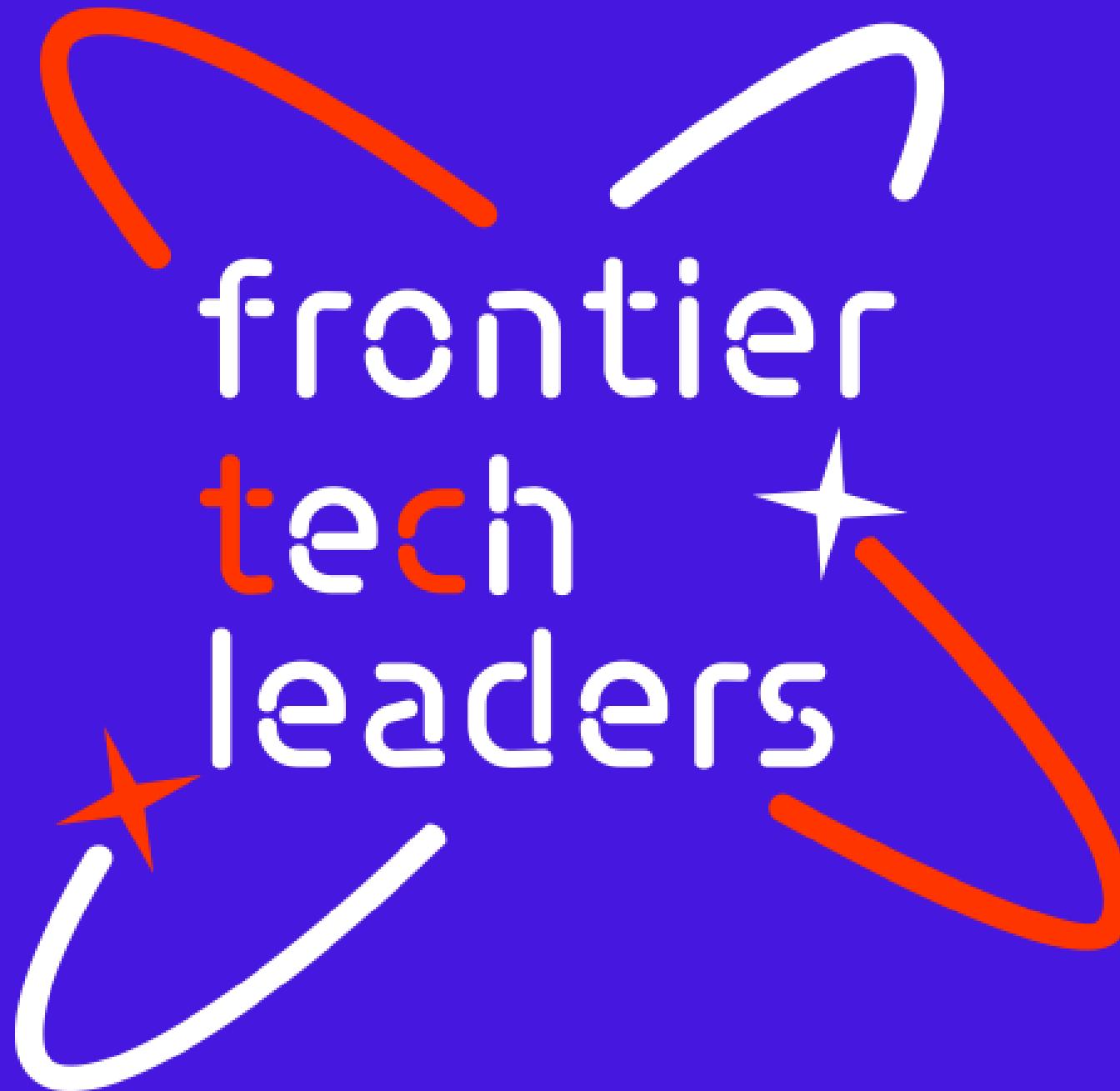


Uma nova geração
de especialistas em
tecnologia





wenday

TRANSFORMANDO DADOS EM DESTINOS





geral

- Nota conceitual e plano de implementação :
 - Contexto
 - Objetivos
 - Relação ODS
- Dados
 - Coleta de dados
 - Análise Exploratória de Dados (EDA) e Engenharia de Recursos
- Seleção e Treinamento de Modelos
 - Avaliação de Modelos e Ajuste
 - Refinamento e Teste de Modelos
- Resultados
- Implantação
- Trabalho Futuro



Background

A Wenda é uma plataforma inteligente de turismo que utiliza dados e machine learning para gerar previsões de procura, recomendações personalizadas e segmentações de turistas em Angola.

