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:** Emanuel Carlos de Oliveira Dias
- **Matrícula:** 20233004020
- **Turma:** Turma 1
- **Data da Avaliação:** 19/10/2026
- **Pontuação Final: 24,4 / 25,0** (Aproveitamento: **97,6%**)
- **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	**4,8**	5,0	96,0%
Q3: Alerta que muda decisões	**4,8**	5,0	96,0%
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	**24,4**	**25,0**	**97,6%**

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

> "Revela que o modelo está em overfitting. O modelo está memorizando padrões do treinamento e perdendo capacidade de generalização. Portanto, a redução da perda de treinamento não garante bom desempenho em dados novos. Devemos considerar early stopping na época 8 e avaliar o modelo em um conjunto de testes independentes."

Parecer do Docente

Diagnóstico impecável: identificou a dissociação das curvas, explicou a perda de capacidade de generalização (overfitting), vetou a entrada em produção baseada apenas no erro de treino e propôs early stopping na Época 8 com validação em teste independente.

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

Resposta Transcrita do Aluno

> "As variáveis numéricas, como idade e sinais vitais, devem ser normalizados e/ou padronizados. As variáveis categóricas podem ser representados por one-hot encoding ou embeddings. Essas entradas são concatenadas e fornecidos à MLP. Durante o treinamento, a retropropagação calcula os gradientes de perda pela regra da cadeia, permitindo atualizar seus pesos e vieses, minimizando erros."

Parecer do Docente

Excepcional clareza e completude: cobriu normalização/padronização de numéricas, codificação por One-Hot/Embeddings para categóricas, concatenação no vetor de entrada e o cálculo formal de gradientes via regra da cadeia no backpropagation.

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

Resposta Transcrita do Aluno

> "São necessários dados representativos de diferentes perfis de pacientes, incluindo urgentes. Devem ser avaliados métricas sensibilidade, especificidade, precisão, etc. Para garantir reprodutibilidade, devem ser registrados sementes aleatórias, versões dos dados, hiperparâmetros, etc."

Parecer do Docente

Pontuação integral. Articulou a necessidade de métricas clínicas adequadas para classes raras/urgentes (sensibilidade/recall e especificidade) e listou os três pilares de reprodutibilidade experimental: fixação de sementes, versionamento de dados e registro de hiperparâmetros.

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

> "Comparar imagens diurnas e noturnas quanto à iluminação e desempenho. Analisar perdas de treino e validação para detectar overfitting. Diferenças nas distribuições para identificar covariate shift. Revisar rótulos para encontrar erros de anotação."

Parecer do Docente

Perfeito! Estabeleceu o protocolo de contraste empírico entre os períodos e delineou a forma de investigar cada uma das três causas: curvas para sobreajuste, deslocamento de distribuição para covariate shift e auditoria de inspeção humana para erros de rotulação.

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

Resposta Transcrita do Aluno

> "CNNs identificam padrões espaciais, como bordas e texturas. Utilizar-se um modelo pré-treinado, congelando inicialmente suas camadas e depois realizando fine-tuning. Aplicar data augmentation com variações de brilho e contraste."

Parecer do Docente

Fundamentação conceitual correta (extração de padrões espaciais invariantes) com pipeline completo de transferência de aprendizado (congelamento + ajuste fino) e perturbações fotométricas de brilho/contraste. Pequeno desconto apenas por não ter desenhado o diagrama esquemático.

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

Resposta Transcrita do Aluno

> "Ampliar a amostra, utilizar validação estratificada e intervalos de confiança. Manter revisão humana até obter resultados confiáveis."

Parecer do Docente

Excelente! Exigiu validação estratificada e intervalos de confiança estatísticos para evitar aprovações espúrias, aliando salvaguarda operacional humana (human-in-the-loop).

Questão 3 — Alerta que muda decisões (4,8 / 5,0 pts)

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

Resposta Transcrita do Aluno

> "O modelo está mal calibrado, pois prevê probabilidades maiores que as frequências reais. Uma previsão de 50% corresponde a 18% de ocorrência. Portanto, o limiar 0,50 não é adequado."

