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:** Matheus Henrique Almeida Silva
- **Matrícula:** 20233004039
- **Turma:** Turma 2
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 23,9 / 25,0** (Aproveitamento: **95,6%**)
- **Legibilidade da Caligrafia:** 94% (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	**4,5**	5,0	90,0%
Q3: Alerta que muda decisões	**4,6**	5,0	92,0%
Q4: Busca de conhecimento para atendimento	**4,8**	5,0	96,0%
Q5: Assistente que pode não responder	**5,0**	5,0	100,0%
TOTAL GERAL	**23,9**	**25,0**	**95,6%**

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

> "Na época 8 em diante há overfitting, isso indica uma baixa generalização, ou seja, é necessário usar o early stopping como forma de teste antes de realmente implantar o modelo."

Parecer do Docente

Nota integral. Identificou a inflexão na Época 8 caracterizando sobreajuste, destacou a baixa capacidade de generalização e prescreveu early stopping com testes rigorosos antes de implantar.

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

Resposta Transcrita do Aluno

> "Os dados devem ser normalizados e codificados (one-hot encoding). A retropropagação calcula os gradientes do erro em relação aos pesos, permitindo ajustá-los pra minimizar perda."

Parecer do Docente

Nota integral. Resposta técnica impecável: normalização numérica, One-Hot Encoding para variáveis categóricas e cálculo de gradientes da função de perda para ajuste ótimo dos pesos.

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

Resposta Transcrita do Aluno

> "Separar dados de treino, validação e teste, avaliar recall, precisão e sensibilidade clínica, pra reproduzir experimentos, fixar a semente aleatória e registrar os parâmetros do treinamento."

Parecer do Docente

Nota integral. Partição tripla de dados, métricas clínicas adequadas (recall, precisão, sensibilidade) e garantia de reprodutibilidade por controle de sementes e parâmetros.

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

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

Resposta Transcrita do Aluno

> "A rotina de investigação seria uma análise das imagens e os resultados do modelo, comparando os dados do turno diurno e noturno. Overfitting: uma diferença significativa entre resultado de treinamento e validação indicaria possível sobreajuste. Covariate shift: se as imagens noturnas apresentarem características diferentes das imagens de treinamento, pode haver mudança na distribuição dos dados. Falha de rotulação: pegaria amostras mal classificadas e pediria uma nova avaliação humana pra verificar se existem rótulos errados ou inconsistentes."

Parecer do Docente

Delineamento experimental primoroso: comparou os turnos diurno e noturno e discriminou com precisão cirúrgica o sobreajuste, o covariate shift na distribuição visual e o ruído de anotação por reavaliação humana.

2.b Arquitetura Convolucional e Transferência de Aprendizado (1,5 / 2,0)

Resposta Transcrita do Aluno

> "Utilizaria uma CNN com transferência de aprendizado aproveitando um modelo pré-treinado com imagens gerais."

Parecer do Docente

Identificou o uso apropriado de CNN e a transferência de aprendizado reaproveitando modelos pré-treinados. Faltou apenas detalhar o congelamento das camadas base e o uso de data augmentation com variações luminosas.

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

Resposta Transcrita do Aluno

> "Separaria um conjunto de teste independente (representando o turno noturno) com rótulos revisados por especialistas, avaliaria métricas como precisão, recall, F1-score e matriz de confusão considerando intervalos de confiança."

Parecer do Docente

Excelente rigor metodológico: base de teste independente para o turno noturno auditada por especialistas, métricas completas e uso de intervalos de confiança estatísticos.

Questão 3 — Alerta que muda decisões (4,6 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "A premissa da direção para fixar o limiar em 0,50 é falha porque previsão não é risco real de 50%, na verdade é 18%, o modelo está mal calibrado."

Parecer do Docente

Nota integral. Extraiu com exatidão quantitativa a probabilidade real associada ao limiar de 0,50 (apenas 18%) a partir da Figura 2, caracterizando a superconfiança do modelo.

3.b Custos Assimétricos e Capacidade da Equipe Médica (1,8 / 2,0)

Resposta Transcrita do Aluno

> "É necessário encontrar um equilíbrio, já que os falsos positivos sobrecarregam e falsos negativos podem colocar pacientes em situação de risco."

Parecer do Docente

Muito boa ponderação entre a sobrecarga operacional decorrente do excesso de alarmes falsos e o risco à vida do paciente em casos de falso negativo.

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

Resposta Transcrita do Aluno

> "Ficar de olho em previsões, erros, calibração e entradas. Dependendo do contexto realizar uma revisão do modelo e até mesmo retreinamento."

Parecer do Docente

Monitoramento adequado de entradas, previsões e calibração com gatilhos de retreinamento. Faltou formalizar um procedimento de contingência/fallback imediato.

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

> "Embeddings estáticos atribuem um valor fixo a cada palavra, enquanto Transformers geram representações contextualizadas. A similaridade do cosseno mede proximidade entre vetores, mas não garante relevância."

Parecer do Docente

Excelente contraste conceitual entre representações estáticas e dinâmicas/contextualizadas em Transformers, apontando a insuficiência do cosseno para garantir relevância normativa.

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

Resposta Transcrita do Aluno

> "Retornar documentos não garante que sejam relevantes. Criaria um conjunto de consultas com documentos relevantes identificados por especialistas e avaliaria com Precision@K, Recall@K e MRR."

Parecer do Docente

Nota integral. Estruturou um benchmark com especialistas (ground-truth) e aplicou as métricas canônicas de avaliação de recuperação de informação: Precision@K, Recall@K e MRR.

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

Resposta Transcrita do Aluno

> "Combinaria busca vetorial com busca lexical BM25, que identifica termos exatos, e filtros de metadados, como data e categoria, normas revogadas seriam marcadas como obsoletas e excluídas por padrão, priorizando documentos atuais e válidos."

Parecer do Docente

Nota integral. Integração canônica da busca híbrida: vetores densos + BM25 para termos exatos associados a filtros de metadados temporais e expurgo automático 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 sistema deve responder somente se o maior escore ultrapassar um limiar mínimo (T1) e a diferença entre os dois maiores escores ultrapassar outro limiar (T2), caso contrário deve se abster e solicitar revisão humana."

Parecer do Docente

Nota integral. Formulação irrepreensível dos critérios canônicos de abstenção: limiar de confiança absoluta ( ) e margem competitiva ( ), acionando human-in-the-loop caso não atendidos.

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

Resposta Transcrita do Aluno

> "Documentos conflitantes devem ser identificados e comparados por confiabilidade, data e autoridade da fonte. Se não for possível resolver o conflito, o sistema deve se abster. A política deve ser validada com casos de teste contendo documentos conflitantes."

Parecer do Docente

Nota integral. Governança por metadados de autoridade e recência temporal, abstenção prudencial e validação empírica prévia com casos de teste com documentos conflitantes.

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

Resposta Transcrita do Aluno

> "Monitoraria a taxa de abstenção, a precisão das respostas e os erros por tipo de consulta. Se as abstenções aumentarem, analisaria exemplos de consultas e documentos pra identificar a causa, ajustando os limiares ou melhorando a recuperação de documentos."

Parecer do Docente

Nota integral. Visão de engenharia completa: monitoramento contínuo da taxa de abstenção e investigação de causa-raiz para recalibração de limiares ou aprimoramento da recuperação.

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