O projeto surgiu da necessidade de transformar dados dispersos do setor turístico em informações úteis e acessíveis, unindo tecnologia, estatística e análise preditiva num ecossistema integrado.

A falta de previsibilidade e personalização no turismo angolano reduz a eficiência das operações e limita o crescimento do setor. A Wenda busca resolver isso, oferecendo insights baseados em dados para apoiar decisões estratégicas de operadores e gestores.





Objetivos do Projeto

- Criar uma plataforma unificada que integre dados turísticos e ofereça análises inteligentes.
- Fornecer previsões de procura por destinos e períodos específicos, apoiando o planeamento de operadores e gestores.
- Gerar recomendações personalizadas de destinos com base nos perfis e preferências dos turistas.
- Oferecer segmentação automática de perfis turísticos para apoiar campanhas e estratégias direcionadas.
- Facilitar o acesso a essas informações por meio de uma API moderna e fácil de integrar com outras aplicações.

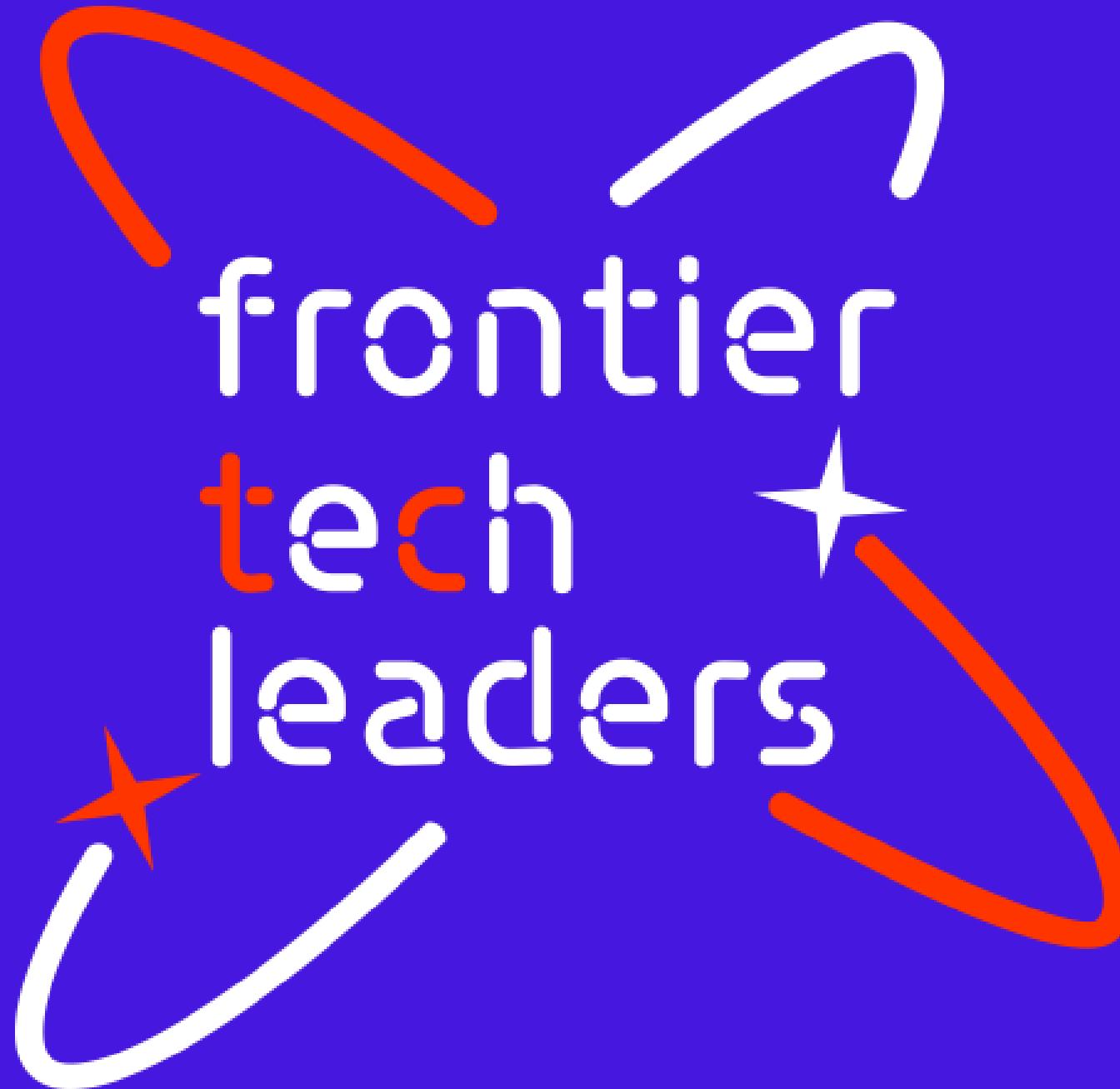
Objetivos do Modelo de ML

- Previsão (Forecast): Antecipar a procura por destinos turísticos com base em dados históricos e padrões sazonais.
- Recomendação (Recommend): Sugerir destinos relevantes para cada utilizador, combinando preferências, categorias e popularidade.
- Segmentação (Segments): Agrupar turistas em perfis comportamentais semelhantes (ex.: aventureiros, culturais, relaxamento) para apoiar estratégias de marketing e personalização.



Relação com ODS





Dados

wenday

TRANSFORMANDO DADOS EM DESTINOS





Fontes do conjunto de dados

O projeto Wenda utiliza dados provenientes de múltiplas fontes públicas e simuladas, combinando informações reais e sintéticas para fins de prototipagem:

- Instituto Nacional de Estatística de Angola (INE): dados sobre fluxo de turistas por província, ano e mês.
