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:** Gustavo Alexandre Silva
- **Matrícula:** 20233004100
- **Turma:** Turma 1
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 21,0 / 25,0** (Aproveitamento: **84,0%**)
- **Legibilidade da Caligrafia:** 91% (Alta Legibilidade)

Desempenho Consolidado por Questão

Questão / Tema	Pontos Obtidos	Valor	Aproveitamento
Q1: Triagem com dados heterogêneos	**4,7**	5,0	94,0%
Q2: Inspeção visual em condições novas	**3,6**	5,0	72,0%
Q3: Alerta que muda decisões	**4,1**	5,0	82,0%
Q4: Busca de conhecimento para atendimento	**4,4**	5,0	88,0%
Q5: Assistente que pode não responder	**4,2**	5,0	84,0%
TOTAL GERAL	**21,0**	**25,0**	**84,0%**

Questão 1 — Triagem com dados heterogêneos (4,7 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "A afirmação da equipe é uma afirmação incorreta a queda contínua da perda de treino indica que o modelo se tornou especialista naqueles dados de treinamento e perdeu a capacidade de generalizar, no entanto a partir da época 8 o modelo começa a sofrer overfitting, ou seja ele começa a errar as predições de novos dados."

Parecer do Docente

Diagnóstico impecável: identificou a dissociação das curvas, apontou a perda de generalização (overfitting) a partir da Época 8 e refutou com solidez a decisão da equipe técnica.

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

Resposta Transcrita do Aluno

> "Dados categóricos devem ser convertidos em numéricos aplicando-se a tecnica de one hot incoding e os dados numéricos devem ser normalizados. O backpropagation serve para ajuste dos pesos no treinamento a cada época, indicando a direção e a necessidade e grau de ajuste"

Parecer do Docente

Excelente domínio: indicou One-Hot Encoding para atributos categóricos, normalização para numéricos e articulou com clareza o papel do backpropagation (cálculo vetorial da direção e magnitude de ajuste dos pesos via gradientes).

1.c Evidências Clínicas de Confiança e Reprodutibilidade (0,8 / 1,0)

Resposta Transcrita do Aluno

> "Antes de avaliar o modelo é necessário extrair as métricas para avaliar precisão, Recall, F1-Score, são métricas que ajudam a verificar a capacidade de previsão do modelo, escolhendo aumentar uma métrica que diminua a consequência caso haja um erro de previsão. Uma medida concreta para garantir reprodutividade é usar earlystop para verificar a validação e parar o treinamento antes de sofrer overfitting"

Parecer do Docente

Correta ênfase no trade-off de métricas clínicas (Recall vs Precisão) e uso de early stopping. Faltou citar a fixação de sementes aleatórias para reprodutibilidade estrita.

Questão 2 — Inspeção visual em condições novas (3,6 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "É necessário uma análise durante o treinamento de treino e validação para não treinar demais ao ponto de não ser generalista. É interessante comparar os dois turnos e verificar o resultado se a rotulação predita está compatível com a real"

Parecer do Docente

Propôs comparar os resultados entre os turnos diurno e noturno. Deixou de discriminar formalmente o deslocamento de covariáveis luminoso (covariate shift) e a inconsistência na anotação dos inspetores humanos.

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

Resposta Transcrita do Aluno

> "A CNN (rede convolucional) é ideal para esse problema porque lida com precisão de imagem, entretanto pelo custo operacional é necessário carregar um modelo pré-treinado e congelar os pesos durante o treinamento, após isso como é apresentado poucos dados para melhorar o modelo é essencial regularizar usando data augmentation, pegando os dados já existentes e rotacionar, aplicar alguns filtros e transformações para aumentar os dados"

Parecer do Docente

Muito boa resposta de engenharia: justificou o uso de CNNs pré-treinadas, congelamento de pesos e uso ostensivo de data augmentation (rotações, filtros e transformações). Pequeno desconto pela ausência do esquema visual.

2.c Cautela Amostral e Validação com Poucos Dados (0,6 / 1,0)

Resposta Transcrita do Aluno

> "É necessário aumentar as amostras para o modelo ter um treinamento melhor e capaz de generalizar mais então coletar mais dados realizar data augmentation contribui para isso. Outro ponto é avaliar as métricas, porque sem uma análise de métricas, pode dar uma falsa sensação de resultado bom."

Parecer do Docente

Alertou sobre a falsa sensação de bom resultado. Faltou detalhar a validação cruzada estratificada, teste cego no turno noturno e cálculo de intervalos de confiança / testes de hipótese estatísticos.

