

THIAGO DA SILVA ADRIANO

POSTECH

IA PARA DEVS

TECH CHALLENGE

FASE 01

TECH CHALLENGE

Tech Challenge é o projeto que engloba os conhecimentos obtidos em todas as disciplinas dessa fase. Esta é uma atividade que, em princípio, deve ser desenvolvida em grupo. É importante atentar-se ao prazo de entrega, uma vez que essa atividade é obrigatória, valendo 60% da nota de todas as disciplinas da fase.

O problema

Você é um profissional encarregado de desenvolver um modelo preditivo de regressão para prever o valor dos custos médicos individuais cobrados pelo seguro de saúde.

A base de dados para este desafio pode ser algo como no exemplo a seguir:

```
idade,gênero,imc,filhos,fumante,região,encargos
56,feminino,29.774373714007336,2,sim,sudoeste,31109.88976
3423336
46,masculino,25.857394655216346,1,não,nordeste,26650.7026
46642694
32,masculino,23.014839993647488,0,não,sudoeste,21459.0379
9039332
```

Você precisa apenas alimentar ela com mais informações ou utilizar uma outra de sua preferência.

Tarefas

Exploração de dados:

- Carregue a base de dados e explore suas características.
- Analise estatísticas descritivas e visualize distribuições relevantes.

Pré-processamento de dados:

- Realize a limpeza dos dados, tratando valores ausentes (se necessário).
- Converta variáveis categóricas em formatos adequados para modelagem.

Modelagem:

- Crie um modelo preditivo de regressão utilizando uma técnica à sua escolha (por exemplo, Regressão Linear, Árvores de Decisão etc).
- Divida o conjunto de dados em conjuntos de treinamento e teste.

Treinamento e avaliação do modelo:

- Treine o modelo com o conjunto de treinamento.

Validação estatística:

- Utilize métricas estatísticas para validar a eficácia do modelo (p-value, intervalos de confiança).

O que avaliaremos:

Apresente resultados visuais, como gráficos de previsões vs. valores reais.

Elabore um relatório que inclua uma análise dos resultados, insights obtidos e validação estatística.

Observações:

Esperamos que o modelo seja capaz de fazer previsões confiáveis dos custos médicos individuais com base nas características fornecidas.

Entregável:

Como entregável, o grupo deve enviar um vídeo junto com o link do github do projeto e o código desenvolvido, apresentando o passo a passo do que foi utilizado como a fonte de dados e como os modelos foram criados.

O vídeo deve estar em uma plataforma como Youtube.

Qualquer dúvida, não deixe de nos chamar no Discord: lá poderemos tirar suas dúvidas e auxiliar nesta tarefa. Boa sorte!



POS TECH