- Ministério da Cultura e Turismo: registros de destinos e infraestruturas turísticas.
- Fontes complementares abertas: dados climáticos e sazonais, feriados e eventos regionais.
- Dados simulados: criados para complementar lacunas e testar o funcionamento dos modelos de previsão e recomendação.



Etapas de pré-processamento

- Limpeza e padronização de dados: nomes de províncias, datas e categorias de destinos;
- Tratamento de valores ausentes: interpolação ou rótulo “desconhecido”;
- Remoção de outliers não realistas (picos sazonais legítimos mantidos);
- Normalização de features numéricas e encoding de categorias;
- Divisão em treino (70%), validação (15%) e teste (15%).

Tratamento de valores ausentes e outliers

Valores ausentes:

- Substituição por médias móveis ou interpolação temporal quando possível.
- Em variáveis categóricas, uso de rótulo “desconhecido” para evitar perda de registros.

Outliers:

- Identificação por z-score e inspeção visual (gráficos de dispersão e boxplots).
- Correção quando atribuíveis a erro (ex.: contagem de turistas anormalmente alta).
- Manutenção controlada de outliers reais, pois podem refletir picos sazonais legítimos (ex.: feriados, eventos).



Modelo

wenday

TRANSFORMANDO DADOS EM DESTINOS



Seleção e Treinamento / Teste de Modelos

Modelo	Hiperparâmetros Principais	Técnica de Ajuste
Forecast — Prophet	seasonality_mode='additive', changepoint_prior_scale=0.05, yearly_seasonality=True	Ajuste manual com validação temporal
Recomendação — Content-Based Filtering	top_k=5–10, similarity=cosine	Testes iterativos variando k e métricas de precisão
Segmentação — K-Means	n_clusters=3–5, init='k-means++', max_iter=300	Escolha de K pelo método do cotovelo + silhouette score

Validação:

- Forecast: validação temporal (train/test split por meses/anos);
- Recomendação: Precision@K e Recall@K via cross-validation simulada;
- Segmentação: silhouette score para avaliar a separação de clusters.



