

P2 # Programação com python

Problema 1:

Neste problema, você implementará várias funções curtas. O principal objetivo deste exercício é familiarizar-se com o Python, mas, como bônus, as funções que você implementará serão úteis em trabalhos de casa subsequentes.

Não importe nenhuma biblioteca externa (por exemplo, numpy). Somente bibliotecas Python padrão e / ou as bibliotecas importadas no código inicial são permitidas.

Se você é novo no Python, os itens a seguir fornecem dicas para vários tutoriais e exemplos da linguagem:

- [Python para programadores](https://wiki.python.org/moin/BeginnersGuide/Programmers) (<https://wiki.python.org/moin/BeginnersGuide/Programmers>)
- [Programas de exemplo de complexidade crescente](https://wiki.python.org/moin/SimplePrograms) (<https://wiki.python.org/moin/SimplePrograms>)

O que esperávamos: Código Python implementando as funções fornecidas em submit.py. Tente fazer seu código o mais limpo e simples possível e certifique-se de escrever suas respostas entre a resposta inicial e os comentários da resposta final.

- uma. [3 pontos] Implementar `find_alphabetically_first_word` no [submissão.py](#).
- b. [3 pontos] Implementar `Distância euclidiana` no [submissão.py](#).
- c. [6 pontos] Implementar `mutate_sentences` no [submissão.py](#).
- d. [4 pontos] Implementar `sparse_vector_dot_product` no [submissão.py](#).
- e. [4 pontos] Implementar `increment_sparse_vector` no [submissão.py](#).
- f. [4 pontos] Implementar `find_nonsingleton_words` no [submissão.py](#).

Problema 2: Identificação de questões éticas

Um dos objetivos deste curso é ensinar como lidar com problemas do mundo real com ferramentas de IA. Mas os problemas do mundo real têm consequências para o mundo real. Junto com as habilidades técnicas, uma habilidade importante que todo praticante de IA precisa desenvolver é a consciência das questões éticas associadas à IA. O objetivo deste exercício é praticar a identificação de possíveis preocupações éticas em aplicações de IA - mesmo as aparentemente inócuas.

Nesta pergunta, você explorará a ética de quatro diferentes cenários do mundo real usando as diretrizes éticas produzidas por um local de pesquisa de aprendizado de máquina, a conferência NeurIPS. [NeurIPSEthical Guidelines](https://neurips.cc/public/EthicsGuidelines) (<https://neurips.cc/public/EthicsGuidelines>) listar dezesseis preocupações não exaustivas em Potenciais Impactos Sociais Negativos e

Conduta Ética Geral (as listas numeradas). Para cada cenário, você escreverá uma declaração de impactos negativos potenciais. Para fazer isso, você primeiro determinará se o algoritmo / conjunto de dados / técnica pode ter um impacto social negativo potencial ou violar a conduta ética geral (novamente, os dezesseis itens numerados retirados do [Diretrizes Éticas NeurIPS](#) página). Se o cenário violar a conduta ética ou tiver potenciais impactos sociais negativos, liste uma preocupação que ele viola e justifique porque você acha que essa preocupação se aplica ao cenário. Se você fizer **não** acho que o cenário tem uma preocupação ética, explique como você chegou a essa decisão. Ao contrário dos problemas anteriores na lição de casa, existem muitas boas respostas possíveis. Se puder justificar sua resposta, você deve se sentir confiante de que respondeu bem à pergunta.

Cada um dos cenários é elaborado a partir de um artigo de pesquisa de IA real. A ética da pesquisa em IA reflete de perto as possíveis consequências no mundo real da implantação de IA, e as lições que você tirará deste exercício certamente serão aplicáveis à implantação de IA em escala. Como nota, você **é não** obrigados a ler os artigos originais, mas colocamos um link para eles para o caso de serem úteis. Além disso, você pode responder a qualquer coisa no artigo vinculado que não seja mencionada no cenário escrito, mas os cenários descritos aqui devem fornecer detalhes suficientes para localizar pelo menos uma preocupação.

