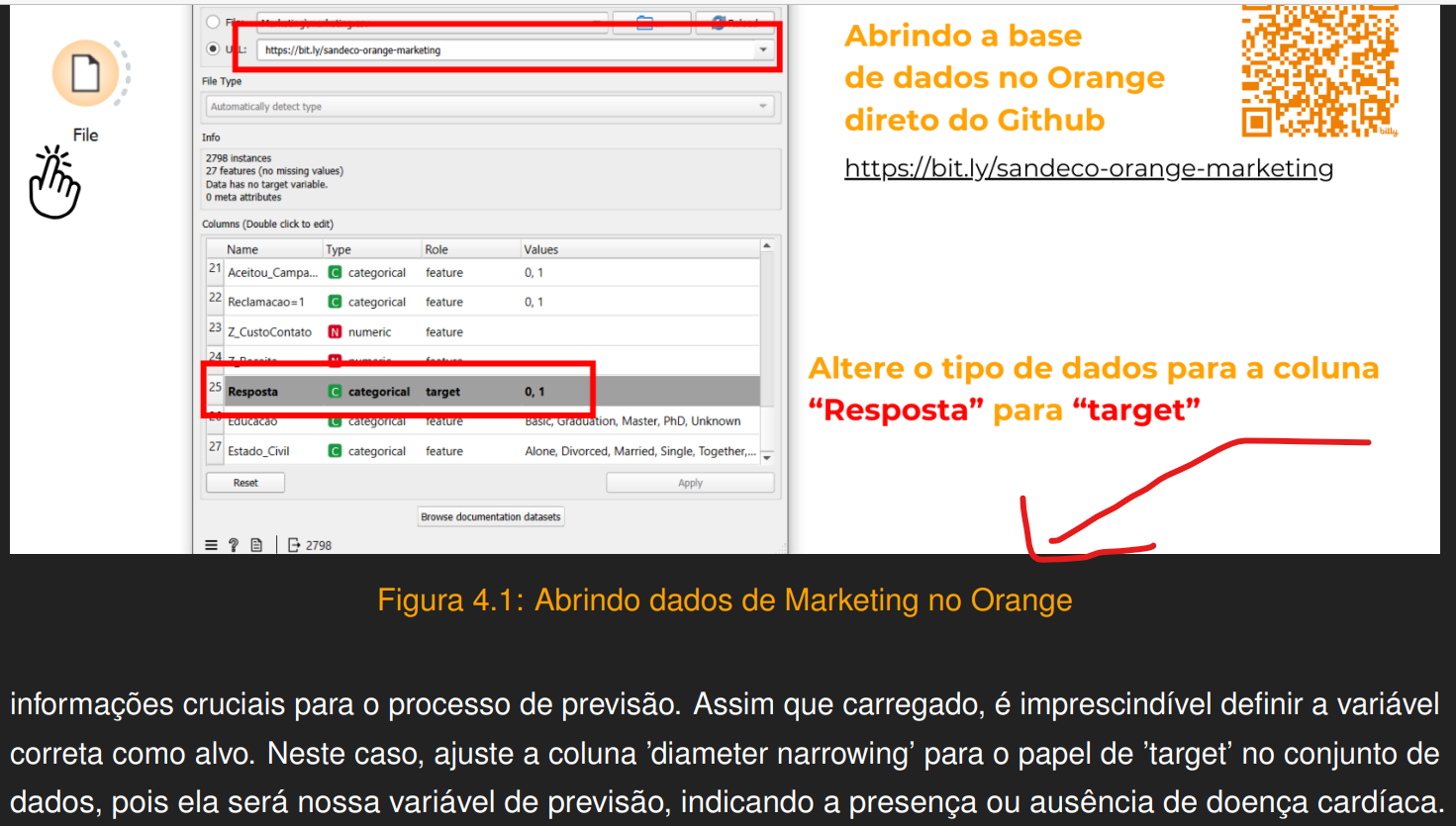
Esto é um teste de controle de versão usando GITHUB

Na pagina 70 o link leva para o dataset de marketing

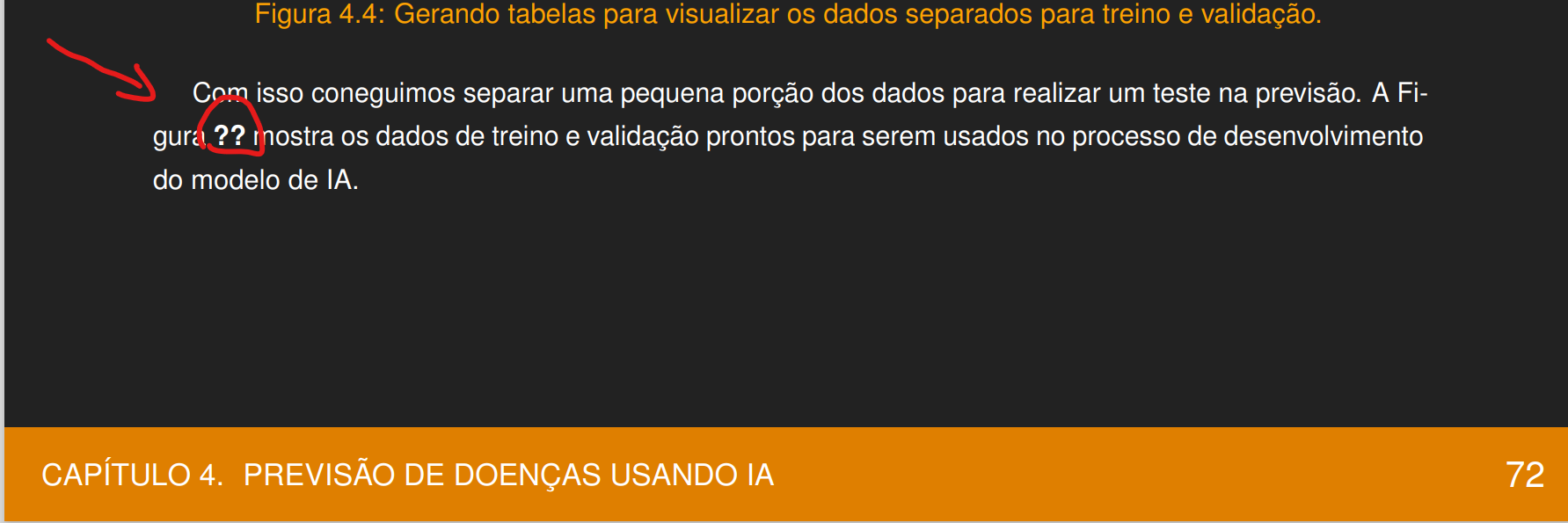
Site

Descrição gerada automaticamente com confiança média

Na pagina 71 a imagem inteira faz referencia ao dataset marketing e o link do Github tambem.

****

**Eu acho que na pagina 72 falta subtituir “??” pelo 4.5**

****

**Deveria ser em maiusculo**

**Texto

Descrição gerada automaticamente**

**Linha do tempo

Descrição gerada automaticamente com confiança média**

Diagrama

Descrição gerada automaticamente com confiança baixa

O Orange Canvas pode ser uma ferramenta extremamente útil **para uma imobiliária, permitindo-lhe realizar desde a análise exploratória de dados até a modelagem preditiva para diversas finalidades.** Aqui estão alguns exemplos reais de como uma imobiliária pode aplicar o Orange Canvas:

### 1. \*Análise de Tendências do Mercado Imobiliário\*

- \*Visualização de Dados\*: Use widgets de visualização para entender as tendências do mercado imobiliário, analisando a evolução dos preços, a demanda por diferentes tipos de propriedades, e as localizações mais valorizadas.

- \*Previsão de Preços\*: Através de modelos de regressão, como regressão linear ou florestas aleatórias, preveja os preços de imóveis com base em características como localização, tamanho, e amenidades. Isso ajuda a aconselhar vendedores sobre o preço de listagem e compradores sobre ofertas justas.

### 2. \*Segmentação de Clientes\*

- \*Clustering\*: Utilize técnicas de clustering para segmentar seus clientes em grupos baseados em preferências de propriedades, capacidade de investimento, e outros critérios relevantes. Isso permite à imobiliária personalizar comunicações e recomendar propriedades que sejam mais relevantes para cada segmento de cliente.

### 3. \*Previsão da Demanda de Mercado\*

- \*Modelos Preditivos\*: Implemente modelos preditivos para antecipar a demanda futura por certos tipos de propriedades ou em determinadas localizações. Isso pode informar estratégias de aquisição de imóveis e ajudar a identificar onde investir em novas construções.

### 4. \*Análise de Localização\*

- \*Geo-Referenciamento\*: Integre dados de localização com mapas e dados demográficos para identificar áreas de alto potencial para diferentes tipos de investimentos imobiliários. Analise fatores como proximidade a escolas, áreas comerciais, e transporte público.

### 5. \*Identificação de Imóveis Subvalorizados\*

- \*Análise Comparativa\*: Utilize o Orange para comparar preços de imóveis com base em suas características e localização, identificando assim imóveis potencialmente subvalorizados. Essa informação pode ser valiosa tanto para investidores quanto para a própria imobiliária.

### Exemplo Prático: Previsão de Preços de Imóveis

Vamos detalhar como a imobiliária pode usar o Orange para prever os preços dos imóveis:

1. \*Coleta de Dados\*: Compile um conjunto de dados de listagens de imóveis anteriores, incluindo características dos imóveis (localização, metros quadrados, número de quartos, etc.) e o preço de venda.

2. \*Pré-processamento de Dados\*: Use widgets de pré-processamento para limpar os dados, lidar com valores ausentes, e transformar variáveis categóricas em formatos utilizáveis.

3. \*Análise Exploratória\*: Empregue widgets de visualização para explorar relações entre características dos imóveis e seus preços de venda.

4. \*Modelagem Preditiva\*: Selecione widgets de regressão para construir modelos preditivos com base nos dados processados. Experimente diferentes algoritmos e compare suas performances.

5. \*Avaliação e Implementação\*: Avalie o modelo mais eficaz usando dados de teste e implemente-o para prever preços de novas listagens, ajudando assim a definir estratégias de precificação mais informadas.

Ao empregar o Orange Canvas com essas abordagens, uma imobiliária pode melhorar significativamente suas operações, desde a tomada de decisão estratégica até a personalização de serviços para clientes, potencializando suas chances de sucesso no mercado.

**Tema Meio Ambiente**

**Insight**

Segmentação de Paises: Utilizar técnicas de clustering para segmentar os países em grupo de altamente poluídos e não poluídos, para trazar estratégias para minimizar o impacto negativo na atmosfera

**Tema Meio Ambiente.** Fatores de conversão, são padrões para calcular a quantidade de dióxido de carbono que os setores econômicos gastam na sua atividade. Seria bom um dataset que nos permitisse avaliar a sua evolução ao longo dos anos e.g. desde 2017, 2018 e 2023, englobando, consumo de energia elétrica, gás natural, água, alimentos, produção de resíduos e energia despendida em viagens pendulares e de trabalho.

Há 6 gases abrangidos no protocolo de Quioto: dióxido de carbono (CO2), metano (CH4), óxido nitroso (N2O), hidrofluorocarbonetos (HFC), compostos perfluorados (PFC) e hexafluoreto de enxofre (SF6). Há dados com isso por industria e por países. É uma ideia, que ainda estou a adaptar com base no que disse o Daniel ontem

As empresas/organizações realizam atividades que geram gases de efeito estufa, por exemplo, durante a fabricação, transmissão ou consumo de energia. Produzem produtos/bens de consumo que também emitem gases de efeito estufa antes, durante e depois de sua vida útil.

Podíamos associar os gases de efeito de estufa às fontes de emissão e sugerir alguns fatores de conversão ou alternativas.

//////////////////////////////////////////

Podemos fazer a comparação entre Setores:, através da comparação das emissões de GEE entre diferentes setores, como agricultura, indústria, transporte e construção. Isso pode ajudar a identificar os setores que são os maiores emissores e que podem exigir mais atenção em termos de políticas de mitigação.

Podemos também fazer uma análise geográfica, analisando as emissões de GEE por país (não temos a região geográfica por país). Isso pode revelar disparidades entre países "desenvolvidos" e "em desenvolvimento", não podemos fazer as variações regionais dentro de um país

O Orange também permite fazer previsões e podíamos fazer projeções futuras, utilizando os dados históricos. Fazer projeções sobre as emissões futuras de GEE em diferentes cenários, considerando fatores como crescimento econômico, mudanças tecnológicas e políticas climáticas (embora ai seja difícil, mas podemos imaginar).

Acho que ainda dará para avaliar como as emissões de GEE em diferentes setores e países contribuem para a vulnerabilidade às mudanças climáticas e assim ajudar a identificar estratégias de adaptação adequadas (só imaginando, pois fica fora do escopo do estudo).

NOTAS  
1. Os gases efeitos estufa impactam no aquecimento global e mudanças climaticas

**Tema 3.**

**Câncer de Pele**

Ao desenvolver um modelo de machine learning para classificar câncer de pele usando imagens no Orange, a seleção de métodos estatísticos e testes depende das várias etapas do processo de desenvolvimento e validação do modelo. Vamos abordar algumas das técnicas e métodos que poderiam ser relevantes para este cenário e solicitar um melhor entendimento com Sandeco:

### Pré-processamento e Análise Exploratória

1. \*Análise Descritiva\*: Para entender a distribuição das características básicas dos dados, como faixa etária dos pacientes, distribuição de classes (tipos de câncer de pele), etc.

2. \*Visualizações de Dados\*: Gráficos e imagens para visualizar as diferenças entre categorias de lesões cutâneas.

### Seleção e Extração de Características

3. \*Análise de Componentes Principais (PCA)\* ou \*Autoencoders\*: Para reduzir a dimensionalidade das imagens e destacar as características mais relevantes.

### Treinamento do Modelo

4. \*Matriz de Confusão e Métricas Associadas\*: Precisão, recall, F1-score e suporte para avaliar a performance do modelo em cada classe.

5. \*Curva ROC e AUC\*: Para cada classe, especialmente em um cenário multiclasse, para entender a capacidade do modelo de discriminar entre classes.

6. \*Validação Cruzada\*: Para garantir que o modelo seja generalizável e robusto, utilizando várias subdivisões dos dados para treinar e testar o modelo.

### Comparação e Otimização de Modelos

7. \*Teste t para Amostras Independentes ou Teste de Wilcoxon\*: Se comparar diferentes modelos ou hiperparâmetros, para determinar se as diferenças no desempenho são estatisticamente significativas.

8. \*ANOVA ou Kruskal-Wallis (para dados não paramétricos)\*: Quando comparar três ou mais modelos ou configurações de hiperparâmetros.

### Avaliação Pós-modelagem

9. \*Técnicas de Interpretabilidade de Modelos\*: Como LIME ou SHAP, para entender como diferentes características das imagens contribuem para as previsões do modelo.

10. \*Teste de Sensibilidade\*: Para avaliar como pequenas alterações nas imagens (como variações de iluminação ou ângulo) afetam as previsões do modelo.

### Testes de Generalização

11. \*Teste de Desempenho em Dados Externos\*: Avaliar o modelo em um conjunto de dados externo, que não foi usado durante o treinamento, para testar sua capacidade de generalização.

Esses métodos ajudam a garantir que o modelo seja preciso, confiável e robusto, minimizando os falsos positivos e negativos, que são particularmente críticos em aplicações médicas como a classificação do câncer de pele. Além disso, a interpretabilidade e a generalização são aspectos cruciais, considerando a importância clínica e a variedade de apresentações do câncer de pele.

Hijo,

Recuerde algo, un amigo es un DIAMANTE, tipo Guilherme y Leo, los de la facultad tu estas conociendo y ellos a ti, no te extralimite con la confianza porque OJOS VEMOS CORAZONES NO SABEMOS.