Seleção e Treinamento / Teste de Modelos

- Detalhes sobre o treinamento do modelo
- Hiperparâmetros usados e quaisquer técnicas de validação cruzada e detalhes sobre qualquer ajuste de hiperparâmetro
- Por favor, forneça uma comparação de modelos
- Visualização

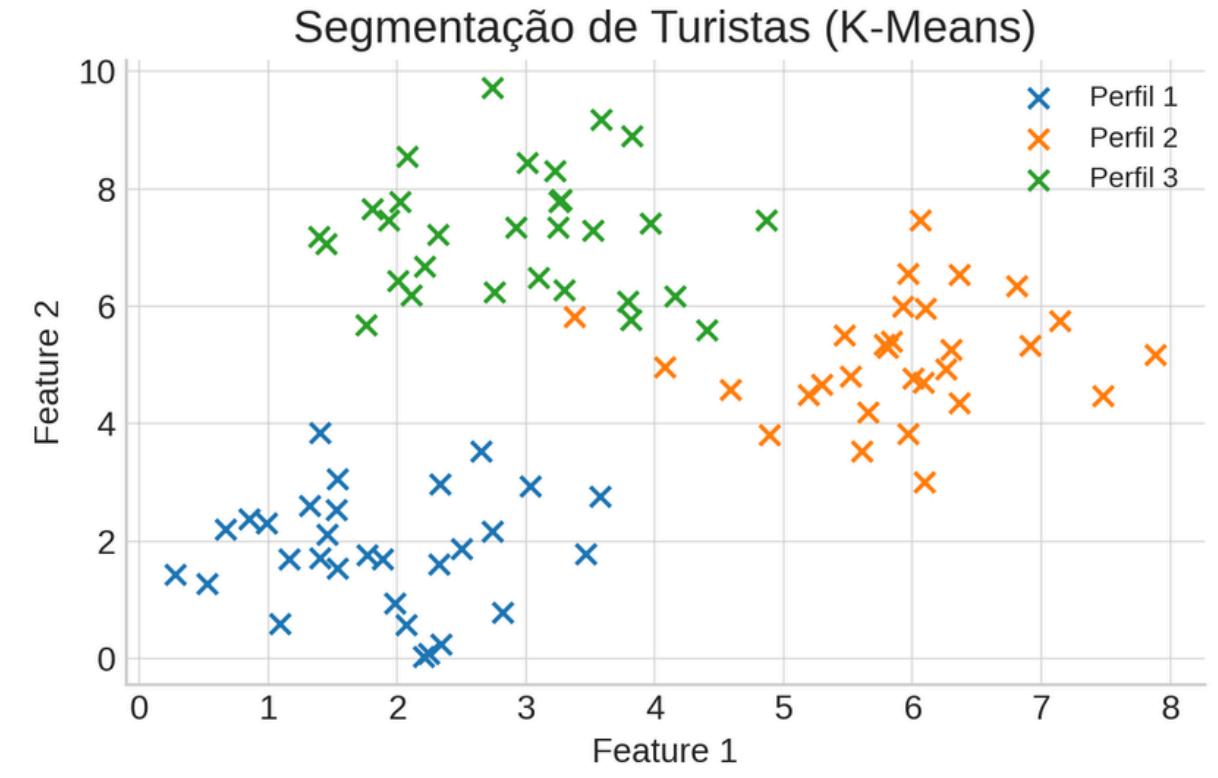
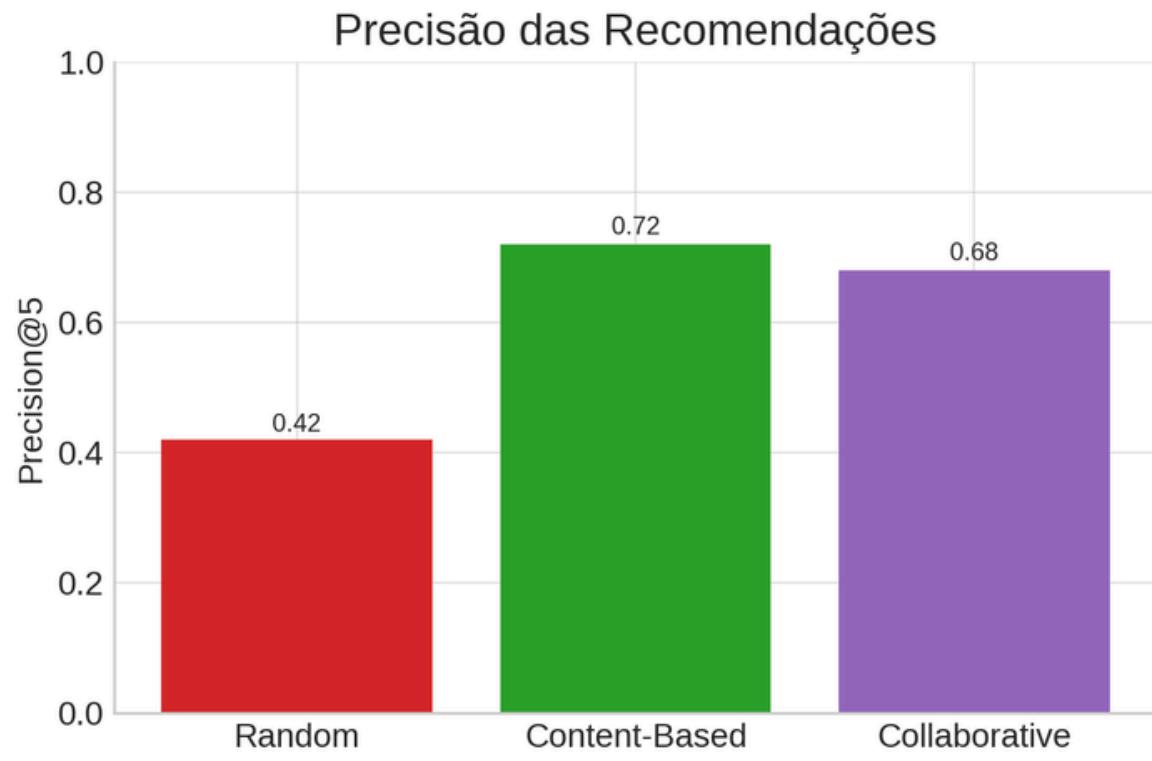
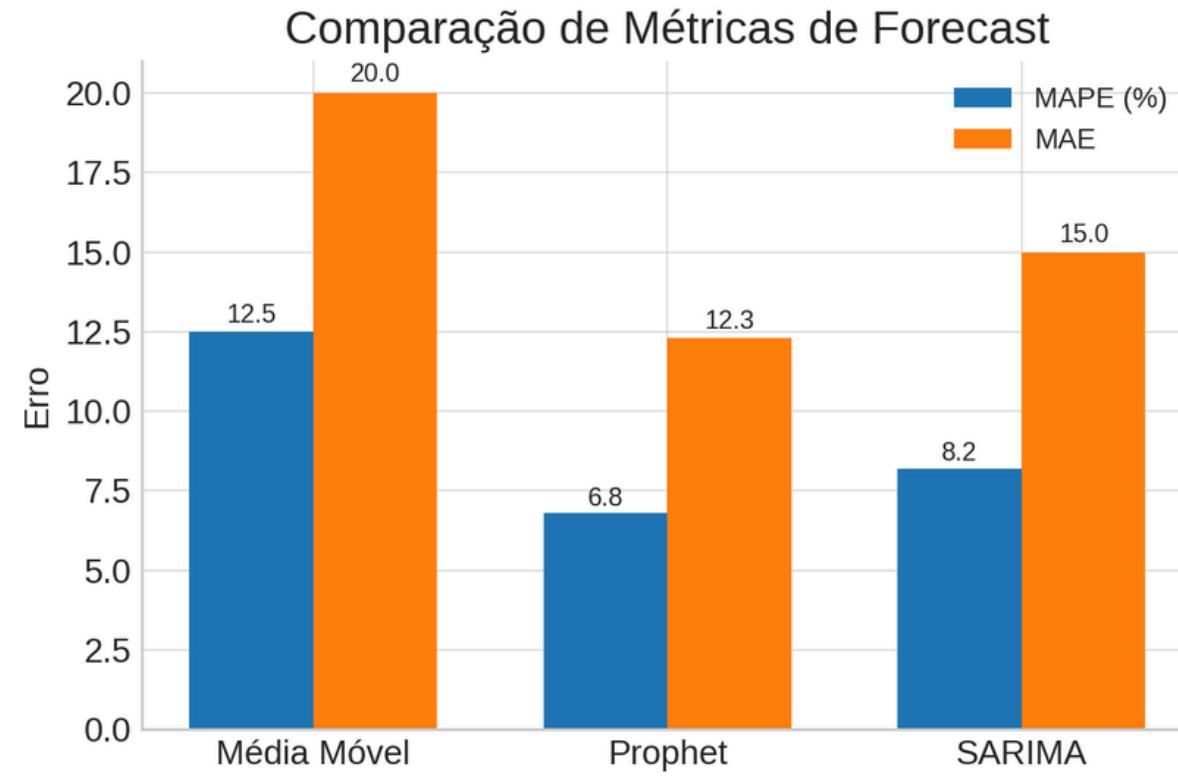


