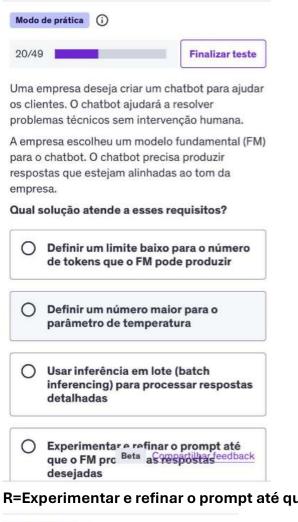


Uma empresa está usando um modelo de linguagem grande pré-treinado (LLM) para construir um chatbot para recomendações de produtos. A empresa precisa que as respostas do LLM sejam curtas e escritas em um idioma específico.

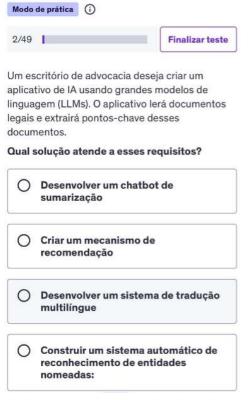
Qual solução alinhará a qualidade das respostas do LLM com as expectativas da empresa?

0	Aumentar a tempe	ratura
0	Escolher um LLM d diferente	e tamanho
0	Ajustar o prompt	
0	Aumentar o valor d	е Тор К
	Beta	Compartilhar feedback

R=AJUSTAR O PROMPT



R=Experimentar e refinar o prompt até que o FM produza as respostas desejadas



R=Desenvolver um chatbot de sumarização



R=AWS CloudTrail



Uma empresa deseja construir uma aplicação de aprendizado de máquina (ML).

Selecione e ordene os passos corretos da lista a seguir para desenvolver uma carga de trabalho de ML bem arquitetada. Cada passo deve ser selecionado apenas uma vez.

Obs: Escreva exatamente conforme a imagem, respeitando Maiúsculas e Minúsculas.

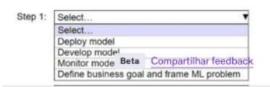
Tradução:

Implantar o modelo (Deploy model)

Desenvolver o modelo (Develop model)

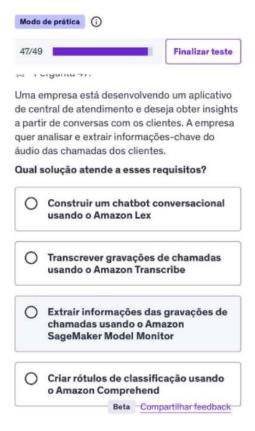
Monitorar o modelo (Monitor model)

Definir metas de negócios e formular o problema de ML (Define business goal and frame ML problem)

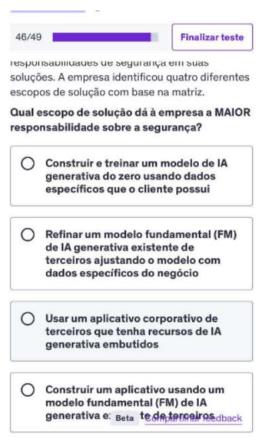


R: Define business goal and frame ML problem,

Develop model, Deploy model, Monitor model



R=Transcrever gravações de chamadas usando o Amazon Transcribe



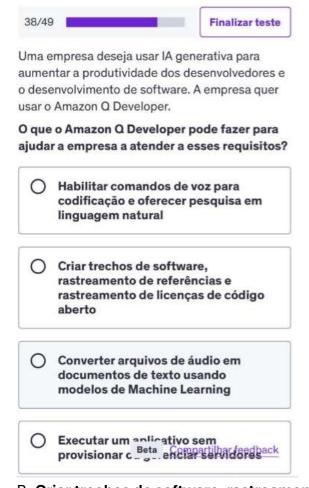
Uma empresa está usando a Matriz de Escopo de Segurança de IA Generativa para avaliar responsabilidades de segurança em suas soluções. A empresa identificou quatro diferentes escopos de solução com base na matriz.

Qual escopo de solução dá à empresa a MAIOR responsabilidade sobre a segurança?

