

Proposta Educativa 11.2: Viés Algorítmico

Reflexão: De que forma o viés algorítmico pode ter impactos negativos e frequentemente prejudiciais na nossa sociedade?

Descrição

Os alunos irão conhecer o impacto do viés algorítmico na nossa sociedade, com o suporte de documentários e outros exemplos, e a forma como este problema poderá ter resultados negativos e prejudiciais.

Objetivos

- Desenvolver a capacidade de identificar, analisar e discutir casos de viés algorítmico.
- Incentivar a reflexão sobre as implicações éticas e sociais e promover a consciencialização sobre a importância da equidade nos algoritmos.

Referencial de Aprendizagem

N2-IA-24 Compreender como o domínio do conhecimento precisa de ser amplo o suficiente para abranger todos os grupos que um aplicativo possa chegar.

N2-IA-33 Examinar características e rótulos de dados de treino para detetar potenciais fontes de preconceito.

Resultados do aluno

- Consigo identificar viés algorítmico.
- Consigo compreender as Implicações éticas e sociais de viés algorítmico, incluindo como ele pode afetar grupos específicos e contribuir para desigualdades sociais.

Vocabulário

Viés algorítmico - erros nos algoritmos, frequentemente causados por conjuntos de dados distorcidos, que criam resultados injustos, como privilegiar um determinado grupo em detrimento de outros, ou discriminar grupos específicos.

Discriminação - tratar um grupo de pessoas de forma diferente de outro grupo de forma injusta.

Olhar codificado - outro termo para "viés algorítmico"; as formas como os algoritmos tratam as pessoas injustamente.

Preparação

Requisitos tecnológicos

- Se estiver a fazer esta aula presencialmente, necessitará de um computador com acesso à Internet e a funcionalidade de partilhar à turma através de um projetor ou similar. Para a abordagem à distância, os alunos necessitarão dos seus próprios dispositivos.

Preparar recursos online

- Testar os vídeos e as hiperligações para garantir que estão a funcionar. A maioria dos vídeos do Youtube foram inseridos através da cópia URL do vídeo.

Preparar materiais

- Para esta proposta educativa não são necessários materiais específicos para a atividade.

Recursos

Atividade

- Top 4 cores

Avaliação

Perguntas incorporadas ao longo da apresentação e dos exercícios.

1 - O que é o viés algorítmico?

2 - Quais são algumas das implicações destes preconceitos no nosso mundo?

Introdução

Através do **diapositivo 3**, destaque que nesta aula será explorado o viés algorítmico e a sua implicação ética e social no mundo real, e o modo como poderá ser desenvolvida a capacidade em analisar este tipo de problemas.

Desenvolvimento

1. Passe para o **diapositivo 4** e partilhe o vídeo sobre o documentário “Coded Bias”, realizado por Dr. Joy Buolamwini.

2. A seguir, no **diapositivo 5**, analise o vídeo com os alunos:

- Qual foi problema que Dr. Joy Buolamwini identificou na tecnologia de reconhecimento facial?

(Sugestão de resposta: Os sistemas de reconhecimento facial são muito bons a reconhecerem pessoas de pele branca, mas não tão bons a reconhecerem pessoas de pele escura.)

- Porque é que isto é um problema?

(Sugestão de resposta: Pensem em algumas formas como o reconhecimento facial é utilizado, por exemplo, Snapchat/filtros de câmara, pesquisas no Google, carros autónomos. Não é justo que uma tecnologia que é anunciada como sendo concebida para toda a gente não funcione tão bem para algumas pessoas. Por exemplo, as pessoas com a pele mais escura não podem utilizar filtros de câmara divertidos ou mesmo os carros autónomos têm dificuldade em ver pessoas com pele mais escura.

Há alguns efeitos positivos: se as câmaras tiverem mais dificuldade em ver pessoas com pele mais escura, há menos hipóteses da sua privacidade ser violada pela vigilância. No entanto, isto também tem alguns riscos - e se uma pessoa inocente for acusada de um crime devido a um mau sistema de reconhecimento?)