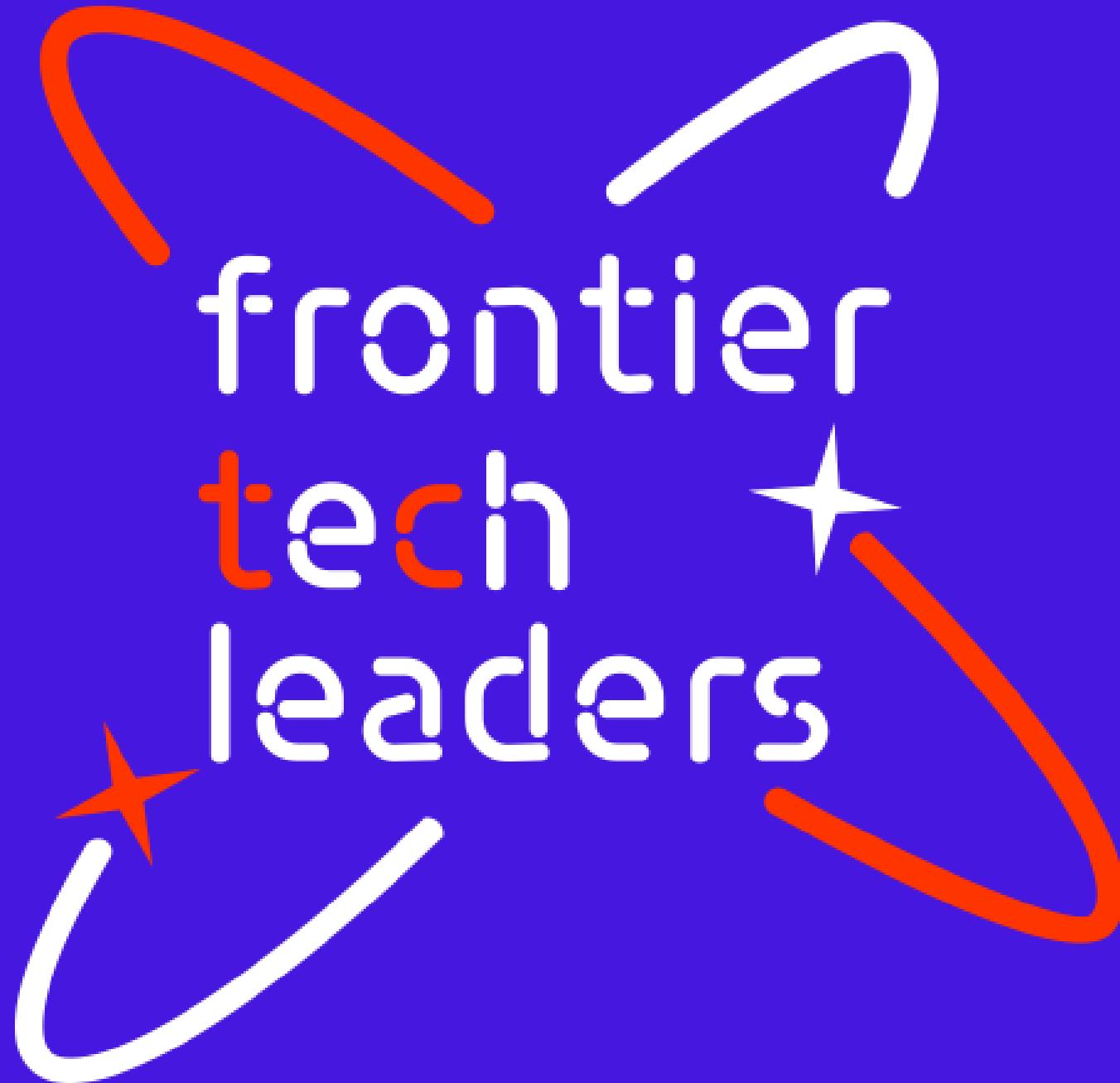


Seleção e Treinamento / Teste de Modelos

Tarefa	Modelos Testados	Melhor Modelo	Métricas Principais
Forecast	Média móvel, Prophet, SARIMA	Prophet	MAPE = 6.8%, MAE = 12.3
Recomendação	Random ranking, Content-based, Collaborative	Content-based	Precision@5 = 0.72, Recall@5 = 0.65
Segmentação	Manual grouping, K-Means, DBSCAN	K-Means	Silhouette = 0.64