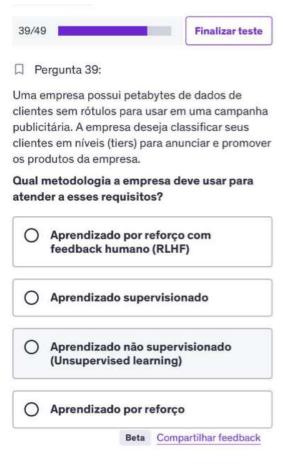
R=Construir e treinar um modelo de IA generativa do zero usando dados específicos que o cliente possui



R=AWS Artifact



R=Criar trechos de software, rastreamento de referências e rastreamento de licenças de código aberto



R=Aprendizado não supervisionado (Unsupervised learning)

	de prática (i)	
35/49		Finalizar teste
□ Pe	ergunta 35:	
Bedro aplica atendo a VPC	ck para desenvolver (tivo está hospedado (er aos padrões de cor	está usando o Amazon um aplicativo de IA. O em uma VPC. Para nformidade regulatória para acessar nenhum
	serviço ou recurso d requisitos?	a AWS atenderá a
0	Internet gateway	
0	Amazon Macie	
0	AWS PrivateLink	
	Amazon CloudFror	



Ativar o registro de invocações do modelo para coletar mensagens.

 Usar Throughput Provisionado para o LLM.

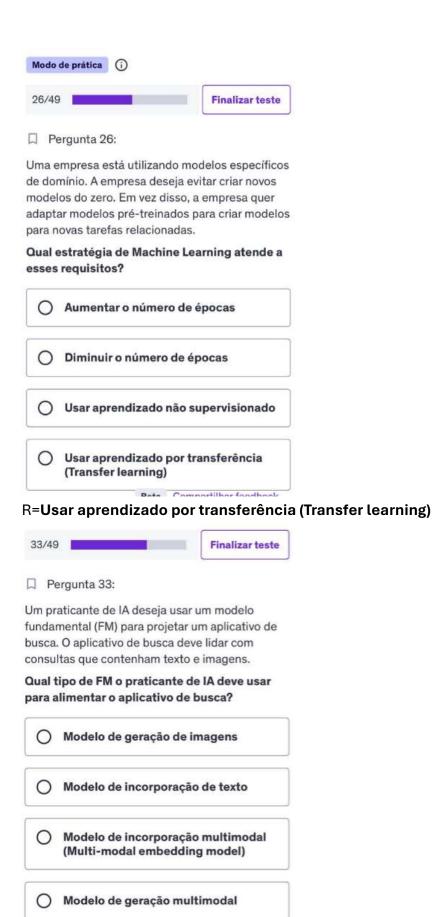
 Usar o Amazon Personalize para salvar o histórico de conversas.

 Adicionar mensagens ao prompt do modelo.

das mensagens anteriores do cliente?

R=Adicionar mensagens ao prompt do modelo.

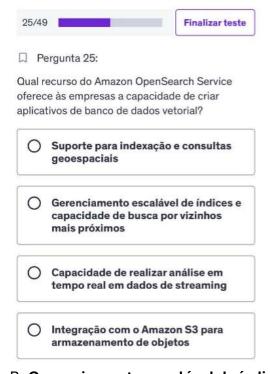
Beta Compartilhar feedback



R=Modelo de incorporação multimodal (Multi-modal embedding model)



R=Amazon SageMaker Feature Store



R=Gerenciamento escalável de índices e capacidade de busca por vizinhos mais próximos

