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:** Lucas Barros Meckler
- **Matrícula:** 20233003953
- **Turma:** Turma 1
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 24,8 / 25,0** (Aproveitamento: **99,2%**)
- **Legibilidade da Caligrafia:** 95% (Alta 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	**5,0**	5,0	100,0%
Q4: Busca de conhecimento para atendimento	**5,0**	5,0	100,0%
Q5: Assistente que pode não responder	**5,0**	5,0	100,0%
TOTAL GERAL	**24,8**	**25,0**	**99,2%**

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

> "O overfitting começa na época 8, a perda de treino cai, ao passo que a validação sobe, o que indica uma baixa generalização, sendo preciso usar early stopping e testar o modelo antes de implementar"

Parecer do Docente

Resposta precisa e concisa: identificou a divergência na Época 8, caracterizou o sobreajuste como perda de capacidade de generalização e prescreveu early stopping com teste rigoroso antes da produção.

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

Resposta Transcrita do Aluno

> "Os dados numéricos devem ser normalizados, já os categóricos, codificados com one-hot-encoding ou embeddings. A retropropagação calcula os gradientes da função de perda, o que permite atualizar os pesos da MLP durante o treinamento"

Parecer do Docente

Pontuação integral. Especificou normalização numérica, One-Hot Encoding ou embeddings para variáveis categóricas e sintetizou perfeitamente o papel da retropropagação no cálculo dos gradientes para atualização dos pesos.

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

Resposta Transcrita do Aluno

> "O modelo deve ser avaliado em dados independentes, separados por treino, teste e validação, usando métricas como sensibilidade e especificidade, para garantir a reprodutibilidade, devem ser registrados os dados que foram utilizados, além dos hiperparâmetros e das sementes aleatórias"

Parecer do Docente

Excelente rigor de engenharia: partição independente tripla, métricas clínicas adequadas (sensibilidade/especificidade) e conjunto completo de reprodutibilidade científica (versionamento, hiperparâmetros e sementes aleatórias).

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

> "É preciso comparar o desempenho nos turnos diurno e noturno, o overfitting é indicado pela diferença entre os resultados de treino e validação, já o covariate shift é indicado por mudanças na distribuição de imagens, enquanto as falhas no processo de rotulação são evidenciadas por inconsistências identificadas na revisão manual dos rótulos."

Parecer do Docente

Discriminação experimental primorosa: contrastou os turnos diurno e noturno, mapeando com precisão a evidência distintiva de sobreajuste, mudança na distribuição luminosa (covariate shift) e ruído na revisão manual de anotação.

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

Resposta Transcrita do Aluno

> "Uma CNN seria adequada para identificar padrões espaciais, como defeitos e bordas, pode-se utilizar um modelo pré-treinado, congelando inicialmente suas camadas e realizando fine-tunning posteriormente, sendo que algumas técnicas de data argumentation ajudam a reduzir o overfitting, como variações de iluminação e contraste"

Parecer do Docente

Fundamentação sólida: justificou a CNN para captura de invariância espacial de bordas/defeitos, com transferência de aprendizado (congelamento + fine-tuning) e aumento sintético de dados com foco em contraste e iluminação. Pequeno desconto pela ausência de esquema gráfico.

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

Resposta Transcrita do Aluno

> "Tem que ser utilizado um conjunto de teste independente, além de calcular intervalos de confiança e realizar a validação cruzada, além disso, também é importante coletar mais imagens rotuladas do turno noturno antes de concluir que houve uma melhora"

Parecer do Docente

Padrão ouro de validação: exigiu partição independente, intervalos de confiança estatísticos, validação cruzada e ampliação da amostragem noturna antes de homologar melhorias.

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

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

Resposta Transcrita do Aluno

> "Uma previsão de 0,5 corresponde à uma frequência real de 18%, o que indica que o modelo pode estar superestimando o risco. Dessa forma, está mal calibrado e o limiar não deve ser escolhido apenas com base na interpretação de que probabilidades acima de 0,5 representam maior certeza"

Parecer do Docente

Leitura quantitativa impecável da Figura 2: extraiu a taxa de 18% real para $\hat{p} = 0{,}50$, diagnosticou a superestimação de risco gerada por má calibração e alertou sobre a falácia de tomar probabilidades nominais como certeza absoluta.