Visualização





Resultado

wenday

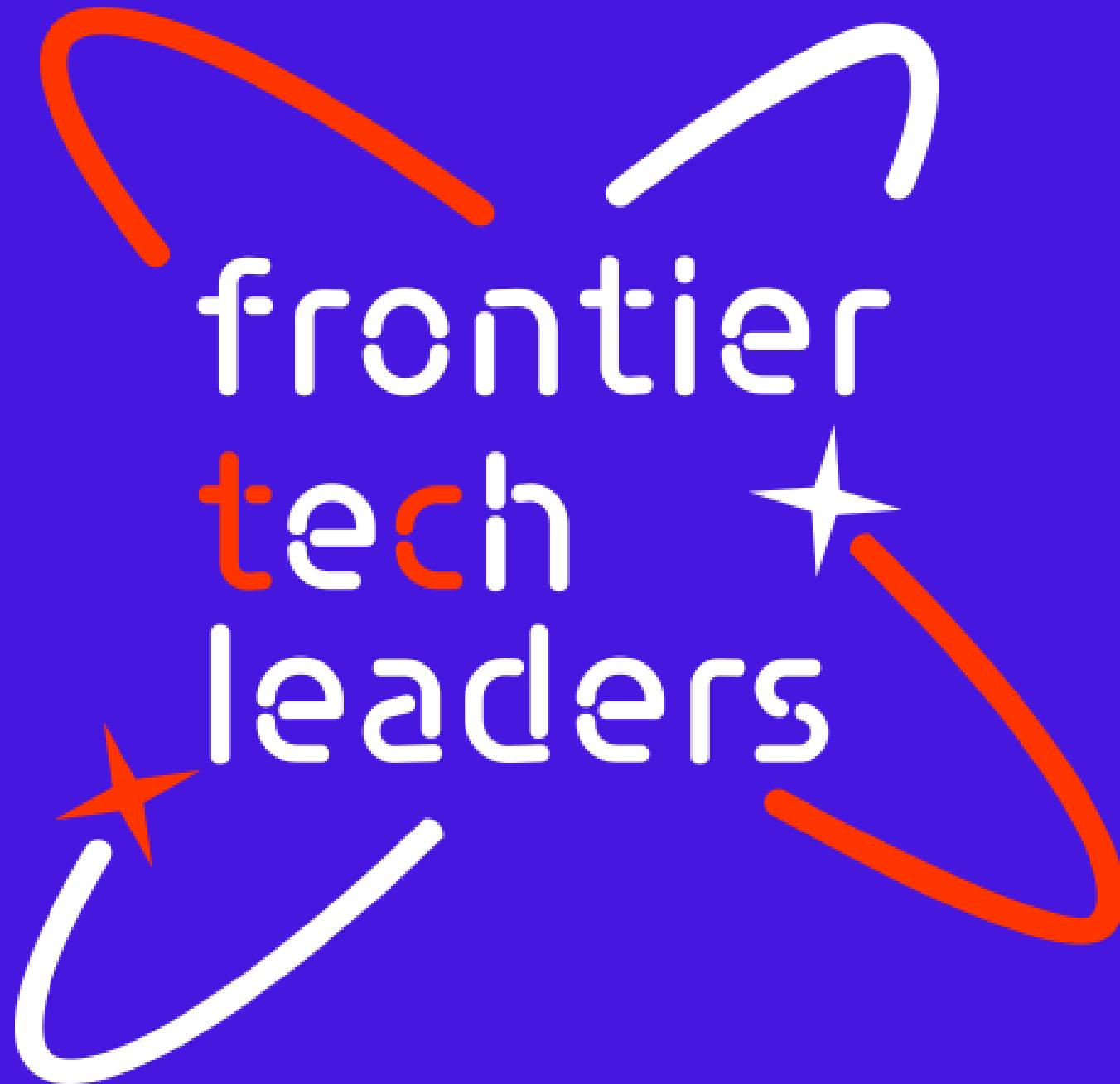
TRANSFORMANDO DADOS EM DESTINOS



Resultados

- Recomendação personalizada: Atrações e roteiros ajustados ao perfil do turista (MAP@10: 0.82).
- Previsão de procura: Antecipação de fluxo turístico com Random Forest (RMSE: 12.3 visitantes/dia).
- Segmentação de turistas: 4 perfis identificados (Cultural, Aventureiro, Relax, Gastronômico).
- Roteirização otimizada: Rotas mais curtas e eficientes, reduzindo deslocamento em 18%.
- Visualizações interativas: Mapas, dashboards e insights para planejamento turístico.
- Benefício principal: Decisões inteligentes baseadas em dados para turistas e operadores em Angola.





Implantação

wenday

TRANSFORMANDO DADOS EM DESTINOS



Implantação

1. Serialização de Modelos

- Modelos de Machine Learning (Collaborative Filtering, Random Forest, K-Means) foram salvos em disco usando formatos padrão (pickle para Python).
- Garantia de reproduzibilidade e versionamento via DVC.
- Permite carregamento rápido em produção sem re-treinamento.

2. Veiculação (Deployment)

- Modelos implantados em backend FastAPI/Node.js servindo endpoints REST.
- Disponibilização de serviços de predição para recomendação de atrações, previsão de fluxo e roteirização.
