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:** Thiago Victor de Jesus Moura
- **Matrícula:** `20233004511`
- **Turma:** Turma 1
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 16,0 / 25,0** (Aproveitamento: **64,0%**)
- **Legibilidade da Caligrafia:** 89% (Boa Legibilidade)

Desempenho Consolidado por Questão

Questão / Tema	Pontos Obtidos	Valor	Aproveitamento
Q1: Triagem com dados heterogêneos	**4,6**	5,0	92,0%
Q2: Inspeção visual em condições novas	**3,1**	5,0	62,0%
Q3: Alerta que muda decisões	**2,6**	5,0	52,0%
Q4: Busca de conhecimento para atendimento	**2,8**	5,0	56,0%
Q5: Assistente que pode não responder	**2,9**	5,0	58,0%
TOTAL GERAL	**16,0**	**25,0**	**64,0%**

Questão 1 — Triagem com dados heterogêneos (4,6 / 5,0 pts)

1.a Diagnóstico da Curva de Aprendizado e Decisão de Produção (1,9 / 2,0)

Resposta Transcrita do Aluno

> "Não está pronto pois há a presença de overfitting, observado pela queda contínua do loss de treino e a separação e crescimento do loss de validação a partir da época 8, indica que o modelo começou a decorar os dados."

Parecer do Docente

Diagnóstico cirúrgico: identificou a divergência na Época 8 com o loss de validação em elevação contínua como marca de sobreajuste/memorização e vetou prontamente a ida para produção.

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

Resposta Transcrita do Aluno

> "Para as features numéricas, utilizar Standard Scaler. Aplicando o fit_transform no treino e transform no teste. Para as features categóricas, One Hot Encoding. Backpropagation é um algoritmo de otimização do gradiente para atualização dos pesos, utiliza a regra da cadeia para calcular o gradiente com base na função loss."

Parecer do Docente

Nota integral. Resposta técnica primorosa: StandardScaler com fit_transform no treino e transform no teste para evitar vazamento, One-Hot Encoding para categóricos e backpropagation com cálculo de gradientes via regra da cadeia.

1.c Evidências Clínicas de Confiança e Reprodutibilidade (0,7 / 1,0)

Resposta Transcrita do Aluno

> "Utilização de métricas como Recall, Acurácia, F1-Score e também matriz de confusão. Para a reprodução do experimento, é essencial salvar os pesos do modelo."

Parecer do Docente

Citou métricas clínicas relevantes (Recall, F1) e o salvamento de pesos para persistência. Faltou mencionar a fixação de sementes aleatórias (random seeds) e uma base independente de teste cego.

Questão 2 — Inspeção visual em condições novas (3,1 / 5,0 pts)

2.a Diagnóstico Experimental do Aumento de Alertas (1,0 / 2,0)

Resposta Transcrita do Aluno

> "Utilizar imagens com diferentes luminosidades, fundo de imagens diferente (indicará se o modelo decorou o fundo). Utilizar peças em posições diferentes, pois o modelo pode ter decorado a posição da peça."

Parecer do Docente

Excelente intuição prática sobre testes com variações de luminosidade, fundo e posicionamento para detectar memorização de artefatos espúrios. Faltou categorizar as três hipóteses concorrentes (covariate shift, sobreajuste e ruído de anotação).

2.b Arquitetura Convolucional e Transferência de Aprendizado (1,6 / 2,0)

Resposta Transcrita do Aluno

> "A CNN é indicada por utilizar filtros deslizantes sobre a imagem, ao invés de conectar todos os pixels a todos os neurônios. Utilizar um modelo já pré-treinado é interessante pelo fato de já haver conhecimento adquirido, sendo necessário uma nova camada apenas para classificar a imagem mais específicas."

Parecer do Docente

Muito boa explicação sobre a vantagem de filtros convolucionais locais (compartilhamento de pesos) e uso de transfer learning aproveitando conhecimento prévio com ajuste da camada classificadora final.

2.c Cautela Amostral e Validação com Poucos Dados (0,5 / 1,0)

Resposta Transcrita do Aluno

> "Realizar novo treinamento com as imagens noturnas é uma opção, tendo uma nova base de treino e testes para cenários específicos noturnos."

Parecer do Docente

Propôs partição dedicada para o cenário noturno. Omitiu a realização de validação cruzada estratificada com testes estatísticos de hipótese e intervalos de confiança.

Questão 3 — Alerta que muda decisões (2,6 / 5,0 pts)

3.a Interpretação da Calibração e Crítica ao Limiar 0,50 (0,6 / 2,0)

Resposta Transcrita do Aluno

> "Nesse contexto, um classificador mal calibrado é um que classifica uma ocorrência necessária real como sendo falsa."

