

O desafio de vocês, em uma frase

Hoje cada grupo sai com seu challenge statement v1 — testado contra duas perguntas: isso é um problema de ML? Isso é produção?

Onde estamos

Essential Question

Em um mundo onde qualquer um publica qualquer coisa, como decidimos em que confiar?

Challenge

Construam um sistema de ML que ajude uma comunidade real a identificar fake news.

Hoje

Transformar o challenge em UM statement por grupo: quem vamos ajudar, a fazer o quê — e por quê isso é um problema de ML.

Rodada-relâmpago: as notícias de vocês



2 minutos por grupo

Mostrem a notícia mais interessante que trouxeram — e a mais difícil de classificar



Como checaram?

Que evidência buscaram? Quanto tempo levou? Chegaram a um veredito?



A 'duvidosa' é o ouro

Se nem vocês bateram o martelo, o que um sistema automático deveria responder nesse caso?

Isso é um problema de ML? — 6 perguntas

1 Existe um PADRÃO a ser aprendido?

2 Existem DADOS — ou um jeito viável de obtê-los?

3 O problema é PREDITIVO? (dada uma entrada, prever algo)

4 As previsões se repetem em ESCALA?

5 O padrão MUDA com o tempo? (regras fixas envelhecem)

6 O CUSTO DO ERRO é tolerável ou gerenciável?

Aplicuem as 6 ao desafio de fake news do SEU grupo. Alguma resposta é 'não'? O que muda no recorte?

Quando ML é a resposta errada



Uma regra simples resolve — comece por ela (é o seu baseline!)



Não há dados, nem jeito realista de consegui-los



Um único erro é inaceitável e não há supervisão humana



É mais barato/robusto resolver sem ML (planilha, if/else, processo)



O uso seria antiético — ainda que tecnicamente possível

Regra de ouro do semestre: baseline simples primeiro. O ML precisa MERECER o lugar dele.

ML em pesquisa × ML em produção

	Pesquisa	Produção
Objetivo	Bater o benchmark (SOTA)	Gerar valor para alguém — e continuar gerando
Dados	Estáticos, limpos, prontos	Bagunçados, mudando o tempo todo, enviesados
Prioridade	Treinar rápido	Responder rápido (latência de inferência)
Pessoas	Pesquisadores	Usuários, mantenedores, stakeholders com objetivos conflitantes
Interpretar & custo	Opcional	Obrigatório: explicar erros, caber no orçamento

O jogo 'Verdade ou Fake' foi pesquisa: dataset fechado, sem usuário. O sistema de vocês é produção.

Oficina: o challenge statement

Ajudar **[QUEM]** a **[FAZER O QUÊ]** em **[ONDE/CANAL]**,
com um sistema que
[O QUE O ML PREVÊ] — e saberemos que funcionou se **[EVIDÊNCIA]**.



Audiência específica

'grupos de família no WhatsApp' >> 'a sociedade'



Predição nomeada

o que entra, o que sai do modelo



Sucesso observável

dá para mostrar evidência na apresentação de 07/10

30 minutos de trabalho em grupo — partam do v0 de vocês.

Statement vago × statement que guia



“Criar uma IA que detecta fake news na internet.”

Quem usa? Que tipo de conteúdo? O que o sistema responde? Como saber se funcionou? — impossível começar.



“Ajudar administradores de grupos de família no WhatsApp a checar correntes de saúde encaminhadas, com um bot que classifica o texto como confiável ou desinformação e aponta os sinais — e saberemos que funcionou se, no teste com 3 grupos reais, ele acertar mais que o baseline de palavras-chave e os admins o usarem antes de repassar.”

Pitch de 1 minuto



1 minuto por grupo

Leiam o statement v1 e digam: qual das 6 perguntas de ML foi a mais difícil de responder?



Feedback dos outros grupos

Duas forças + um risco. Sem 'tá ótimo': apontem o que está vago.



Ajuste final

5 minutos para incorporar o que fizer sentido. O v1 ajustado é a entrega de hoje.

Saídas de hoje



Challenge statement v1

Entregue por escrito ao professor no fim da aula



Checklist de ML respondido

As 6 perguntas aplicadas ao desafio do grupo



Para 19/08

Tragam 3 guiding questions: o que vocês precisam DESCOBRIR para começar? (dados, usuário, requisitos — Cap. 2)

Próxima aula (19/08) · Investigate: guiding questions, requisitos e delimitação — Cap. 2