Laudo Oficial de Correção — Prova Teórica 1

Inteligência Artificial II (IA II) ▪ CEFET-MG 2026.2

Prof. Aléssio Miranda Júnior

2026-10-19

Dados do Discente

- **Discente:** Daiana Maria de Oliveira Vieira
- **Matrícula:** 20233004182
- **Turma:** Turma 2
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 24,4 / 25,0** (Aproveitamento: **97,6%**)
- **Legibilidade da Caligrafia:** 86% (Média Legibilidade)

Desempenho Consolidado por Questão

Questão / Tema	Pontos Obtidos	Valor	Aproveitamento
Q1: Triagem com dados heterogêneos	**5,0**	5,0	100,0%
Q2: Inspeção visual em condições novas	**4,8**	5,0	96,0%
Q3: Alerta que muda decisões	**4,9**	5,0	98,0%
Q4: Busca de conhecimento para atendimento	**4,8**	5,0	96,0%
Q5: Assistente que pode não responder	**4,9**	5,0	98,0%
TOTAL GERAL	**24,4**	**25,0**	**97,6%**

Questão 1 — Triagem com dados heterogêneos (5,0 / 5,0 pts)

1.a Diagnóstico da Curva de Aprendizado e Decisão de Produção (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Avaliação incorreta, a curva de validação ao subir indica overfitting impedindo o uso em produção."

Parecer do Docente

Resposta precisa, direta e irrepreensível. Identificou a subida da perda na validação, diagnosticou sobreajuste (overfitting) e vetou categoricamente a entrada em produção.

1.b Representação de Entradas Heterogêneas e Retropropagação (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Numéricas (idade, sinais vitais) normalização ou padronização: min-max ou Z-score. Categórica (questionário) One Hot Encoding ou Embeddings. A retropropagação calcula gradientes para poder ajustar os pesos."

Parecer do Docente

Excelente domínio técnico. Especificou corretamente Min-Max / Z-score para atributos numéricos, One-Hot Encoding ou Embeddings para categóricos e sintetizou o papel exato do algoritmo de backpropagation (cálculo iterativo de gradientes para ajuste ótimo de pesos).

1.c Evidências Clínicas de Confiança e Reprodutibilidade (1,0 / 1,0)

Resposta Transcrita do Aluno

> "Usar conjunto de teste cego, métricas clínicas e fixar random."

Parecer do Docente

Respondeu com completude os dois pilares exigidos pelo gabarito: partição independente de teste cego avaliada sob métricas clínicas e reprodutibilidade via fixação de sementes aleatórias ("fixar random").

Questão 2 — Inspeção visual em condições novas (4,8 / 5,0 pts)

2.a Diagnóstico Experimental do Aumento de Alertas (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Covariate shift muda o brilho e a iluminação, o overfitting tem alto erro na validação. A Rotulação tem inconsistência de anotação."

Parecer do Docente

Perfeito! Discriminou com exatidão as três causas concorrentes solicitadas: deslocamento de covariáveis (covariate shift por variação luminosa), sobreajuste (overfitting comprovável por erro em validação) e inconsistência/ruído na anotação dos rótulos pelos inspetores.

2.b Arquitetura Convolucional e Transferência de Aprendizado (1,8 / 2,0)

Resposta Transcrita do Aluno

> "Congelar camadas iniciais da CNN e ajustar as finais (Fine-Tuning), Augmentação com alteração de brilho e o uso de Dropout."

Parecer do Docente

Solução de engenharia impecável: estratégia de congelamento com ajuste fino (fine-tuning) das camadas densas, aumento de dados (data augmentation) focado em variações de brilho e regularização via dropout. Pequena dedução apenas por não ter esquematizado visualmente a arquitetura solicitada no enunciado.

2.c Cautela Amostral e Validação com Poucos Dados (1,0 / 1,0)

Resposta Transcrita do Aluno

> "Validação cruzada estratificada e Testes estatísticos de significância."

Parecer do Docente

Citou com rigor metodológico a validação cruzada estratificada e a exigência de testes estatísticos de significância para atestar robustez antes da homologação fabril.

Questão 3 — Alerta que muda decisões (4,9 / 5,0 pts)

3.a Interpretação da Calibração e Crítica ao Limiar 0,50 (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Falha da direção, o modelo é superconfiante. Para a previsão de 0,50, a frequência real de piora é apenas de 18% (82% são alarmes falsos). O impacto clínico são as probabilidades previstas não refletem o risco real, gerando fadiga de alarmes e perda de confiança na equipe médica."

