todos os outros tópicos ou tiveram 100% de aproveitamento ou, no máximo, uma questão foi respondida de forma errada.

Dessa forma, considerando os testes realizados e a análise feita neste trabalho, é possível responder as questões de pesquisa da seguinte forma: QP1 - sim, os LLMs testados são aptos a responder questões sobre tópicos básicos de redes de computadores, entretanto, é possível que cometam erros eventuais, como esperado pela própria natureza estatística dos LLMs; QP2 - sim, os LLMs apresentam potencial para ajudar estudantes iniciantes em redes de computadores, porém é interessante que seja um uso supervisionado e que os estudantes tenham o hábito da checagem em relação às respostas obtidas, para mitigarem os erros que podem acontecer. Testes com estudantes iniciantes reais podem confirmar ainda mais essa afirmação.

5. Limitações e trabalhos futuros

Alguns pontos que podem ser citados como limitações do trabalho: i) foram analisadas somente questões a respeito de tópicos básicos de redes de computadores, não sendo analisadas respostas sobre assuntos considerados mais avançados; ii) a análise das respostas foi feita pelos autores, e não necessariamente refletem a percepção, de fato, de estudantes iniciantes; iii) foram submetidas aos LLMs quantidades diferentes de questões para cada tópico, o que pode influenciar na taxa de acertos e erros em determinados assuntos.

Como trabalhos futuros é possível realizar uma análise a respeito das respostas dos LLMs em relação a tópicos avançados de redes de computadores, assim, é possível verificar se o comportamento dos LLMs se mantém. É possível também realizar testes com estudantes, propondo o uso dos LLMs em atividades definidas e, posteriormente, coletar suas opiniões e impressões. O resultado obtido neste trabalho também propicia uma reflexão prática em relação ao tipo de atividades que devem ser propostas, considerando o uso dos LLMs no ensino.

6. Conclusão

Este trabalho teve como objetivo analisar as respostas dadas por quatro LLMs em relação ao conteúdo introdutório de redes de computadores. Esse tipo de análise é importante, uma vez que estudantes iniciantes, de qualquer área, utilizam cada vez mais essa tecnologia em seus estudos. Para isso, foram coletadas 40 questões objetivas sobre redes de computadores, provenientes de concursos públicos. As questões foram utilizadas em três versões, totalizando 120 questões submetidas a cada LLM. O resultado mostrou que os LLMs possuem um bom desempenho ao responder essas questões, gerando uma taxa de sucesso igual ou superior a 97,5%. Conclui-se, portanto, que os LLMs possuem grande potencial como auxiliares no aprendizado das questões introdutórias de redes de computadores, entretanto, para uma experiência totalmente proveitosa, a supervisão docente é recomendada.

Referências

Arora, U., Garg, A., Gupta, A., Jain, S., Mehta, R., Oberoi, R., Prachi, Raina, A., Saini, M., Sharma, S., Singh, J., Tyagi, S., and Kumar, D. (2025). Analyzing llm usage in an advanced computing class in india. In *Proceedings of the 27th Australasian Computing Education Conference, ACE '25*, page 154–163, New York, NY, USA. Association for Computing Machinery.

