



Prova Escrita 1 — 2026-2 — 19/10/2026

Disciplina: Inteligência Artificial II

Professor: Prof. Aléssio Miranda Júnior

Nome: _____

Curso: Engenharia de Computação

Duração: 100 minutos

Valor: 25,0 pontos

Matrícula: _____

Modelo A

Instruções. Responda de forma justificada e técnica, explicitando suposições e identificando claramente cada subitem (ex.: [a], [b], [c]). Não há solicitação de código ou chamadas de API.

Questão 1 — Triagem com dados heterogêneos**(5,0 pontos)**

Uma rede de clínicas quer priorizar revisão humana de pedidos de exame. Para cada pedido, há idade, sinais vitais e respostas categóricas de um questionário; o desfecho histórico indica se houve necessidade de atendimento urgente. A equipe propõe uma MLP e apresenta as curvas de aprendizado da Figura 1, afirmando que, como a perda de treinamento diminui a cada época, o sistema está pronto para implantação em produção.

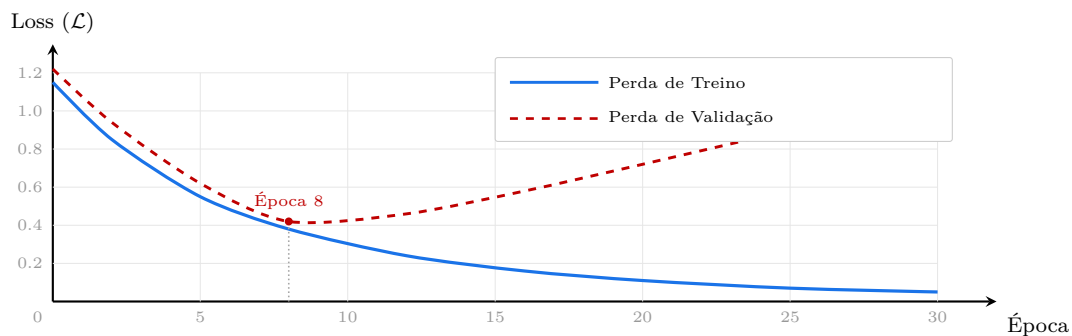


Figura 1: Curvas de perda de treino e validação observadas nas primeiras 30 épocas.

- a) **Figura 1 e Decisão de Produção:** Avalie criticamente a afirmação da equipe. O que o comportamento das curvas de perda a partir da Época 8 revela sobre o modelo, e por que a queda contínua da perda de treino não garante que o sistema esteja pronto para produção? **(2,0 pontos)**
- b) **Representação e Retropropagação:** Como as entradas heterogêneas (numéricas e categóricas) devem ser representadas para alimentar a MLP? Explique também o papel real da retropropagação (*backpropagation*) no ajuste dos parâmetros durante o treinamento. **(2,0 pontos)**
- c) **Evidências de Confiança e Reprodutibilidade:** Quais evidências adicionais (particionamento de dados e métricas clínicas) são indispensáveis antes de confiar no modelo? Cite uma medida concreta para garantir que os experimentos sejam reprodutíveis. **(1,0 pontos)**

Resposta do item (a):

Questão 1 — Continuação dos Campos de Resposta

Página 2 / Modelo A

Resposta do item (b):

Resposta do item (c):

Questão 2 — Inspeção visual em condições novas

(5,0 pontos)

Uma fábrica treinou uma CNN para identificar defeitos em imagens de peças produzidas durante o dia. Na primeira semana noturna, a taxa de alertas cresce muito, embora as peças aprovadas pela inspeção humana pareçam semelhantes. Há poucas imagens rotuladas do turno noturno e um modelo pré-treinado em imagens gerais está disponível.

- a) **Diagnóstico Experimental:** Proponha uma rotina de investigação para o aumento repentino de alertas. Quais resultados práticos distinguiriam entre as hipóteses de sobreajuste (*overfitting*), mudança de distribuição (*covariate shift*) e falha no processo de rotulação? (2,0 pontos)
- b) **Arquitetura e Transferência de Aprendizado:** Justifique por que uma CNN é apropriada para esse problema e desenhe uma estratégia com o modelo pré-treinado (congelamento e ajuste fino) e regularização/*data augmentation* para mitigar a escassez de dados. (2,0 pontos)
- c) **Cautela Amostral:** Diante do baixo volume de imagens rotuladas do turno noturno, como evitar conclusões precipitadas ao validar melhorias no modelo? (1,0 pontos)

Resposta do item (a):

Questão 2 — Continuação dos Campos de Resposta

Página 3 / Modelo A

Resposta do item (b):

Resposta do item (c):

Questão 3 — Alerta que muda decisões

(5,0 pontos)

Um hospital pretende usar a saída probabilística de um classificador para acionar uma equipe de resposta rápida diante de piora clínica de pacientes. O diagrama de calibração levantado na base de validação é apresentado na Figura 2. A direção quer fixar o limiar operacional em 0,50 porque “acima disso o modelo está mais certo que errado”. Após a implantação, os profissionais relatam muitos alarmes pouco úteis e, dois meses depois, o perfil dos pacientes muda.

