

Lista 1

Inteligência Artificial

Nome: Vitor de Meira Gomes

Matrícula: 800643

Questão 1:

1.1) Viés Algorítmico é quando sistemas de IA produzem resultados injustos de forma sistemática, amplificando preconceitos existentes. Isso pode acontecer devido à falta de diversidade nos dados, classificação incorreta de atributos, dados insuficientes e decisões de design. Esses vieses podem surgir durante o desenvolvimento de modelos de IA por erros na coleta, preparação ou seleção de dados, bem como por escolhas inadequadas nos algoritmos e nas métricas de avaliação.

1.2) No documentário, a pesquisadora Joy Buolamwini, descobriu que sistemas de reconhecimento facial não detectavam seu rosto por ela ser mulher e negra. O algoritmo só funcionava corretamente quando ela colocava uma máscara branca. Isso ocorreu devido aos modelos terem sido treinados principalmente com imagens de pessoas brancas, resultando em taxas de erro muito mais altas para mulheres e pessoas negras.

1.3) Duas possíveis estratégias são: aumento da diversidade e qualidade dos dados de treinamento, e anonimização ou remoção de atributos sensíveis. A primeira opção sugere garantir que os conjuntos de dados utilizados representem adequadamente diferentes gêneros, etnias, idades e contextos socioeconômicos, reduzindo a probabilidade de resultados enviesados. Já a segunda, retirar ou mascarar dados que possam introduzir viés, como nome, etnia ou gênero, quando não forem relevantes para a tarefa, evitando que o algoritmo aprenda correlações discriminatórias.

Questão 2:

2.1) A conferência aconteceu no verão de 1956 no Dartmouth College, em Hanover, New Hampshire (EUA). Os principais participantes foram John McCarthy, Marvin Minsky, Claude Shannon e Nathaniel Rochester e a proposta central era investigar como máquinas poderiam simular aspectos da inteligência humana, partindo da hipótese de que todo aspecto do aprendizado ou qualquer outra característica da inteligência poderia ser descrito com precisão e simulado por uma máquina.

2.2) A reunião comemorativa foi realizada em 2006, no mesmo local da conferência de 1956, no Dartmouth College. O encontro reuniu pesquisadores para refletir sobre os desafios propostos no evento original, avaliar o progresso da área e discutir as limitações ainda presentes. Foram abordados avanços técnicos e teóricos, além de questões científicas e filosóficas sobre o futuro da IA. Também se discutiu a necessidade de maior poder computacional, melhores algoritmos e novas abordagens para enfrentar problemas ainda não resolvidos.

2.3) Em 1956, os participantes da conferência tinham uma visão bastante otimista, acreditando que problemas complexos da inteligência poderiam ser resolvidos em poucas décadas, com máquinas capazes de pensar de forma semelhante a humanos. Já em 2006, reconheceu-se que, embora houvesse avanços significativos em áreas específicas, a inteligência artificial geral ainda estava distante. O foco havia mudado para o desenvolvimento de sistemas especializados e eficazes em domínios restritos, enquanto desafios como compreensão de linguagem natural, raciocínio avançado e aprendizado autônomo pleno permaneciam abertos.