O que esperávamos: Um parágrafo de 2 a 5 frases para cada um dos cenários em que você A. identifica pelo menos uma preocupação ética do [Diretrizes Éticas NeurIPS](#) e justifique porque você acha que se aplica, ou B. declare que você não acha que existe uma preocupação e justifique porque é esse o caso. Os cenários escolhidos podem ter de zero a várias preocupações correspondentes, mas você só precisa escolher uma preocupação (se houver) e justificar sua decisão de acordo. Além disso, copie e sublinhe o item da lista de verificação ética ao qual você está se referindo como parte de sua resposta (ou seja: Danificar gravemente o meio ambiente) Também incluímos uma citação na solução de exemplo abaixo, mas você não é obrigado a adicionar citações à sua resposta.

Cenário de Exemplo: Você trabalha para um hospital dos EUA que implementou recentemente um novo programa de intervenção que inscreve pacientes em risco em programas para ajudar a resolver seus problemas médicos crônicos de forma proativa, antes que os pacientes acabem no hospital. O programa de intervenção identifica automaticamente os pacientes em risco, prevendo os escores de risco dos pacientes, que são medidos em termos de custos de saúde. No entanto, você observa que, para um determinado nível de pontuação de risco, os pacientes negros ficam consideravelmente mais doentes quando inscritos do que os pacientes brancos, embora sua pontuação de risco de doença atribuída seja idêntica. Você reatribui manualmente as pontuações de risco dos pacientes com base em seus sintomas atuais e percebe que a porcentagem de pacientes negros que seriam inscritos aumentou de 17% para mais de 45%^[1].

Solução de exemplo: Este algoritmo provavelmente codificado, contém ou potencialmente exacerba o preconceito contra pessoas de uma determinada raça ou etnia uma vez que o algoritmo prevê custos de saúde. Como o acesso a cuidados médicos nos EUA é desigual, os pacientes negros tendem a ter custos de saúde mais baixos do que seus colegas brancos^[2]. Assim, o algoritmo irá prever incorretamente que eles estão em menor risco.

uma. [2 pontos] Uma empresa de investimento desenvolve um modelo simples de aprendizado de máquina para prever se um indivíduo tem probabilidade de inadimplir em um empréstimo devido a uma variedade de fatores, incluindo localização, idade, pontuação de crédito e registro público. Depois de examinar seus resultados, você descobre que o modelo prevê principalmente com base na localização e que o modelo aceita principalmente empréstimos urbanos

centros e nega empréstimos de candidatos rurais [3]. Além disso, observando o gênero e a etnia dos candidatos, você descobre que o modelo tem uma taxa de falsos positivos significativamente mais alta para candidatos negros e homens do que para outros grupos. Em uma previsão de falso positivo, um modelo classifica erroneamente alguém que não está inadimplente como provável inadimplente.

- b. [2 pontos] Estilometria é uma forma de prever o autor de um texto contestado ou anônimo, analisando os padrões de escrita no texto anônimo e em outros textos escritos pelos autores em potencial. Recentemente, algoritmos de aprendizado de máquina altamente precisos foram desenvolvidos para essa tarefa. Embora esses modelos sejam normalmente usados para analisar documentos históricos e literatura, eles podem ser usados para retirar a anonimato de uma ampla gama de textos, incluindo códigos[4].
- c. [2 pontos] Um grupo de pesquisa retirou milhões de rostos de celebridades das imagens do Google para desenvolver tecnologia de reconhecimento facial[5]. As celebridades não deram permissão para que suas imagens fossem usadas no conjunto de dados e muitas das imagens estão protegidas por direitos autorais. Para fotos protegidas por direitos autorais, o conjunto de dados fornece links de URL para a imagem original, juntamente com caixas delimitadoras para o rosto.
- d. [2 pontos] Pesquisadores criaram recentemente um modelo de aprendizado de máquina que pode prever espécies de plantas automaticamente, diretamente de uma única foto[6]. O modelo foi treinado usando fotos enviadas para o aplicativo iNaturalist por usuários que consentiram em usar suas fotos para fins de pesquisa, e o modelo só é usado dentro do aplicativo para ajudar os usuários a identificar plantas que possam encontrar na natureza.

MELHOR!!!