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

Resposta Transcrita do Aluno

> "Os falsos positivos geram alarmes desnecessários, já os falsos negativos podem atrasar o atendimento de pacientes em risco, é importante que o limiar estabeleça um equilíbrio entre custos dos erros e a capacidade operacional da equipe"

Parecer do Docente

Perfeita compreensão da assimetria clínica: contrapôs a gravidade do atraso no socorro (FN) com a restrição real de capacidade operacional da equipe médica frente à fadiga de alarmes desnecessários (FP).

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

Resposta Transcrita do Aluno

> "Os dados de entrada, os erros, as previsões e a calibração do modelo devem ser monitorados, sendo necessário investigar a causa ao detectar anomalias, recalibrando e/ou retreinando o modelo"

Parecer do Docente

Telemetria em quatro dimensões (entradas, erros, previsões e calibração contínua), associada a protocolos de investigação de causas com recalibração ou retreino.

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

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

Resposta Transcrita do Aluno

> "O Word2Vec gera vetores de palavras fixos, ao passo que modelos como BERT produzem representações contextualizadas. A similaridade do cosseno mede proximidade entre vetores, porém, não garante que o documento recuperado seja relevante, correto ou atualizado"

Parecer do Docente

Raciocínio teórico límpido: diferenciou embeddings estáticos de contextuais profundos e explicitou que o cosseno mede proximidade geométrica vetorial, mas é cego à relevância substantiva e à atualização normativa.

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

Resposta Transcrita do Aluno

> "Apenas retornar um documento não significa que ele seja útil, a avaliação deve usar consultas com documentos relevantes previamente identificados, além de métricas como MRR e nDCG"

Parecer do Docente

Pontuação máxima. Refutou a ilusão de utilidade pelo simples retorno forçado e citou a metodologia com base anotada de relevância avaliada por métricas canônicas de ranking (MRR e nDCG).

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

Resposta Transcrita do Aluno

> "A busca híbrida combina a compreensão semântica dos vetores com a busca lexical BM25, que identifica termos exatos, além de permitir filtros por metadados. Normas revogadas devem ser identificadas e excluídas das respostas atuais"

Parecer do Docente

Pontuação integral. Especificou a arquitetura híbrida clássica (vetores densos + BM25 léxico para identificadores) aliada a filtros por metadados para exclusão estrita de normas revogadas.

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

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

Resposta Transcrita do Aluno

> "O maior escore não garante que uma resposta esteja correta, o sistema deve responder somente quando o primeiro resultado superar o limiar mínimo T1 e sua diferença para o segundo superar T2, caso contrário, deve se abster"

Parecer do Docente

Formulação canônica irrepreensível: estabeleceu o limiar absoluto de confiança mínima ( ) e a margem de separação competitiva ( ) para desarmar ambiguidades, abstendo-se fora dessas condições.

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

Resposta Transcrita do Aluno

> "É imprescindível verificar a autoridade, a validade e a data dos documentos, se houver conflito não resolvido, o sistema deve se abster e encaminhar a consulta para revisão humana, a política deve ser testada com consultas respondíveis e não respondíveis, avaliando erros e abstenções"

Parecer do Docente

Governança documental impecável: verificação de autoridade, validade e data, com encaminhamento humano (human-in-the-loop) e conjunto de validação prévia contendo consultas respondíveis e propositalmente sem resposta.

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

Resposta Transcrita do Aluno

> "Devem ser monitoradas a taxa de abstenção, os escores de recuperação e os conflitos entre documentos, um aumento anormal nas abstenções pode indicar mudanças no perfil das consultas ou no acervo documental, exigindo revisão das fontes e dos limiares"

Parecer do Docente

Visão sistêmica madura: correlacionou a telemetria da taxa de abstenção e de conflitos com desvios na formulação de consultas ou defasagem do acervo documental, orientando a recalibração de limiares.

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