Parecer do Docente

Definição superficial de má calibração associada a erros de classificação. Não interpretou a curva da Figura 2, onde $\hat{p} = 0{,}50$ corresponde a apenas 18% de frequência real de piora clínica.

3.b Custos Assimétricos e Capacidade da Equipe Médica (1,5 / 2,0)

Resposta Transcrita do Aluno

> "Deve-se priorizar o aumento de verdadeiros positivos e diminuir o número de falsos positivos, proporcionando melhor atendimento a quem realmente precisa. O custo de falsos positivos é ruim, pois estoura a capacidade operacional da equipe."

Parecer do Docente

Excelente correlação com a restrição de capacidade operacional: identificou com clareza que o excesso de alarmes falsos esgota o teto de atendimento da equipe médica.

3.c Telemetria Contínua e Deriva de Conceito (0,5 / 1,0)

Resposta Transcrita do Aluno

> "Acompanhamento frequente das métricas e matrizes de confusão, se necessário, retreinando o modelo com mais dados."

Parecer do Docente

Preconizou acompanhamento contínuo de matrizes de confusão e retreinamento. Faltou detalhar a detecção de concept/data drift e um plano de contingência hospitalar seguro.

Questão 4 — Busca de conhecimento para atendimento (2,8 / 5,0 pts)

4.a Representação Semântica e Limitações da Similaridade de Cosseno (1,3 / 2,0)

Resposta Transcrita do Aluno

> "No Word2Vec palavras são mapeadas para vetores, e palavras relacionadas são próximas entre si com base em seu contexto. Entretanto, o vetor não carrega em si próprio, o seu contexto. A similaridade do cosseno pode falhar ao não considerar totalmente o contexto da palavra."

Parecer do Docente

Compreendeu a limitação do Word2Vec, onde o vetor estático perde o contexto específico da sentença, afetando o cálculo do cosseno. Faltou citar Transformers e a questão da vigência temporal da norma.

4.b Avaliação Rigorosa de Recuperação (Information Retrieval) (0,5 / 1,5)

Resposta Transcrita do Aluno

> "Porque o modelo pode errar e trazer respostas que não fazem sentido com aquilo que foi pedido. É necessário se utilizar a união de métricas específicas como suporte, confiança e lift"

Parecer do Docente

Percebeu a fragilidade do retorno bruto. Contudo, sugeriu métricas típicas de mineração de regras de associação (suporte, confiança, lift) em vez de métricas de ranking em Information Retrieval (MRR, nDCG, Recall@K).

4.c Busca Híbrida e Governança Normativa (1,0 / 1,5)

Resposta Transcrita do Aluno

> "A abordagem híbrida surge com efeito pois, além de considerar o resultado semântico, ela também leva em consideração a aparição léxica de palavras da query em questão."

Parecer do Docente

Definiu corretamente a essência da busca híbrida: combinação de proximidade semântica com correspondência léxica pontual da consulta. Faltou associar filtros por metadados de vigência.

Questão 5 — Assistente que pode não responder (2,9 / 5,0 pts)

5.a Contestação do Maior Escore e Desenho da Abstenção (1,3 / 2,0)

Resposta Transcrita do Aluno

> "É necessário fixar limiares de comparação, por mais que um documento possa ter o maior escore entre todos em uma consulta, não significa que tem uma resposta correta. Se não atingir os limiares, deve haver abstenção."

Parecer do Docente

Contestou adequadamente o critério de maior escore e prescreveu a abstenção fundamentada em limiares. Faltou detalhar a formulação dos dois limiares complementares (absoluto e de margem competitiva).

5.b Conflito de Fontes e Validação Prévia da Abstenção (0,8 / 2,0)

Resposta Transcrita do Aluno

> "Utilizando um validador básico, com métricas de suporte, lift e confiança antes de responder ao usuário, abstenção de resultados que não atingem limiares básicos."

Parecer do Docente

Propôs camada de validação prévia antes do envio da resposta e abstenção seletiva. Faltou estruturar a governança por metadados de vigência/autoridade e banco de testes com consultas propositalmente sem resposta.

5.c Telemetria da Taxa de Abstenção em Produção (0,8 / 1,0)

Resposta Transcrita do Aluno

> "Através de gráficos, por exemplo. Se houver muitas abstenção, é indicativo de que os documentos não fazem sentido de resposta para as perguntas."

Parecer do Docente

Boa percepção de monitoramento: relacionou o aumento da taxa de abstenção à perda de aderência entre os documentos do acervo e as perguntas dos usuários.