Parecer do Docente

Leitura quantitativa exata da curva de calibração na Figura 2 (18% de piora real para escore nominal de 50%), diagnosticando a superestimação de probabilidades e a inadequação do limiar padrão.

3.b Custos Assimétricos e Capacidade da Equipe Médica (1,9 / 2,0)

Resposta Transcrita do Aluno

> "O limiar deve equilibrar falsos positivos, que sobrecarregam a equipe, e falsos negativos, que podem comprometer os pacientes. Devem ser considerados os custos dos erros e a capacidade operacional."

Parecer do Docente

Perfeita compreensão da dinâmica de custos assimétricos: contrapôs o risco vital do falso negativo com a restrição real de capacidade operacional médica provocada pelo excesso de alarmes falsos.

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

Resposta Transcrita do Aluno

> "Monitorar distribuição dos dados, probabilidades, taxas de alerta e desempenho. Ao detectar deriva, investigar, recalibrar e/ou treinar o modelo com dados recentes."

Parecer do Docente

Plano de telemetria completo: monitoramento de entradas, escores e alarmes em produção, com protocolo de contingência e recalibração/retreino periódico.

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

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

Resposta Transcrita do Aluno

> "Word2Vec utiliza embeddings estáticos, enquanto BERT considera o contexto das palavras. A similaridade do cosseno mede proximidade semântica, mas não garante relevância, validade ou atualização dos documentos."

Parecer do Docente

Excelente discernimento técnico ao contrastar modelos estáticos com representações contextuais profundas e evidenciar que a métrica de cosseno é cega à vigência, relevância jurídica e temporalidade das normas.

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

Resposta Transcrita do Aluno

> "A produção de consultas com retorno não garante qualidade. Deve-se avaliar documentos relevantes usando métricas como Precision@K e Recall@K"

Parecer do Docente

Refutou com firmeza o argumento ingênuo de que ter retorno forçado é sinônimo de sucesso e indicou as métricas clássicas de Information Retrieval Precision@K e Recall@K.

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

Resposta Transcrita do Aluno

> "Combinar busca vetorial com BM25 melhora a recuperação semântica e lexical. Filtros de metadados devem priorizar documentos vigentes, excluindo normas revogadas das respostas atuais."

Parecer do Docente

Pontuação máxima. Apresentou a arquitetura híbrida clássica (vetores densos + BM25 lexical) com mecanismo imperativo de filtragem por metadados para expurgar 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 uma resposta correta. O sistema deve responder apenas quando o escore máximo superar um limiar mínimo (τ_1) e a diferença para o segundo resultado superar a margem (τ_2). Senão, deve se abster."

Parecer do Docente

Formulação matemática impecável. Enunciou a política de abstenção padrão da engenharia de RAG, combinando o limiar absoluto de confiança mínima ( ) com a margem mínima de separação competitiva ( ).

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

Resposta Transcrita do Aluno

> "Documentos conflitantes devem ser analisados conforme à vigência e autoridade. Persistindo a dúvida, o sistema deve se abster. A validação deve incluir consultas respondíveis e não respondíveis, avaliando precisão, cobertura e taxa de abstenção."

Parecer do Docente

Raciocínio maduro de governança documental (critérios de vigência e autoridade), postura de abstenção prudencial e metodologia de validação contendo casos respondíveis e propositalmente sem resposta no acervo.

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

Resposta Transcrita do Aluno

> "Monitorar taxas de abstenção, escores e margens entre resultados. Aumentos inesperados podem indicar mudanças nas consultas e documentos, exigindo revisão do sistema."

Parecer do Docente

Mapeou os indicadores-chave de desempenho em produção (taxa de abstenção e margem inter-resultados) e correlacionou oscilações anômalas com drifts de comportamento dos usuários ou modificações na base regulatória.

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