Questão 3 — Alerta que muda decisões (4,1 / 5,0 pts)

3.a Interpretação da Calibração e Crítica ao Limiar 0,50 (1,9 / 2,0)

Resposta Transcrita do Aluno

> "Porque ouve uma falsa sensação de aceitável, 0,5 não quer dizer 50% na verdade é somente 18% do real o que deixa passar ainda muitos erros por estar mal calibrado"

Parecer do Docente

Leitura quantitativa cirúrgica da Figura 2: extraiu o valor de 18% real para o escore 0,50, denunciou a falsa sensação de segurança e diagnosticou o desvio de calibração.

3.b Custos Assimétricos e Capacidade da Equipe Médica (1,4 / 2,0)

Resposta Transcrita do Aluno

> "A matriz de confusão deve ser bem analisada para o contexto específico, então a escolha do ponto deve levar em conta os efeitos colaterais de um erro. Dessa forma, deve priorizar a analise de uma métrica que reduza o dano."

Parecer do Docente

Compreendeu a minimização do dano em custos assimétricos. Faltou explicitar a diminuição do limiar para priorizar Recall e ponderar a restrição de capacidade operacional da equipe de saúde.

3.c Telemetria Contínua e Deriva de Conceito (0,8 / 1,0)

Resposta Transcrita do Aluno

> "É necessário um monitoramento das previsões, qual os resultados, monitorar as entradas e caso disparado uma anomalia revisar imediatamente o modelo de treinamento"

Parecer do Docente

Monitoramento de previsões e entradas em tempo real, com gatilho de revisão em anomalias. Faltou citar o plano de contingência e fallback assistencial.

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

4.a Representação Semântica e Limitações da Similaridade de Cosseno (1,9 / 2,0)

Resposta Transcrita do Aluno

> "no uso de Transformers/BERT é levado em conta contextos já WORD2VEC gera vetores fixos. O cosseno calcula a similaridade semântica com base na direção do vetor mas não garante a relevância real de documentos"

Parecer do Docente

Excelente contraste entre modelos contextuais (BERT) e estáticos (Word2Vec) e diagnóstico preciso da limitação do cosseno (mede apenas direção vetorial, sem garantir relevância factual/legal).

4.b Avaliação Rigorosa de Recuperação (Information Retrieval) (1,3 / 1,5)

Resposta Transcrita do Aluno

> "Porque você não tem uma garantia apenas pela analise visual é necessário validações como 'ground truth' para analisar as métricas de relevância dos documentos com base nos top k itens"

Parecer do Docente

Muito bom! Exigiu a existência de uma base rotulada ('ground truth') e métricas de relevância sobre os Top-K itens recuperados. Faltou apenas nomear as métricas padrão como NDCG ou MRR.

4.c Busca Híbrida e Governança Normativa (1,2 / 1,5)

Resposta Transcrita do Aluno

> "Porque a busca híbrida analisa tanto buscas semânticas como também busca léxicas o que faz com que haja um desempenho melhor que busca vetoriais puras não cobrem."

Parecer do Docente

Identificou com clareza o ganho de cobertura unindo recuperação léxica e semântica. Faltou explicitar os filtros por metadados de vigência para governança de normas revogadas.

Questão 5 — Assistente que pode não responder (4,2 / 5,0 pts)

5.a Contestação do Maior Escore e Desenho da Abstenção (1,6 / 2,0)

Resposta Transcrita do Aluno

> "O maior escore não vai garantir uma resposta correta, é necessário um limiar mínimo para que o agente não alucine e cuspa uma resposta incorreta mas porque teve que dar uma resposta por não ter um limiar mínimo"

Parecer do Docente

Refutou com firmeza o maior escore e associou a ausência de limiar mínimo à indução forçada de alucinações. Faltou apenas formalizar a margem de separação contra ambiguidade ( ).

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

Resposta Transcrita do Aluno

> "Deve ser validado, tanto com metricas, na qual especialistas vão dizer quais documentos são relevantes e atribuir metricas a partir do resultado gerado, quanto os documentos devem carregar metadados de origem. Deve ser feito consultas: queris respondíveis e não respondíveis, no qual quando não respondível, o score abaixo do limiar dizer que não tem informações sobre isso e não trazer respostas alienadas"

Parecer do Docente

Excelente resposta! Previu avaliação com especialistas e ground truth, metadados de origem e desenho de testes com consultas propositalmente respondíveis e não respondíveis, validando o mecanismo de abstenção.

5.c Telemetria da Taxa de Abstenção em Produção (0,6 / 1,0)

Resposta Transcrita do Aluno

> "Deve ser monitorado através dos metadados que carregam apresentando uma serie de informações do documento retornado"

Parecer do Docente

Citou acompanhamento dos metadados dos documentos. Faltou acompanhar a métrica operacional central: a taxa de abstenção e seu significado diante da desatualização da base normativa.