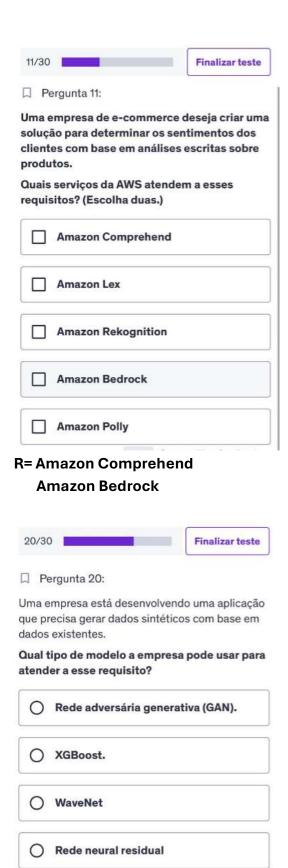
18/49	Finalizar teste	
	rinalizar teste	
□ Pe	ergunta 18:	
plicat	mpresa deseja usar lA para proteger seu ivo contra ameaças. A solução de lA preci ar se um endereço IP é de uma fonte ta.	sa
Qual s	olução atende a esses requisitos?	
0	Criar um sistema de previsão de fraude	
0	Criar um sistema de reconhecimento de entidades nomeadas (NLP)	
0	Desenvolver um sistema de detecção de anomalias	
0	Criar um sistema de reconhecimento de fala	
(=D(esenvolver um sistema de	detecção de anomaι ¬
13/30	Finalizar teste	
□ P	ergunta 13:	
mode mode	empresa de mídia social deseja usar um elo de linguagem grande (LLM) para rração de conteúdo. A empresa quer ir as saídas do LLM quanto a preconceito sível discriminação contra grupos ou	
e pos indiví Qual	duos específicos. fonte de dados a empresa deve usar par	
e pos indiví Qual avalia	5	
e pos indiví Qual avalia	fonte de dados a empresa deve usar par ır as saídas do LLM com o MENOR esforo	
e pos indiví Qual avalia	fonte de dados a empresa deve usar par ur as saídas do LLM com o MENOR esforo nistrativo?	
e pos indiví Qual avalia	fonte de dados a empresa deve usar par ir as saídas do LLM com o MENOR esfor nistrativo? Logs de moderação.	

7/30		Finalizar teste
	//	Timanzar teste
☐ Pe	rgunta 7:	
earnin	mpresa construiu um mode g para detecção de objeto: o em produção.	
	rocesso de IA ocorre qua a uma nova imagem para s?	
0	Treinamento (Training)	
0	Implantação do modelo deployment)	(Model
0	Inferência (Inference)	
0	Correção de viés (Bias co	orrection)
R=Inf	erência (Inference)	
17/30		Finalizar teste
□ Pe	ergunta 17:	
Amazo orçam compo	empresa deseja criar um a on Bedrock. A empresa po nento limitado e prefere fle romisso de longo prazo.	ossui um exibilidade sem
Amazo orçam compo Qual i	on Bedrock. A empresa po nento limitado e prefere fle	ossui um exibilidade sem do Amazon
Amazo orçam compo Qual i	on Bedrock. A empresa po nento limitado e prefere fle romisso de longo prazo. modelo de precificação e	ossui um exibilidade sem do Amazon sitos?
Amazo orçam compo Qual i Bedro	on Bedrock. A empresa po nento limitado e prefere fle romisso de longo prazo. modelo de precificação ock atende a esses requi	exibilidade sem do Amazon sitos? delo (Model
Amazo orçam compi Qual i Bedro	on Bedrock. A empresa po nento limitado e prefere fle romisso de longo prazo. modelo de precificação ock atende a esses requi Personalização de mod customization)	do Amazon sitos? delo (Model

R=Sob demanda (On-Demand)

10/30		Finalizar teste
model diagno transp	empresa médica está persi lo fundamental (FM) para f óstico. A empresa precisa parente e explicável para a sitos regulatórios.	ins de que o modelo seja
0	Gerar métricas simples exemplos usando o Am SageMaker Clarify	
0	Coletar mais dados. Us Rekognition para adicio personalizados aos dad	onar rótulos
0	Configurar a segurança conformidade usando o Inspector	
0	Criptografar e proteger treinamento usando o A	
R=Ge	rar métricas simpl	es, relatórios
1/30		Finalizar teste
interaç	ções passadas e recursos	s online.
	estratégia de aprendiza capacidade de autoaper	
Altern	nativas traduzidas:	
0	Aprendizado por refor recompensas para fee dos clientes	
0	Aprendizado supervis conjunto de dados ma organizado de boas e	nualmente
0	Aprendizado supervis banco de dados de FA continuamente atualiz	a
0	Aprendizado não supe encontrar clusters de similares dos clientes	consultas

R=Aprendizado por reforço com recompensas para feedback positivo dos clientes



R=Rede adversária generativa (GAN).