- Dakshit, S. (2024). Faculty perspectives on the potential of rag in computer science higher education. In *Proceedings of the 25th Annual Conference on Information Technology Education, SIGITE '24*, page 19–24, New York, NY, USA. Association for Computing Machinery.
- DeepSeek-AI, Liu, A., Feng, B., Xue, B., Wang, B., Wu, B., Lu, C., Zhao, C., Deng, C., Zhang, C., Ruan, C., Dai, D., Guo, D., Yang, D., Chen, D., Ji, D., Li, E., Lin, F., Dai, F., and Pan, Z. (2024). Deepseek-V3 technical report. *arXiv preprint* arXiv:2412.19437. Disponível em: <https://arxiv.org/abs/2412.19437>.
- Donadel, D., Marchiori, F., Pajola, L., and Conti, M. (2024). Can LLMs Understand Computer Networks? Towards a Virtual System Administrator . In *2024 IEEE 49th Conference on Local Computer Networks (LCN)*, pages 1–10, Los Alamitos, CA, USA. IEEE Computer Society.
- Filho, L. P., Souza, T., and Paula, L. (2023). Análise das respostas do chatgpt em relação ao conteúdo de programação para iniciantes. In *Anais do XXXIV Simpósio Brasileiro de Informática na Educação*, pages 1738–1748, Porto Alegre, RS, Brasil. SBC.
- Google, Pichai, S., and Hassabis, D. (2023). Introducing gemini: our largest and most capable ai model. <https://blog.google/technology/ai/google-gemini-ai/#introducing-gemini>. Acessado em 7 de junho de 2025.
- Kurose, J. F. and Ross, K. W. (2021). *Redes de Computadores e a Internet (coedição Bookman e Pearson)*. Bookman Editora.
- Lo, C. K. (2023). What is the impact of chatgpt on education? a rapid review of the literature. *Education Sciences*, 13:410.
- Lyu, W., Wang, Y., Chung, T. R., Sun, Y., and Zhang, Y. (2024). Evaluating the effectiveness of llms in introductory computer science education: A semester-long field study. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale, L@S '24*, page 63–74, New York, NY, USA. Association for Computing Machinery.
- Ma, X. and Wang, J. (2024). WIP: Active Learning Through Prompt Engineering and Agentic AI Simulation-A Pilot Project in Computer Networks Education . In *2024 IEEE Frontiers in Education Conference (FIE)*, pages 1–5, Los Alamitos, CA, USA. IEEE Computer Society.
- Marques, D. and Morandini, M. (2024). Uso do chatgpt no contexto educacional: Uma revisão sistemática da literatura. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*, pages 1784–1795, Porto Alegre, RS, Brasil. SBC.
- Microsoft (2025). What is microsoft copilot? <https://www.microsoft.com/en-us/microsoft-copilot/copilot-101/what-is-copilot>. Acessado em 7 de junho de 2025.
- Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., and Gao, J. (2025). Large language models: A survey.
- OpenAI (2025). How chatgpt and our foundation models are developed. <https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed>. Acessado em 7 de junho de 2025.
- Raihan, N., Siddiq, M. L., Santos, J. C., and Zampieri, M. (2025). Large language models in computer science education: A systematic literature review. In *Proceedings of the*

56th ACM Technical Symposium on Computer Science Education V. 1, SIGCSETS 2025, page 938–944, New York, NY, USA. Association for Computing Machinery.

Rangel, M., Rodrigues, J., and Mocarzel, M. (2018). Fundamentos e princípios das opções metodológicas. metodologias quantitativas e procedimentos quali-quantitativos de pesquisa. *OMNIA Revista Interdisciplinar de Ciências e Artes*, 8:5–11.

Soares, B. J., Franco, D., Sabino, B., and Eguchi, M. (2023). Implicações da inteligência artificial na educação. *TECCOGS – Revista Digital de Tecnologias Cognitivas*, pages 76–86.

Tanenbaum, A. S. and Wetherall, D. J. (2011). *Redes de Computadores*. Pearson, São Paulo, SP, 5 edition. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 22 ago. 2025.

Tenable Security Response Team, Narang, S., and Miles, N. (2025). Frequently asked questions about deepseek large language model (llm). <https://www.tenable.com/blog/frequently-asked-questions-about-deepseek-large-language-model-llm-v3-r1>. Acessado em 7 de junho de 2025.

Wang, B., Zhang, Y., and Zeng, Q. (2024). Resources Construction and LLM Fine-tuning for Education of Computer Science . In *2024 IEEE International Conference on Big Data (BigData)*, pages 1655–1663, Los Alamitos, CA, USA. IEEE Computer Society.