- Como podemos ajudar a evitar este viés algorítmico?

(Sugestão de resposta: Analisar com mais cuidado o próprio algoritmo e não aceitar cegamente os seus resultados. Certificar-se de que o conjunto de dados é mais equitativo. Testar os algoritmos para garantir que não contêm enviesamentos.

3. No **diapositivo 6**, partilhe e analise o vídeo com os alunos. Questione sobre o tipo de tecnologia que está a ser utilizada (assistentes virtuais).

4. No **diapositivo 7**, partilhe os assistentes virtuais e aguarde pelas respostas dos alunos sobre a respetiva identificação dos mesmos. Destaque a importância de garantir que assistentes virtuais sejam capazes de compreenderem uma variedade de sotaques (mesmo que o exemplo do vídeo sirva “apenas” como quebra-gelo). Está intrinsecamente ligada à promoção da acessibilidade e inclusão na interação homem-máquina. Diferenças linguísticas, expressas através de sotaques, são uma parte valiosa da diversidade cultural e individual. Quando estas tecnologias são sensíveis a estas nuances linguísticas, proporcionam uma experiência mais inclusiva para um público global.

Ao não reconhecerem adequadamente diferentes sotaques, os assistentes virtuais podem excluir involuntariamente grupos de usuários, criando barreiras de comunicação. Isso não apenas limita a eficácia dessas tecnologias, mas também perpetua desigualdades e exclusões digitais. Portanto, é crucial que os desenvolvedores se empenhem em treinar e aprimorar estes assistentes virtuais para que sejam proficientes em compreenderem e responderem a uma ampla gama de pronúncias.

5. Através do **diapositivo 8** explore a pesquisa da palavra “Maçã”. Questione os alunos sobre o que poderá aparecer nas pesquisas ao nível das imagens. Aguarde as respostas.

6. No **diapositivo 9**, partilhe e analise com os alunos o conjunto de imagens que surgiram na pesquisa.

- Pergunte aos alunos se acham a pesquisa justa. Porque é que o algoritmo está a sugerir principalmente maçãs vermelhas?

- Reflita com os alunos como isso poderá influenciar a maneira como as pessoas veem as maçãs e se reflete corretamente sobre a variedade de cores das maçãs no mundo real.

- Introduza a ideia de que os algoritmos aprendem com os dados que recebem, e se esses dados não forem diversificados, o resultado também será limitado.

Sugestão de síntese: Pensar na pesquisa online por "maçã". Os algoritmos aprendem com imagens que lhes são mostradas, e se a maioria dessas imagens são de maçãs vermelhas, a pesquisa irá mostrar mais dessas maçãs. No entanto, isso não representa todas as maçãs, como as verdes ou amarelas. Assim, podemos perceber que se os algoritmos não aprenderem com diferentes tipos de maçãs, eles podem mostrar resultados que não são justos.

Nota: Este exemplo poderá ajudar os alunos a compreenderem que, se os algoritmos não são treinados de forma adequada e diversificada, podem gerar resultados distorcidos. Além disso, esta analogia poderá facilitar a discussão sobre a importância de incluir uma variedade de dados ao treinar algoritmos para garantir que sejam mais justos e equitativos.

7. Realize a atividade “Top 4 cores” (**diapositivo 10**). As instruções para a realização desta atividade estão na respetiva pasta “Proposta Educativa 12 - Inteligência Artificial”.

Sugestão

- Através do exemplo do diapositivo 23, poderão ser adicionados outras tecnologias para fazer a respetiva análise dos prós e contras.

Créditos

Esta proposta educativa foi traduzida e adaptada do projeto *MIT Media Lab & i2 Learning*



Atribuição-NãoComercial-
Compartilha Igual 4.0 Internacional
(CC BY-NC-SA 4.0)