□ P	ergunta 15:				
Uma	empresa construiu um ch	atbot que pode			
respo	nder a perguntas em ling	uagem natural com	n		
	ens. A empresa deseja ga				
não re	etorne imagens inadequa	das ou indesejadas	5.		
_	II - PERMIA - MANAGEMENT DE LA CAMPANA DE LA				
O	Implementar APIs de r (Implement moderation				
	V		J		
0	Automotizar a integra	aão do foodback			
O	Automatizar a integra do usuário:	,ao do leedback			
	PORTO PERSONAL PORTO POR		J		
_	B	dete	ì		
0	Realizar validação do	nodelo			
			7		
0	Treinar novamente o n				
	conjunto de dados pú	olico geral:			
rever nemó rogra prend lesen lados lados	mpresa de dispositivos dig a demanda do cliente por l ria. A empresa não possui e imação ou conhecimento d dizado de máquina (ML) e p volver um modelo preditivo . A empresa precisa realizar internos e externos.	hardware de experiência em e algoritmos de recisa orientado por r análises em			
Guai s	olução atendera a esses	equisitos:			
0	Importe os dados para o SageMaker Canvas. Con- modelos de ML e previso demanda selecionando d dados do SageMaker Ca	strua ies de os valores nos			
0	Importe os dados para o SageMaker Data Wrangl modelos de ML e previsõ demanda usando uma re Personalize Trending-No	er. Construa ies de eceita Amazon			

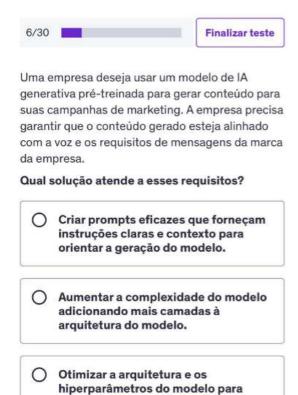
R=Importe os dados para o Amazon SageMaker Canvas. Construa modelos de ML e previsões de demanda selecionando os valores nos dados do SageMaker Canvas.

Importe os d. Beta para o Amazon de SageMaker Data Wrangler. Crie

12/30 Finalizar teste	
atender a esses requisitos? (Escolha duas.)	Uma empresa de empréstimos está
Garantir que o modelo seja executado frequentemente	construindo uma solução baseada em IA generativa para oferecer descontos a novos solicitantes com base em critérios específicos de
Usar a técnica Recall-Oriented Understudy for Gisting Evaluation (ROUGE) para garantir que o modelo seja 100% preciso.	negócios. A empresa deseja construir e usar um modelo de IA de forma responsável para minimizar vieses que possam afetar negativamente alguns clientes.
Avaliar o comportamento do modelo para que a empresa possa fornecer transparência aos stakeholders.	Quais ações a empresa deve tomar para atender a esses requisitos? (Escolha duas.)
Garantir que o tempo de inferência do modelo esteja dentro dos limites aceitáveis	
Detectar desequilíbrios ou disparidades nos dados Beta Compartilhar feedback	

R=Avaliar o comportamento do modelo para que a empresa possa fornecer transparência aos stakeholders.

Detectar desequilíbrios ou disparidades nos dados



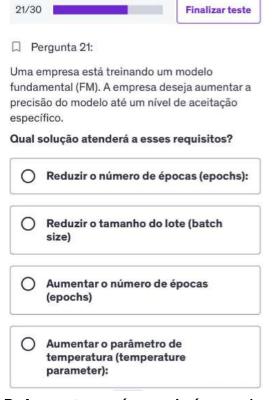
melhorar o desempenho geral do

Selecionar um conjunto de dados

Beta Compartilhar feedback

modelo.

R=Criar prompts eficazes que forneçam instruções claras e contexto para orientar a geração do modelo.



R=Aumentar o número de épocas (epochs)



Um praticante de IA está desenvolvendo um modelo para gerar imagens de humanos em várias profissões. O praticante de IA descobriu que os dados de entrada estão enviesados e que atributos específicos afetam a geração de imagens, criando vieses no modelo.

Qual técnica resolverá o problema?