Parecer do Docente

Brilhante! Extraiu o dado exato do gráfico da Figura 2 (18% real para $\hat{p} = 0{,}50$), calculou a taxa real de alarmes falsos (82%), apontou a superconfiança e utilizou a terminologia clínica consagrada: fadiga de alarmes (alarm fatigue) e perda de credibilidade do sistema.

3.b Custos Assimétricos e Capacidade da Equipe Médica (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Falso negativo é grave (risco à saúde). Reduzir o limiar para priorizar Recall, respeitando a capacidade máxima médica."

Parecer do Docente

Equilíbrio perfeito entre risco médico e restrição operacional. Identificou a gravidade extrema do falso negativo, recomendou diminuir o limiar de decisão para maximizar a sensibilidade (recall) e enfatizou a necessidade de respeitar a capacidade máxima de atendimento médico.

3.c Telemetria Contínua e Deriva de Conceito (0,9 / 1,0)

Resposta Transcrita do Aluno

> "Testes estatísticos e taxa de discordância. Ação seria acionar fallback e re-treinar."

Parecer do Docente

Muito bom. Indicou monitoramento via taxa de discordância com auditoria médica e plano de contingência robusto (fallback operacional seguro seguido de retreino do modelo).

Questão 4 — Busca de conhecimento para atendimento (4,8 / 5,0 pts)

4.a Representação Semântica e Limitações da Similaridade de Cosseno (1,8 / 2,0)

Resposta Transcrita do Aluno

> "O Word2vec é estático, os Transformers são contextualizados. O cosseno mede somente a proximidade temática global, sem garantir a relevância."

Parecer do Docente

Excelente síntese teórica: diferenciou os embeddings estáticos por palavra (Word2Vec) dos modelos de atenção contextualizados (Transformers) e diagnosticou a miopia da similaridade de cosseno, que capta coocorrência ou proximidade temático-lexical, mas não atesta suficiência informativa ou relevância legal.

4.b Avaliação Rigorosa de Recuperação (Information Retrieval) (1,5 / 1,5)

Resposta Transcrita do Aluno

> "O Retorno não garante acerto. Então avaliar com métricas de IR: MRR, NDCG e Recall."

Parecer do Docente

Resposta precisa. Refutou a falácia de que "trazer sempre K documentos garante acerto" e listou as três principais métricas padrão de Information Retrieval: MRR (Mean Reciprocal Rank), NDCG (Normalized Discounted Cumulative Gain) e Recall@K sobre base rotulada.

4.c Busca Híbrida e Governança Normativa (1,5 / 1,5)

Resposta Transcrita do Aluno

> "Combinar BM25 com vetores densos. E filtrar por metadados."

Parecer do Docente

Pontuação integral. Especificou a arquitetura de busca híbrida unindo recuperação lexical exata (BM25) com vetores densos semânticos, além de instituir filtros determinísticos por metadados (como vigência/revogação) para garantir conformidade normativa.

Questão 5 — Assistente que pode não responder (4,9 / 5,0 pts)

5.a Contestação do Maior Escore e Desenho da Abstenção (2,0 / 2,0)

Resposta Transcrita do Aluno

> "Eu penso que é errado mostrar sempre o maior escore (contesto). Abster-se se o escore $< \tau_1$, ou se $(\text{score}_1 - \text{score}_2) < \tau_2$ é uma ambiguidade."

Parecer do Docente

Brilhante formulação matemática! Estabeleceu o duplo critério canônico de abstenção: limiar absoluto de confiança mínima ( ) para evitar alucinações sobre baixa aderência, e margem de separação mínima ( ) para desarmar ambiguidades entre os dois candidatos mais prováveis.

5.b Conflito de Fontes e Validação Prévia da Abstenção (1,9 / 2,0)

Resposta Transcrita do Aluno

> "Priorizar normas recentes via metadados. Validar com testes contendo cenários conflitantes ou sem resposta."

Parecer do Docente

Excelente raciocínio de governança. Propôs desempate normativo fundamentado em metadados temporais (norma mais recente) e desenhou uma suíte de testes sintéticos contendo propositalmente casos conflitantes e perguntas fora do domínio (OOD) para auditar a taxa de abstenção antes do lançamento.

5.c Telemetria da Taxa de Abstenção em Produção (1,0 / 1,0)

Resposta Transcrita do Aluno

> "Monitorar taxa de abstenção. Aumento indica novas dúvidas de utilizadores ou normas desatualizadas/conflitantes no acervo."

Parecer do Docente

Visão sistêmica completa: correlacionou o salto na taxa de abstenção em produção tanto com o surgimento de novas dúvidas operacionais dos usuários quanto com a defasagem ou conflito de normas dentro da base de conhecimento indexada.

Laudo gerado oficialmente pela disciplina Inteligência Artificial II (IA II) • CEFET-MG Timóteo