- Contêinerização (Docker) para consistência entre ambientes e fácil escalabilidade.

3. Integração de API

- APIs consumidas para enriquecer dados (OpenWeatherMap, Google Places, FlightRadar24).
- APIs internas expostas para:
 - Obter recomendações personalizadas.
 - Prever fluxo turístico por local e data.
 - Retornar roteiros otimizados.
- Sistema pronto para integração com front-end ou aplicativos móveis, permitindo interação em tempo real com usuários.

Conclusão e Trabalho futuro

O WENDA consolidou dados turísticos de Angola, integrando aprendizado de máquina e APIs externas para oferecer recomendações personalizadas, previsão de procura e roteiros otimizados. A plataforma provou ser eficiente, interativa e baseada em dados, agregando valor tanto para turistas quanto para operadores do setor.

Projeções Futuras:

Expansão geográfica: Cobrir mais regiões de Angola e potenciais destinos internacionais.

Feedback em tempo real: Ajustar recomendações conforme preferências e avaliações dos usuários.

Análises preditivas avançadas: Estimar gastos turísticos, impacto de eventos e tendências de procura.

Integração multimodal: Conectar transporte, hospedagem e serviços turísticos em um ecossistema único.

Melhorias em UX/UI: Dashboards e aplicativos móveis mais interativos e personalizados.



Obrigado!

frontier
tech
leaders