0	Aumento de dados para classes desbalanceadas (Data augmentation
	for imbalanced classes)

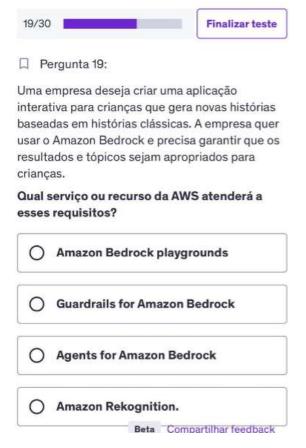
- Monitoramento do modelo para distribuição de classes (Model monitoring for class distribution)
- O Geração aumentada por recuperação
- Detecção de marca d'água para imagens (Watermark detection for images)

 Beta Compartilhar feedback

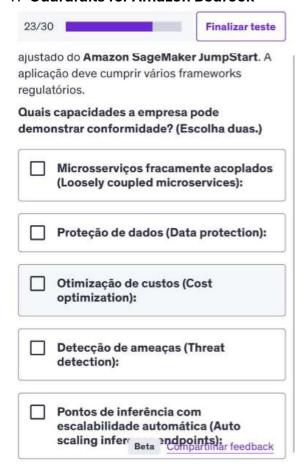
Um praticante de IA está
desenvolvendo um modelo para
gerar imagens de humanos em
várias profissões. O praticante de IA
descobriu que os dados de
entrada estão enviesados e que
atributos específicos afetam a
geração de imagens, criando
vieses no modelo.

Qual técnica resolverá o problema?

R=Aumento de dados para classes desbalanceadas (Data augmentation for imbalanced classes)



R=Guardrails for Amazon Bedrock



Uma empresa deseja implantar um chatbot conversacional para responder às perguntas dos clientes. O chatbot é baseado em um modelo ajustado do Amazon SageMaker JumpStart. A aplicação deve cumprir vários frameworks regulatórios.

Quais capacidades a empresa pode demonstrar conformidade? (Escolha duas.)

R=Proteção de dados (Data protection): Detecção de ameaças (Threat detection):

P	ergunta 28:
igua	as empresas podem usar modelos de gem grande (LLMs) com segurança no on Bedrock?
0	Projetar prompts claros e específicos. Configurar funções e políticas do AWS Identity and Access Management (IAM) usando o princípio de menor privilégio.
0	Usar o Amazon CloudWatch Logs para tornar os modelos explicáveis e monitorar vieses
0	Habilitar trabalhos automáticos de avaliação de modelos no Amazon Bedrock
0	Habilitar o A\ Beta iditiManager paracetrabalhos automáticos de avaliação

R=Projetar prompts claros e específicos. Configurar funções e políticas do AWS Identity and Access Management (IAM) usando o princípio de menor privilégio.

		Finalizar teste
□ Pe	ergunta 27:	
model base a empre	empresa precisa construir s lo de linguagem de grande apenas nos dados privados esa está preocupada com o ental do processo de treina	porte (LLM) com da empresa. A impacto
	tipo de instância Amazon DR impacto ambiental ao	
0	Amazon EC2 P series	
0	Amazon EC2 Trn series	
0	Amazon EC2 G series	
	Amazon EC2 C series	

R= Amazon EC2 Trn series

26/30		Finalizar teste
□ Pe	ergunta 26:	
copiar	studante em uma universion ndo conteúdo de uma IA g ver redações.	20 TH TO THE TOTAL STATE OF
	desafio de uso responsá ativa esse cenário repre	
0	Plagiarism (Plágio)	
0	Privacy (Privacidade)	
0	Hallucinations (Alucina	ıções).
0	Toxicity (Toxicidade).	

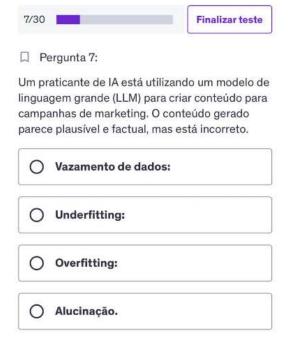
R=**Plagiarism (Plágio)**



R= Amazon Q no Amazon QuickSight

0	Tokens são os prompts específicos ou instruções fornecidos a um modelo de IA generativa para gerar saídas:
0	Tokens são as unidades básicas de entrada e saída com as quais um modelo de IA generativa opera, representando palavras, subpalavras ou outras unidades linguísticas:
0	Tokens são os pesos pré-treinados de um modelo de IA generativa ajustados para tarefas específicas:
0	Tokens são as representações matemáticas de palavras ou conceitos us Beta em modelos de la generativa:

R=Tokens são as unidades básicas de entrada e saída com as quais um modelo de IA generativa opera, representando palavras, subpalavras ou outras unidades linguísticas:



R=Alucinação

8/30		Finalizar teste
de sup dias. A respor	ande varejista recebe milha porte ao cliente sobre prod As consultas precisam ser p ndidas rapidamente. A emp mentar Agents for Amazon	utos todos os processadas e presa deseja
0	Geração de modelos fundamentais personalizados (FMs) para prever as necessidades dos clientes.	
0	Seleção do modelo fundamental (FM) com base em critérios e métricas predefinidos.	
0	Automação de tarefas repetitivas e orquestração de fluxos de trabalho complexos.	
0	Chamadas automáticas a múltiplos modelos fundamentais (FMs) e consolidação dos resultados. Beta Compartilhar feedback	
R=Aı	utomação de tarefa	
□ Pe	ergunta 10:	
promp Amazo exemp vez ao	mpresa está utilizando few-s oting em um modelo base hos on Bedrock. O modelo atualme olos no prompt. O modelo é indi dia e está apresentando bom resa deseja reduzir o custo m	spedado no ente usa 10 vocado uma n desempenho.
0	Aumentar o número de tok prompt	ens no
0	Personalizar o modelo usar tuning:	ndo fine-
0	Usar Throughput Provision	ado

R=Reduzir o número de tokens no prompt.

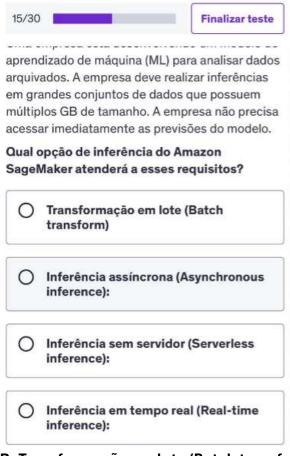
Beta Compartilhar feedback

Reduzir o número de tokens no

prompt.

□ Pe	ergunta 11:				
Uma e ao uso para g	empresa quer avaliar os cu o de um modelo de linguaç erar inferências. A empre on Bedrock para criar apli	gem grande (LLM sa deseja usar o			
Qual f inferê	ator irá influenciar os cu ncia?	istos de			
0	Tempo total de treinamento				
0	O Valor da temperatura				
0	Ouantidade de dados usados para treinar o LLM:				
0	Número de tokens con	sumidos			
12/30	ergunta 12:	Finalizar teste			
curríci que a curríci empre autom PDF p proces	empresa revisa manualmentulos enviados em formato Fempresa cresce, espera-se ulos exceda a capacidade desa. A empresa precisa de ulatizado para converter os cara o formato de texto simpossamento adicional.	PDF. À medida e que o volume de le revisão da im sistema currículos em oles para			
Quals	erviço da AWS atende a es	se requisito?			
0	Amazon Lex:				
0	Amazon Personalize:				
0	Amazon Textract				
0	Amazon Transcribe:				

R=Amazon Textract



Uma empresa está
desenvolvendo um modelo de
aprendizado de máquina (ML)
para analisar dados arquivados.
A empresa deve realizar
inferências em grandes
conjuntos de dados que
possuem múltiplos GB de
tamanho. A empresa não precisa
acessar imediatamente as
previsões do modelo.

Qual opção de inferência do Amazon SageMaker atenderá a esses requisitos?

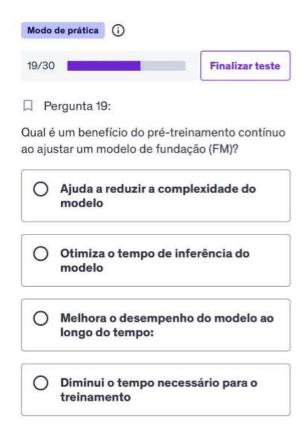
R=Transformação em lote (Batch transform)

Finalizar teste



R=Amazon S3

18/30



R=Melhora o desempenho do modelo ao longo do tempo: