



Uma nova geração  
de especialistas em  
**tecnologia**

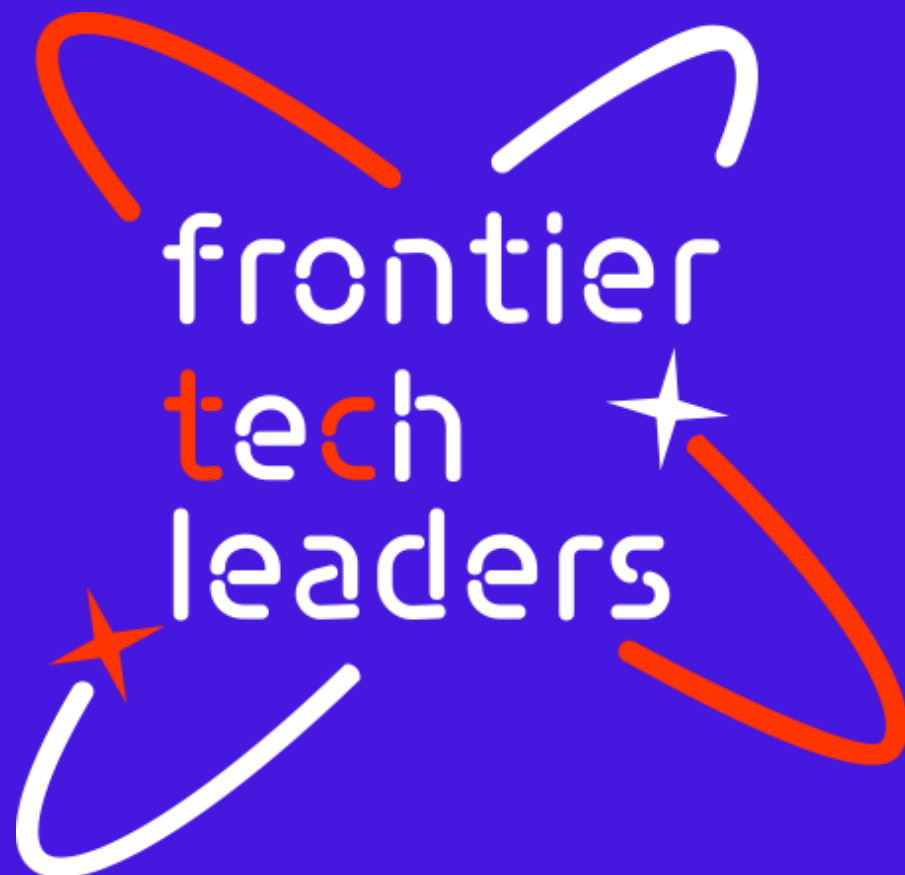




## Grupo 6

Alberto Pereira  
Elizabeth Lopes  
Júlio Cesar





# Prevtour

Previsão de demanda  
turística



# geral

O projeto **Prevtour** tem como objetivo prever a taxa de ocupação de hotéis, a procura por atrações e os fluxos de visitantes em pontos turísticos específicos com alguns dias de antecedência.

Essa previsão permite que os gestores de turismo e hotelaria planejem recursos, otimizem operações e tomem decisões estratégicas baseadas em dados.

# Background

O turismo é um dos setores mais importantes da economia, mas sofre com grande variabilidade na demanda.

Antecipar picos de procura ajuda a reduzir desperdícios e melhorar o planejamento operacional.

O projeto aplica técnicas de **aprendizado de máquina** para modelar a demanda turística, integrando dados históricos e variáveis contextuais.

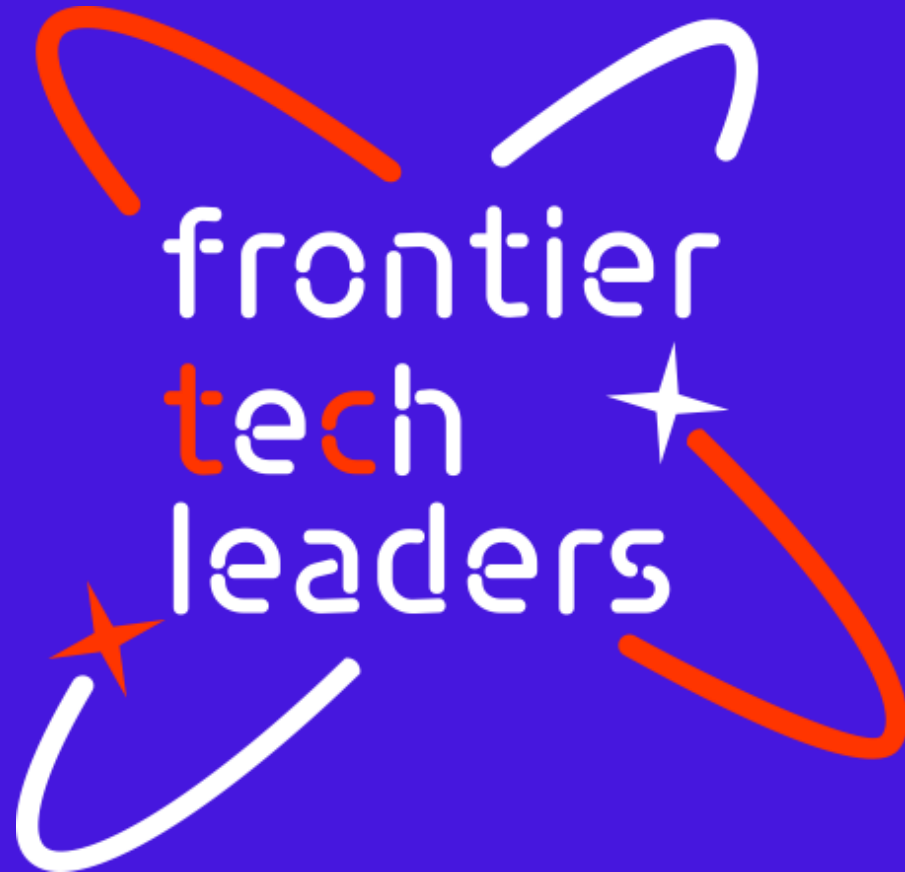
# Objetivos

- Prever a taxa de ocupação e demanda turística com alguns dias de antecedência.
- Avaliar diferentes modelos de machine learning e deep learning para determinar o mais eficaz.
- Apoiar a tomada de decisão no setor de turismo, promovendo sustentabilidade

## Relação com ODS

Este projeto contribui para os ODS 8 (Trabalho Decente e Crescimento Econômico) e ODS 12 (Consumo e Produção Responsáveis), permitindo o uso mais racional de recursos e evitando excesso de infraestrutura turística.

Ao prever a demanda, o projeto reduz desperdícios e melhora a sustentabilidade do setor.



**Dados**





# Dados

- **Fonte:** Dados públicos do Kaggle (formato .csv).
- **Pré-processamento:**
  - Codificação categórica via *OneHotEncoder*.
  - Limpeza de valores ausentes e outliers.
  - Seleção de features relevantes por correlação.
- **Observação:** Uma das colunas (“estado\_reserva”) foi removida por causar *overfitting*

# Coleta de Dados

Os dados foram obtidos de plataformas abertas (Kaggle) e processados para se adequar ao modelo.

Foram analisados e transformados para manter consistência e qualidade, garantindo que os conjuntos de treinamento, validação e teste tivessem proporções equilibradas.

# Análise Exploratória de Dados(EDA) e Engenharia de Recursos

Durante a EDA, identificaram-se padrões de sazonalidade e variáveis correlacionadas à ocupação hoteleira.

Foram criadas novas features baseadas em tempo, localização e histórico de ocupação.

As técnicas de seleção passo a passo e filtro estatístico bivariado foram utilizadas para escolher os atributos mais relevantes.



# Modelo



# Seleção e Treinamento / Teste de Modelos

Foram testados os seguintes algoritmos:

- **K-Nearest Neighbors (KNN)**
- **XGBoost (Árvore de Decisão)**
- **Regressão Logística**

Técnicas utilizadas:

- **Validação cruzada:** *cross\_val\_score (cv=3)*.
- **Ajuste de hiperparâmetros:** *GridSearchCV*.
- **Métricas avaliadas:** Precisão, Revogação, Acurácia e F1-Score.

# Avaliação de Modelos

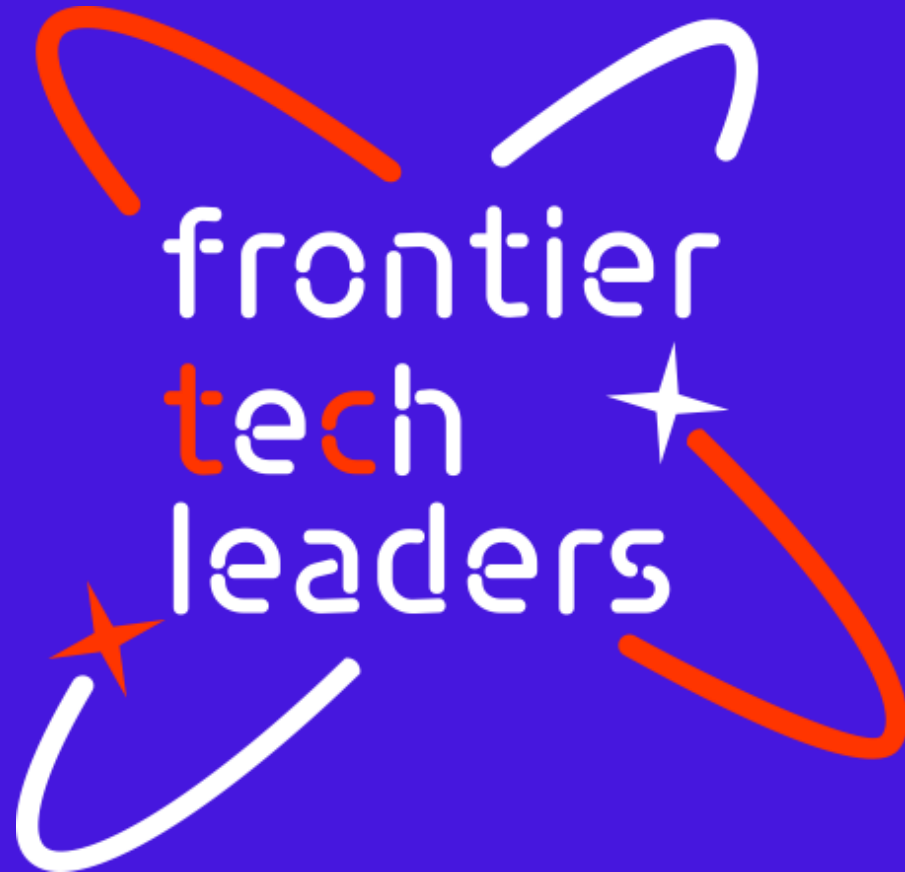
- O melhor modelo obteve:
  - **Precisão:** 76%
  - **Revogação:** 67%
  - **Acurácia:** 80%
  - **F1-Score:** 71%
- Na validação cruzada, obteve média de **65%**, o que mostra consistência, mas ainda espaço para refinamento.

# Refinamento e Teste de modelos

A técnica **GridSearchCV** foi usada para otimizar os hiperparâmetros.

Após eliminar variáveis redundantes e ajustar o número de grupos na validação cruzada (cv=3), o modelo atingiu equilíbrio entre generalização e performance.

O *overfitting* foi mitigado com seleção manual e redução de colunas correlacionadas.



# Resultado



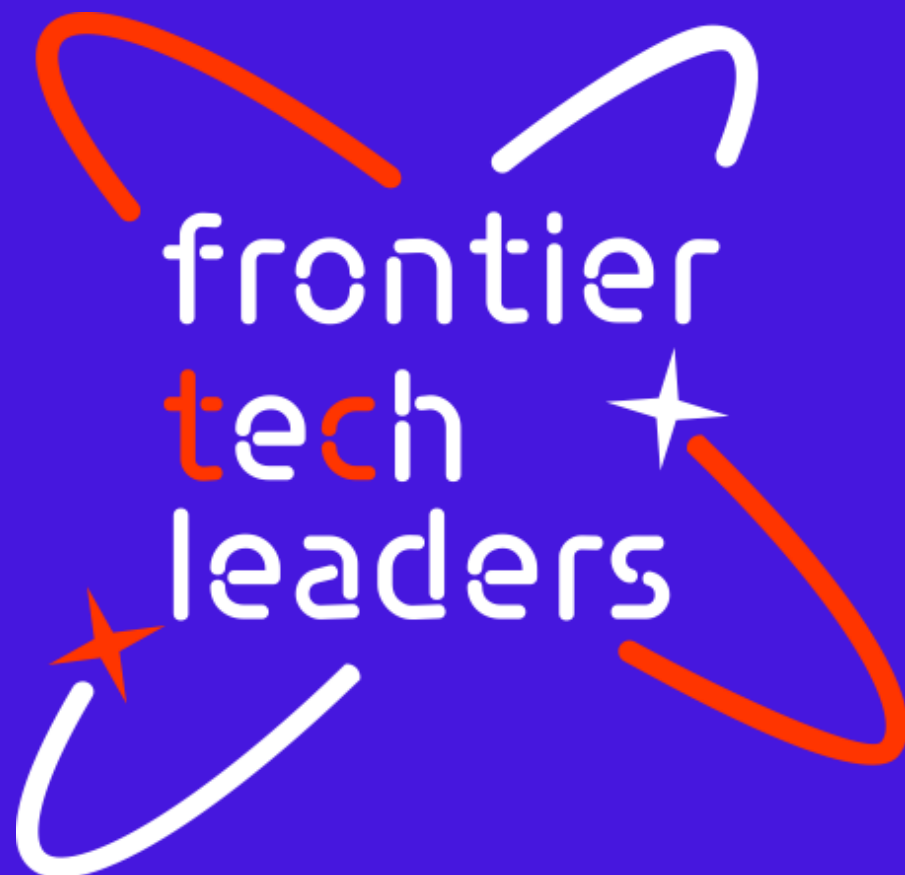


# Resultados

Nos testes, o modelo obteve:

- **Precisão:** 73%
- **Revogação:** 67%
- **F1-Score:** 70%

A ligeira queda de desempenho em relação ao treino indica boa generalização e estabilidade do modelo.



# Implantação



# Implantação

O modelo pode ser colocado em produção dentro de um **container**, integrado a uma **plataforma de gestão hoteleira** via API REST.

Isso permitirá previsões automáticas e visualização em tempo real da demanda esperada, ajudando gestores a agir de forma proativa.

## Conclusão e Trabalho futuro

O projeto demonstrou o potencial do aprendizado de máquina para apoiar decisões estratégicas no turismo.

Os principais desafios foram o tratamento de *bias* nos dados e o controle de *overfitting*.

Como próximos passos:

- Ampliar a base de dados com variáveis climáticas e eventos locais.
- Testar modelos de deep learning como **LSTM** e **Transformers**.
- Criar um painel interativo com previsões em tempo real.



# Obrigado!