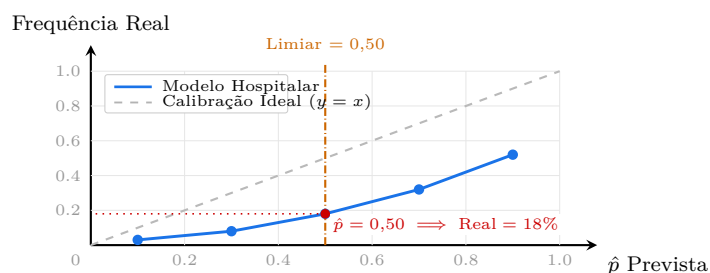


Figura 2: Diagrama de calibração do classificador no conjunto de validação.

- a) **Interpretação da Calibração e Crítica ao Limiar:** Com base na Figura 2, por que a premissa da direção para fixar o limiar em 0,50 é falha? O que significa um classificador estar mal calibrado nesse contexto clínico? (2,0 pontos)

Resposta do item (a):

Questão 3 — Continuação dos Subitens e Respostas

Página 4 / Modelo A

- b) **Custos Assimétricos e Capacidade Operacional:** Como os custos de Falsos Positivos e Falsos Negativos, combinados com o teto de capacidade da equipe médica, devem ser ponderados na escolha do ponto operacional (limiar)? (2,0 pontos)

Resposta do item (b):

- c) **Telemetria e Deriva de Conceito:** Descreva um plano de monitoramento em produção capaz de detectar deriva nos dados (*data drift*) ou nas decisões, e qual ação técnica imediata deve ser tomada ao detectar anomalias. (1,0 pontos)

Resposta do item (c):

Questão 4 — Busca de conhecimento para atendimento

(5,0 pontos)

Um portal interno deve recuperar documentos para perguntas como “posso alterar a senha fora da rede?”. A base contém manuais antigos, normas recentes e textos curtos de chamados. A equipe considera representar cada pergunta e documento por vetores, ordenar por similaridade do cosseno e medir sucesso apenas pela proporção de consultas que retorna algum resultado.

- a) **Representação Semântica e Limitações do Cosseno:** Compare o uso de *embeddings* estáticos (como Word2Vec) versus contextualizados (como Transformers/BERT) para esse cenário. Quais são as limitações da similaridade de cosseno em assegurar a relevância real de um documento recuperado? (2,0 pontos)
- b) **Avaliação Rigorosa de Recuperação:** Por que medir o sucesso pela mera proporção de consultas com retorno é inadequado? Proponha um protocolo formal de avaliação com métricas consagradas de recuperação da informação (*Information Retrieval*). (1,5 pontos)
- c) **Busca Híbrida e Governança:** Como uma abordagem híbrida (combinando vetores densos com busca lexical/BM25 e filtros de metadados) superaria as limitações da busca vetorial pura? Inclua o tratamento de normas antigas revogadas. (1,5 pontos)

Resposta do item (a):

Questão 4 — Continuação dos Campos de Resposta

Página 5 / Modelo A

Resposta do item (b):

Resposta do item (c):

Questão 5 — Assistente que pode não responder

(5,0 pontos)

A mesma organização quer que o portal dê uma resposta automática somente quando houver evidência suficiente. Consultas ambíguas, documentos conflitantes e mudanças frequentes nas normas são comuns. Um gestor pede que o sistema sempre mostre o documento de maior escore, pois isso elimina a necessidade de regras adicionais.

- a) **Contestação do Maior Escore e Desenho da Abstenção:** Defenda ou conteste a proposta do gestor. Desenhe conceitualmente uma política de recuperação e abstenção baseada em limiar mínimo de escore (τ_1) e margem para o segundo resultado (τ_2). (2,0 pontos)
- b) **Conflito de Fontes e Validação Prévia:** Como tratar documentos com normas conflitantes antes de emitir resposta? Como validar essa política de abstenção em ambiente experimental usando consultas respondíveis e não respondíveis? (2,0 pontos)
- c) **Telemetria e Impacto Operacional:** Como monitorar a taxa de abstenção em produção e qual sinal ela fornece aos engenheiros de IA caso o perfil de consultas ou o acervo documental se modifique? (1,0 pontos)

Resposta do item (a):

Questão 5 — Continuação dos Campos de Resposta

Página 6 / Modelo A

Resposta do item (b):

Resposta do item (c):

Espaço para Rascunho ou Cálculos Adicionais (Opcional)*Destinado a rascunhos, esquemas e deduções. O conteúdo deste espaço não substitui os campos oficiais de resposta.*

Fim da prova — Revise suas respostas antes da entrega!