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:** Carlos Matheus de Almeida Eufrazio
- **Matrícula:** 20233007620
- **Turma:** Turma 1
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 16,6 / 25,0** (Aproveitamento: **66,4%**)
- **Legibilidade da Caligrafia:** 87% (Média Legibilidade)

Desempenho Consolidado por Questão

Questão / Tema	Pontos Obtidos	Valor	Aproveitamento
Q1: Triagem com dados heterogêneos	**3,4**	5,0	68,0%
Q2: Inspeção visual em condições novas	**3,4**	5,0	68,0%
Q3: Alerta que muda decisões	**3,2**	5,0	64,0%
Q4: Busca de conhecimento para atendimento	**2,9**	5,0	58,0%
Q5: Assistente que pode não responder	**3,7**	5,0	74,0%
TOTAL GERAL	**16,6**	**25,0**	**66,4%**

Questão 1 — Triagem com dados heterogêneos (3,4 / 5,0 pts)

1.a Diagnóstico da Curva de Aprendizado e Decisão de Produção (1,8 / 2,0)

Resposta Transcrita do Aluno

> "A partir da época 8 a curva de validação começa a subir e a de perda continua descendo, isso significa que a validação chegou em seu ponto mínimo de análise e após isso não consegue avaliar precisamente as épocas, podendo até mesmo ter sido causado pelo overfitting. Sendo assim, é incorreto afirmar que o sistema esteja pronto para implantação, pois a partir da época 8 não é possível validar a perda do treinamento precisamente."

Parecer do Docente

Diagnóstico muito bom. Identificou a divergência das curvas na Época 8, diagnosticou o sobreajuste (overfitting) e refutou a implantação em produção.

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

Resposta Transcrita do Aluno

> "Elas devem sofrer uma passagem na camada de entrada, passando pela camada oculta e saindo na camada de saída, tudo isso com a representação vetorial. O backpropagation permite os ajustes dos valores de acordo com os resultados obtidos nas camadas, melhorando o sistema com o passar do tempo"

Parecer do Docente

Mencionou representação vetorial e camadas densas, além de associar a retropropagação ao ajuste iterativo dos parâmetros. Contudo, não especificou o pré-processamento obrigatório para entradas heterogêneas (escalonamento numérico e One-Hot Encoding / embeddings para categóricas) nem o cálculo dos gradientes via regra da cadeia.

1.c Evidências Clínicas de Confiança e Reprodutibilidade (0,6 / 1,0)

Resposta Transcrita do Aluno

> "A análise da precisão e do recall auxiliam no entendimento de como o modelo se comportou no geral. Mas uma medida concreta é a acurácia, que vai, geralmente, identificar o quanto o modelo conseguiu reproduzir os dados corretamente ou, com a ajuda das métricas anteriores, como ele apenas decorou as informações ao invés de aprender."

Parecer do Docente

Citou métricas essenciais (precisão, recall e acurácia). Todavia, no contexto de triagem médica desbalanceada, a acurácia é uma métrica ilusória, sendo o recall/curva PR a evidência primária. Omitiu também requisitos de reprodutibilidade experimental (sementes aleatórias e versionamento).

Questão 2 — Inspeção visual em condições novas (3,4 / 5,0 pts)

2.a Diagnóstico Experimental do Aumento de Alertas (1,8 / 2,0)

Resposta Transcrita do Aluno

> "É necessário fazer uma investigação nos dados entregues ao modelo para realizar o pré-treinamento, eles podem apresentar algum padrão que o sistema acabou decorando e causou overfitting. Também precisamos verificar se não houve nenhuma alteração nas formas de analisar cada item ou se alguma informação foi invertida com outra o que poderia causar o covariate shift. Por fim, analisar se tudo foi corretamente classificado antes do treinamento, para termos certeza se não há falhas no processo de rotulação"

Parecer do Docente

Excelente articulação! Abordou com discernimento as três hipóteses fundamentais discriminadas no gabarito: sobreajuste (memorização de padrões espúrios), mudança de distribuição (covariate shift) e inconsistência/falha no processo de rotulação humana.

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

Resposta Transcrita do Aluno

> "As redes convolucionais são melhores quando se trata de dados e aprendizado com base em imagens, porém é necessário técnicas como o Dropout para evitar que o sistema apenas decore os modelos, mesmo com escassez de dados é possível realizar um treino adequado. Dropout: Elimina randomicamente alguns dados."

Parecer do Docente

Correto ao prescrever regularização com Dropout para mitigar o sobreajuste em cenários de poucos dados. Faltou fundamentar a arquitetura convolucional pela invariância por translação e detalhar a estratégia prática de transfer learning (congelamento de camadas extratoras e data augmentation coerente com ruído/iluminação noturna).

2.c Cautela Amostral e Validação com Poucos Dados (0,4 / 1,0)

Resposta Transcrita do Aluno

> "Quando se faz um comparativo entre diferentes épocas e decide um valor mínimo de aceitação para determinar a possível melhoria do modelo, são maneiras seguras de validá-lo."

Parecer do Docente

Propôs critérios de parada/comparação entre épocas, mas não abordou a cautela amostral estatística essencial diante do volume escasso de imagens noturnas (intervalos de confiança ou bootstrap para evitar declarar ganhos estatisticamente não significativos).

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

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

Resposta Transcrita do Aluno

> "Porque a calibração de 0.50 não é 'forte' suficiente para validar corretamente a operação, ela não vai ter uma distinção suficiente para diferenciar os pacientes padrões com os que tiveram pioras clínicas. Neste contexto, um classificador mal calibrado é o que não consegue separar os pacientes padrões dos que tiveram pioras clínicas corretamente."

Parecer do Docente

Confundiu calibração com discriminação/separabilidade. Não interpretou o diagrama da Figura 2, onde para p_hat = 0,50 a taxa real observada de urgência é de apenas 18% (sobreconfiança acentuada gerando disparos falsos maciços).

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

Resposta Transcrita do Aluno

> "No contexto hospitalar, os falsos positivos (FP), apesar de serem assustadores, geralmente não causam tantos problemas aos pacientes quanto os falsos negativos, que podem levar a morte. É necessário ponderar sobre o limiar para decidir um valor que tenha o menor número possível de FP e FN, tendo como analisadores da situação a equipe médica, que possuem o conhecimento necessário da área para auxiliar os desenvolvedores na escolha."

Parecer do Docente

Excelente resposta. Demonstrou plena consciência clínica da severidade assimétrica (o falso negativo pode resultar em óbito) e envolveu diretamente a equipe médica para calibrar o ponto de corte operacional e respeitar o limite operacional do pronto atendimento.

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

Resposta Transcrita do Aluno

> "O plano é sempre avaliar o andamento da estrutura dos gráficos, se houver uma alteração não prevista o correto é parar o processamento e alterar o valor do limiar até a identificação do problema."

Parecer do Docente

Plano muito prudente de contingência operacional: monitorar o comportamento contínuo dos gráficos, pausar o processamento automatizado diante de quebras não previstas e reajustar o limiar preventivamente.

Questão 4 — Busca de conhecimento para atendimento (2,9 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "A similaridade de cosseno apesar de conseguir realizar uma ordenação agradável, ela só consegue trabalhar em contextos de embeddings estáticos, tendo dificuldade quando os vetores necessitam de um certo contexto para ser feita a correta classificação."

Parecer do Docente

Registrou corretamente a fórmula matemática do cosseno nas notas marginais e apontou a carência de contexto nos embeddings estáticos. Equivocou-se, contudo, ao afirmar que o cosseno só funciona com vetores estáticos (é a métrica padrão também para vetores densos contextualizados), omitindo que o cosseno mede proximidade angular mas não garante vigência ou precisão factual.

4.b Avaliação Rigorosa de Recuperação (Information Retrieval) (0,7 / 1,5)

Resposta Transcrita do Aluno

> "Os itens terem retornado corretamente não quer dizer necessariamente que eram os documentos exigidos e que o desenvolvedor procurava. O protocolo de avaliação necessita levar em consideração a busca semântica da consulta, pois é preciso ter cuidado com todas as similaridades."

Parecer do Docente

Correto ao rejeitar a métrica rasa de volume retornado. Não especificou, contudo, o desenho formal de avaliação de Information Retrieval (anotação de ground truth de relevância e cálculo de métricas consagradas como Recall@K, Precision@K, MRR ou NDCG).

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

Resposta Transcrita do Aluno

> "A busca híbrida é a ideal, enquanto a léxica vai buscar as palavras parecidas entre os documentos e encontrar a central, a busca semântica vai analisar a central e procurar as semelhantes, permitindo com essa combinação encontrar qualquer tipo de documento, o mais ideal possível de preferência, sem excluir as palavras-chaves ou sinônimos."

Parecer do Docente

Boa intuição sobre a complementaridade da busca léxica (palavras-chave exatas) com a busca semântica (sinônimos e conceitos aproximados). Faltou detalhar a camada de governança por filtros de metadados rígidos (data/status de revogação) para impedir que normas revogadas superem as vigentes.

Questão 5 — Assistente que pode não responder (3,7 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "O documento de maior escore não significa automaticamente o melhor, pois apesar de já levar em consideração coisas como as consultas ambíguas e os documentos conflitantes, ela não analisa os sinônimos. O primeiro limiar mínimo deveria ser algo relacionando a encontrar os documentos que tratam a respeito do mesmo assunto (para que não haja possível fuga do tema) e o segundo limiar busca o que apresentar o maior escore."

Parecer do Docente

Contestou acertadamente a premissa simplista do gestor e propôs uma política com dois limiares para assegurar pertinência e evitar alucinação em perguntas ambíguas. Faltou formular formalmente o critério da margem de separação entre o primeiro e o segundo resultado.

5.b Conflito de Fontes e Validação Prévia da Abstenção (1,4 / 2,0)

Resposta Transcrita do Aluno

> "Eu zero ambos e passo-os para o processo de abstenção. Na abstenção eles são avaliados individualmente e o que fizer maior escore de acordo com as consultas respondíveis e não respondíveis pode retornar ao escopo, enquanto o outro é descartado da busca atual."

Parecer do Docente

Proposta muito interessante de aplicar abstenção preventiva diante do conflito entre fontes e validar o comportamento do sistema com consultas respondíveis e intencionalmente não respondíveis.

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

Resposta Transcrita do Aluno

> "O monitoramento pode ser feito através dos gráficos de abstenção e o sinal que eles dão aos engenheiros de modificação são as quedas bruscas ou aumentos bruscos irregulares que não condizem com o processamento contínuo do gráfico."

Parecer do Docente

Identificou com acerto o papel da telemetria contínua: anomalias e aumentos bruscos na taxa de abstenção sinalizam desvios de regime (query drift ou lacunas documentais criadas por atualizações do acervo).
