

Gestão de demanda aérea: um estudo com árvores de decisão e modelos de “ensemble”

Rafael Carneiro da Silveira^{1*}; Ana Julia Righetto²

¹ Lab Minds. Consultor em soluções em pricing & dados. São Paulo, SP, Brasil

² Alvaz. Head in R&D and Customer Experience. Av. Ayrton Senna da Silva, nº 600 - Sala 602 – Gleba Fazenda Palhano; 86050-460 Londrina, PR, Brasil

*autor correspondente: rcarneirosil@gmail.com

Gestão de demanda aérea: um estudo com árvores de decisão e modelos de “ensemble”

Resumo

A revolução tecnológica e os avanços em inteligência artificial estão transformando radicalmente as empresas, especialmente no que se refere à otimização de estratégias de posicionamento de mercado e gestão de demanda em companhias aéreas. Este estudo explora o potencial de modelos supervisionados de “machine learning” — incluindo Árvore de Decisão, “Random Forest” e “XGBoost” — na automação de decisões tarifárias, com foco na eficiência e precisão em cenários de mercado altamente dinâmicos e competitivos. Ao aplicar esses modelos, foi possível observar como cada abordagem se adapta às variações de demanda, reforçando a importância de técnicas de “ensemble” na captura de padrões complexos e na maximização da competitividade. Os resultados mostram que os modelos “ensemble”, particularmente o “XGBoost”, são poderosos aliados na adaptação às flutuações de demanda, oferecendo maior precisão e capacidade de generalização em múltiplos domínios. Com isso, este estudo fornece insights estratégicos valiosos para companhias aéreas que buscam otimizar suas práticas de “revenue management”, utilizando tecnologias avançadas para ajustar rapidamente suas estratégias tarifárias e fortalecer sua posição no mercado.

Palavras-chave: gestão de receita, companhias aéreas, aprendizado de máquina, otimização de sistemas.

Introdução

“Revenue management” refere-se à gestão de receita e demanda em empresas que operam em mercados dinâmicos, como nos setores de aviação e hotelaria. Trata-se de uma prática essencial para otimizar a receita em indústrias de capacidade fixa, que possuem uma oferta limitada de serviços ou produtos e, portanto, não permitem a flexibilidade de ajustar preços conforme o volume em resposta às variações na demanda (Talluri & Van Ryzin, 2004). Como esse setor é caracterizado por dinâmicas oligopolistas, estratégias econômicas cuidadosamente planejadas são cruciais para a sobrevivência das empresas (Phillips, 2011). A concorrência limitada, típica de um oligopólio, torna a gestão de tarifas um aspecto fundamental para garantir competitividade no mercado. Assim, a precisão na definição de preços e a habilidade de adaptar-se às flutuações do mercado são indispensáveis para maximizar receitas e sustentar uma posição estratégica sólida.

Sob essa perspectiva, existe um desafio único comparado a outros setores quanto à precificação, devido à natureza dinâmica da diferenciação tarifária e segmentação de clientes (Talluri & Van Ryzin, 2004). A complexidade do negócio exige a implementação de ferramentas de gestão, como sistemas conhecidos, a exemplo de “SABRE” e “PROS”, que já integram algoritmos de otimização (Phillips, 2011). Nesse contexto, o uso de algoritmos de

“machine learning” é amplamente presenciado na literatura do setor aéreo, com modelos preditivos que utilizam redes neurais, regressões e séries temporais para estimação de variáveis-chave (Talluri & Van Ryzin, 2004).

Especificamente no Brasil, a parte da gestão de demanda em companhias aéreas envolve forte intervenção humana na definição de classes de tarifa, especialmente na gestão dos voos no curto prazo, com antecedências variando de 0 a 30 dias antes da decolagem. As alterações diárias consistem no ajuste das classes tarifárias de cada voo, baseadas em parâmetros específicos. Esse modelo de gestão por classes foi inspirado na implementação bem-sucedida da “American Airlines” em 1985, que estabeleceu um método eficaz para gerenciar a capacidade limitada de assentos em resposta às dinâmicas de mercado. A criação de classes tarifárias com mecanismos específicos de alocação de assentos e preços tornou-se um exemplo de sucesso dessas estratégias (Anderson, Bell & Kaiser, 2003).

A gestão de demanda, portanto, está no coração do “revenue management” e, segundo Phillips (2011), representa um dos pilares fundamentais para maximização de receita em mercados de capacidade fixa. É também destacada a importância de utilizar algoritmos que permitam ajustes frequentes e leitura em tempo real das tarifas, garantindo que a aplicação das tarifas esteja alinhada às variações de demanda e ocupação.

O objetivo deste trabalho foi avaliar, de forma prática e comparativa, como os modelos baseados em árvores como Árvore de Decisão, “Random Forest” e “XGBoost” podem otimizar a gestão de demanda no setor aéreo, com foco na definição e ajuste das classes tarifárias em voos de curta antecedência. Ao investigar a aplicação desses modelos no ajuste diário de tarifas, o estudo pretende fornecer percepções importantes sobre como tecnologias de “machine learning” podem melhorar a tomada de decisão em “revenue management” como um todo, maximizando a receita e fortalecendo a posição competitiva das companhias aéreas.

Material e Métodos

Para enfrentar os desafios da tarifação dinâmica, este estudo emprega duas principais técnicas de aprendizado supervisionado de “ensemble”: “Bagging” e “Boosting”. O “Bagging”, exemplificado pelo modelo “Random Forest” (Breiman, 1996), funciona ao combinar previsões de várias árvores de decisão, o que ajuda a reduzir a variância e a melhorar a robustez do modelo. Já o “Boosting”, com o “XGBoost” como principal representante (Chen & Guestrin, 2016), corrige erros sequenciais, refinando a acurácia ao focar em padrões complexos de dados. Ambas as abordagens foram escolhidas pela capacidade de responder rapidamente às oscilações de demanda, o que é essencial para ajustar tarifas em ambientes de alta volatilidade, como o setor aéreo. Este estudo também utiliza a árvore de decisão simples

(“Decision Tree”) como ponto de referência (Quinlan, 1986), por sua simplicidade e interpretabilidade, permitindo uma análise comparativa da eficácia dos métodos.

O estudo utiliza também uma base de dados real anonimizada fornecida por uma companhia aérea, contendo informações sobre voos dentro de uma antecedência de curto prazo de 30 dias corridos. A base é composta por 11.368 registros, abrangendo uma variedade de voos e suas respectivas classes tarifárias.

Para garantir a integridade das análises, a base de dados passou por um processo de “data wrangling”, no qual foram aplicadas transformações para padronizar e tratar os dados de forma adequada. Posteriormente, a base foi dividida em conjuntos de treino e teste, utilizando uma divisão estratificada de 70% para treino e 30% para teste (James et al., 2013). A estratificação foi realizada com base na variável dependente ‘Classe’, que representa as classes tarifárias, assegurando que a proporção de cada classe fosse mantida em ambos os conjuntos. Isso é essencial para garantir que o modelo receba uma amostra representativa e equilibrada durante o treinamento e a avaliação (Kuhn & Johnson, 2013).

Neste projeto, os modelos de Árvore de Decisão, “Random Forest” e “XGBoost” foram aplicados em paralelo, utilizando os dados previamente estratificados. A fase de teste foi realizada não apenas nos 30% da base original, mas também em uma amostra externa contendo dados de um domínio, ou mercado, distinto. A inclusão dessa amostra externa permitiu uma avaliação mais robusta da capacidade de generalização dos modelos (Hastie, Tibshirani, & Friedman, 2009), analisando seu desempenho em contextos de mercado variados, além de validar sua eficácia na precificação dinâmica e segmentação tarifária.

Os modelos foram avaliados utilizando acurácia para medir o desempenho geral, enquanto a matriz de confusão forneceu percepções detalhadas sobre os erros de classificação (sobretudo em classes adjacentes). Precisão e especificidade foram essenciais para identificar a qualidade das previsões positivas e negativas, respectivamente. O F1-score, por sua vez, garantiu uma avaliação equilibrada em casos de classes desbalanceadas. Finalmente, o índice de competitividade integrou essas métricas com fatores diretamente relacionados ao negócio e sua realidade operacional.

Variáveis do Estudo

Dada a natureza do negócio, foi utilizada uma gama de variáveis que, na rotina diária, auxilia os analistas na tomada de decisão sobre a alteração de tarifas. Entre elas, destaca-se a variável CMT (Indicador de competitividade atual da classe), que indica se a tarifa praticada no momento presente ou em $d-1$ ¹ está em linha com a concorrência. Essa variável é de vital

¹ O termo $d-1$ refere-se ao dia imediatamente anterior à data de referência no estudo, enquanto $d+1$ indica o dia subsequente. Essa nomenclatura é amplamente utilizada em análises de séries temporais para representar variações temporais relativas à data base ($d0$).

importância dentro do contexto do setor aéreo, onde o monitoramento constante das tarifas dos competidores pode garantir que a empresa mantenha sua competitividade no mercado, ajustando suas próprias tarifas de forma estratégica (Phillips, 2011).

As demais variáveis são de classificação binária (“dummy”) ou categóricas mais amplas com representações numéricas adequadas. Vale destacar também a presença de variáveis de “delay” na composição da base (CMs e DMs) que permite a interpretação das defasagens para a captura correta dos impactos de eventos passados.

Tabela 2. Variáveis do Modelo

Variável	Descrição
CLASSE	Variável dependente - Classe a ser aplicada para tarifação
DTD	Quantidade de dias até a decolagem
ADVP	Agrupamento da antecedência do voo
CAP	Capacidade da aeronave
Season	Flag de temporada: alta ou baixa
DOW	Dia da semana
CLUSTER	Cluster de tarifação inicial
BKD	Booked Load Factor - Ocupação de assentos em %
EXP	Expectativa - Cálculo de projeção de demanda em %
VAR_EXP	Variação de Expectativa versus dia anterior
Revenue	Receita
RASK	Indicador de receita do negócio
TM	Tarifa média
FX	Projeção de expectativa
DM1	Ocupação de assentos d-1 em pontos percentuais
DM2	Ocupação de assentos d-2 em pontos percentuais
DM3	Ocupação de assentos d-3 em pontos percentuais
LW	Ocupação de assentos versus semana anterior em pontos percentuais
TLW	Variação das transações de passagem versus semana anterior em p.p.
CNX_percent	Abertura para venda de conexões em %
CL	Classe mínima para abertura de conexão
CM1	Classe aplicada em d-1
CM2	Classe aplicada em d-2
CM3	Classe aplicada em d-3
CMT_0	Último indicador de competitividade – variável binária
Match_min	Classe em que se precisa chegar para ficar igual ao competidor em tarifa
CMT	Indicador de competitividade atual da classe – variável binária
Fluxo	Flag de fluxo: fluxo, contrafluxo ou indiferente

Fonte: Base de dados originais

Outro ponto importante a acrescentar é a exclusão de valores atrelados à classe de tarifa, a variável dependente, dado que o processo de gestão de demanda envolve uma técnica chamada “steering”, que consiste na ação de alocar diferentes classes tarifárias conforme as condições do mercado. Essa técnica permite que a companhia ajuste

dinamicamente as classes de forma a otimizar a ocupação dos voos, deslocando a oferta para maximizar a receita. A escolha da classe atribui uma tarifa em um segundo momento, subordinada a uma estrutura própria que não será abordada neste trabalho.

As classes de tarifa, por sua vez, são enumeradas de 0 a 12, com 0 correspondendo à maior e 12 à menor. Essas 13 opções são alocadas conforme a necessidade das demais variáveis, como indicado na Tabela 2, onde se encontra o conjunto completo de variáveis que compõem a base de dados. As variáveis foram escolhidas de acordo com a relevância no processo de gestão de demanda, ou seja, somente constam aquelas que possuem relação direta com a tarifação e impactam o ajuste das classes tarifárias. Dessa forma, a base é composta por fatores como sazonalidade, taxa de ocupação e características da rota, elementos essenciais para a definição de estratégias do setor aéreo (Phillips, 2011).

Base de dados

A Base de dados consiste em dois domínios de mercados de operação distintos, representados por amostras de duas regiões de atuação da companhia aérea no Brasil. Nos dois casos das bases aplica-se a lógica da divisão entre treino (70%) e teste (30%) e uma amostra de teste distinta, denominada “cross-domain” (Ganin, Y. et al. 2016), que contém os dados do mercado do domínio 2 e será aplicada sobre os modelos das duas bases. Essa massa de teste servirá como ponto de partida para testar o quanto os modelos são generalistas e como se comportam com dados distintos sob mesma estrutura.

Essa divisão se configura da seguinte forma:

- a. **Base 1:** representação integral do domínio 1 com 11.368 registros;
Nessa base o objetivo é mensurar a capacidade dos modelos em entender os padrões de comportamento dos dados dentro de um mesmo domínio, o que serve de parâmetro na avaliação de escolha entre ter um modelo mais abrangente ou um modelo mais específico ao próprio domínio.
- b. **Base 2:** representação parcial dividida 50% domínio 1 e 50% domínio 2 com 6.218 registros;
Nesse cenário a base 2 serve de referência para identificar se é possível um mix de domínios para se obter um modelo mais generalista que atenda a situações distintas de comportamento dos dados.

Aplicação dos Modelos

A árvore de decisão é um modelo de aprendizado supervisionado que classifica ou prediz resultados ao segmentar dados em subconjuntos, baseando-se em regras de decisão simples. Cada nó representa uma pergunta ou condição sobre uma variável, e cada ramo segue um possível resultado dessa pergunta, conduzindo a novos nós ou a uma previsão final (Quinlan, 1986).

O “Random Forest” é um modelo supervisionado de aprendizado de máquina que combina várias árvores de decisão para melhorar a precisão das previsões. Cada árvore é treinada em uma amostra aleatória dos dados, e em cada etapa de divisão, apenas alguns dos recursos são considerados. Depois que todas as árvores são construídas, elas fazem previsões de forma independente. Para problemas de classificação, que é o caso deste trabalho em si, o modelo escolhe a classe final com base na votação da maioria das árvores (Breiman, 2001).

Já o “XGBoost” é um modelo supervisionado de aprendizado de máquina que constrói múltiplas árvores de decisão de forma sequencial. Diferente do “Random Forest”, ele ajusta as árvores seguintes com base nos erros cometidos pelas árvores anteriores, aprimorando a capacidade do modelo de corrigir previsões erradas. Esse processo contínuo de ajuste torna o “XGBoost” adequado em capturar padrões complexos nos dados (Chen & Guestrin, 2016).

A seguir, são apresentados os hiperparâmetros de “Random Forest” na Tabela 3 e “XGBoost” na Tabela 4, utilizados no treinamento dos dois principais modelos, selecionados com base no critério de melhor acurácia observada na fase inicial de teste. Vale destacar que a acurácia se refere à taxa de acerto do modelo, ou seja, a proporção de previsões corretas feitas pelo modelo treinado quando aplicado à parte restante da base de dados (30%) utilizada para teste. Essa métrica é fundamental para avaliar o desempenho do modelo em dados que ele não viu durante o treinamento, refletindo sua capacidade de generalização.

Tabela 3. Hiperparâmetros do Modelo “Random Forest”

Parâmetro	Valor	Descrição
“n_estimators”	500	Número de árvores
“random_state”	123	Parâmetro para garantir reprodutibilidade
“max_depth”	25	Profundidade máxima de cada árvore
“min_samples_split”	4	Número mínimo de amostras para divisão
“max_features”	sqrt	Número máximo de recursos para divisão

Fonte: Base de dados originais

Tabela 4. Hiperparâmetros do Modelo “XGBoost”

Parâmetro	Valor	Descrição
“n_estimators”	1000	Número de árvores
“random_state”	123	Parâmetro para garantir reprodutibilidade
“use_label_encoder”	False	Desabilita o codificador de rótulos
“eval_metric”	mlogloss	Métrica de avaliação
“learning_rate”	0,01	Taxa de aprendizado do modelo
“max_depth”	8	Profundidade máxima de cada árvore
“subsample”	0,8	Fração de amostras usadas para cada Árvore
“colsample_bytree”	0,8	Fração de recursos usados para cada Árvore
“reg_alpha”	0,5	Regularização L1 para evitar overfitting
“reg_lambda”	1,5	Regularização L2 para evitar overfitting

Fonte: Base de dados originais

Tanto o modelo de Árvore de Decisão, “Random Forest” e “XGBoost” passaram pelo teste “K-fold” (parâmetro 5) para a validação das variações dentro da própria base com treino e teste e o resultado apresentou dispersão baixa. Sobre a execução com os hiperparâmetros, é reconhecido que é possível combinar diversos parâmetros para encontrar a acurácia desejada, no entanto, como foi constatado a seguir a acurácia para esse segmento de negócio não é o único e principal componente de avaliação, visto que a diferença entre classes adjacentes na etapa teste não reflete necessariamente em erro.

Uma preocupação adicional e complementar a esse fato é sobre o risco do modelo apresentar “overfitting”, que ocorre quando o modelo se ajusta excessivamente aos dados de treino, capturando ruídos e padrões específicos que não generalizam bem para novos dados, resultando em um desempenho inferior em bases de teste ou em produção. Por outro lado, há também o risco de “underfitting”, que acontece quando o modelo é incapaz de capturar a complexidade dos dados de treino, resultando em um ajuste insuficiente. Nesse caso, o modelo se torna excessivamente simplificado, deixando de aprender padrões importantes, o que prejudica tanto o desempenho nos dados de treino quanto nos novos dados. Portanto, o desafio é encontrar um equilíbrio adequado entre esses dois extremos, de modo que o modelo seja capaz de generalizar bem para diferentes conjuntos de dados ou, até mesmo, identificar se o melhor caminho seria especializar um modelo para cada domínio (Hastie, Tibshirani, & Friedman, 2009).

Resultados e Discussão

Foi possível visualizar uma evolução de acurácia a partir do modelo simples de Árvore de Decisão até chegar no mais moderno “XGBoost” nos testes com as Bases 1 e 2, conforme ilustrado nas Figuras 1 e 2. Nota-se, também, indícios de “overfitting” nos modelos treinados somente com a Base 1, visível na etapa de teste de validação cruzada. Já a Base 2

apresentou acurácia razoavelmente próxima quando treinada com a parcela de 30% de teste e na etapa de validação cruzada.

Esses resultados sugerem que a escolha do modelo e o ajuste de seus hiperparâmetros devem ser orientados pela finalidade específica da aplicação. Em situações em que o modelo precisa ser especializado em um mercado particular, o ajuste mais focalizado pode ser necessário, enquanto modelos destinados a múltiplos domínios se beneficiam de uma abordagem mais generalista. Assim, é essencial definir se o objetivo é capturar padrões complexos em um contexto amplo ou maximizar a precisão para um segmento específico.

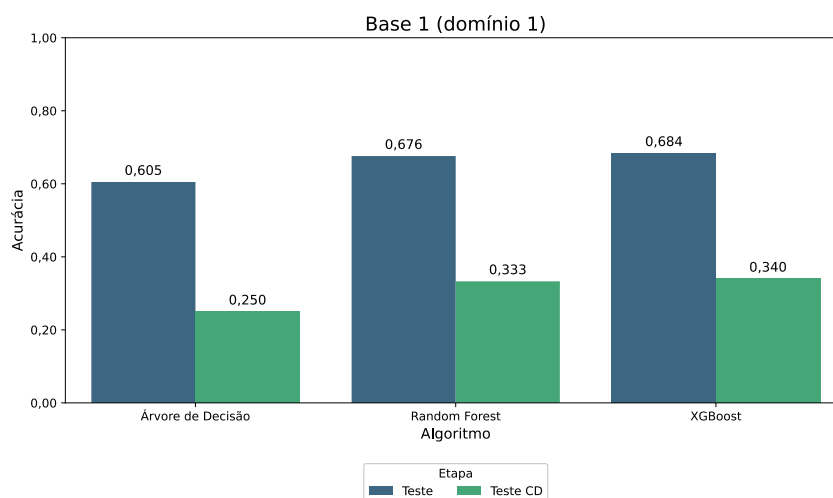


Figura 1. Acurácia dos modelos de Base 1
Fonte: Resultados originais da pesquisa

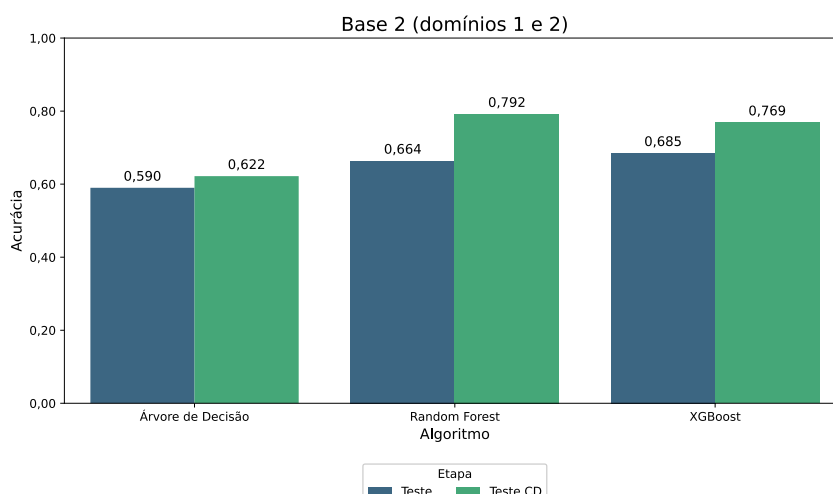


Figura 2. Acurácia dos modelos de Base 2
Fonte: Resultados originais da pesquisa

A faixa de acurácia entre 60% e 70% não é problemática, considerando que a determinação das 13 classes envolve múltiplas decisões estratégicas, permitindo a alocação em classes adjacentes e mantendo, assim, índices de competitividade com os concorrentes de mercado. A Figura 3, a seguir, apresenta a matriz de confusão como exemplo para o modelo “XGBoost” da Base 1, ilustrando a leve dispersão em torno dos valores de teste.

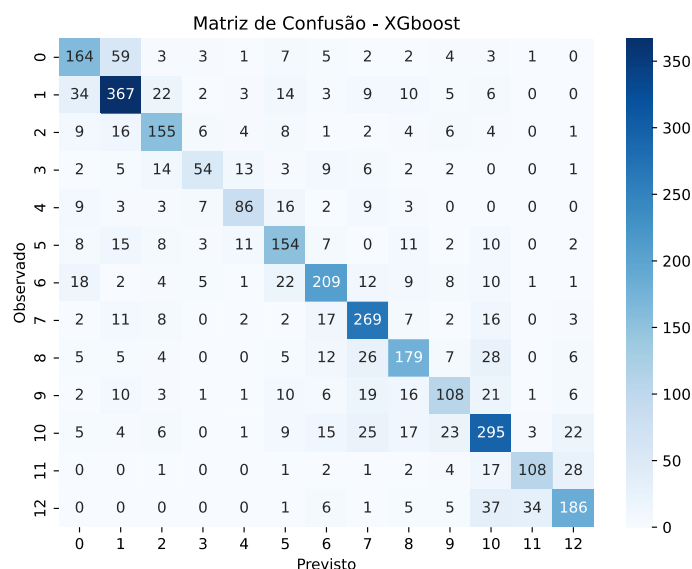


Figura 3. Matriz de Confusão em mapa de calor

Fonte: base de dados originais sob elaboração própria

Nota: intensidade de cores de 0 a 350

Os demais parâmetros de avaliação: Precisão, Sensibilidade e “F1-Score” foram ponderados para conseguir apresentar com facilidade a comparação entre os modelos dentro das 13 classes disponíveis, conforme apresentado nas Tabelas 5 e 6.

Na Base 1 é possível ver o melhor desempenho de “XGBoost” entre os modelos nas três métricas, o que indica maior eficácia na identificação correta das classes e na consistência dos resultados, seguido do modelo “Random Forest” e da Árvore de Decisão. Na Base 2, o modelo “XGBoost” também lidera a avaliação, porém, com uma queda geral nos valores das métricas, o que indica que a combinação dos domínios aumenta a variabilidade e os desafios de generalização². Essa alteração para menos é vista em todos os modelos,

² Esse fenômeno de adaptação dos modelos ao contexto de múltiplos domínios é discutido por Ganin et al. (2016), que exploram como abordagens de aprendizado multi-domínio, mesmo que aplicadas originalmente a redes neurais, podem se relacionar a modelos de “ensemble” ao permitir uma maior generalização ao custo de ajustes específicos.

supondo que os dados capturados em dois domínios distintos apresentam uma complexidade a mais para o treinamento, interferindo na apuração das métricas e no resultado.

Ademais, o modelo de referência da Árvore de Decisão apresentou maiores dificuldades entre os modelos ao lidar com a variabilidade das duas bases de dados, evidenciando suas limitações em capturar padrões complexos e adaptar-se às especificidades de diferentes domínios. Essa característica reduz sua capacidade de generalização quando comparado a modelos “ensemble”, que conseguem lidar melhor com essa variabilidade adicional (Quinlan, 1986; Breiman, 2001).

Tabela 5. Parâmetros de Avaliação - Base 1 (domínio 1)

Modelo	Precisão	Sensibilidade	F1-Score
Árvore de Decisão	0,593	0,593	0,593
"Random Forest"	0,673	0,669	0,666
"XGBoost"	0,682	0,677	0,678

Fonte: resultados originais da pesquisa

Tabela 6. Parâmetros de Avaliação - Base 2 (domínios 1 e 2)

Modelo	Precisão	Sensibilidade	F1-Score
Árvore de Decisão	0,587	0,583	0,584
"Random Forest"	0,657	0,634	0,642
"XGBoost"	0,674	0,643	0,653

Fonte: resultados originais da pesquisa

Conforme mencionado anteriormente, não são considerados, de maneira generalista, erros de classificação uma variação entre classes adjacentes, ou até mesmo dentro de uma escala maior; o negócio utiliza uma divisão em que se aplica um “mix” dentro dessa classificação: “mix high”, “mix mid” e “mix low”; essa divisão é apresentada na Tabela 7. Os agrupamentos atendem a certas estratégias de aplicação de classes de tarifa e ajudam a dar um contorno mais acessível ao entendimento da assertividade nas escolhas de classe pelo analista. Assim sendo, quando foi avaliada a acurácia por agrupamentos foi observado um índice maior em todo grupo conforme as Tabelas 8 e 9.

Tabela 7. Grupos de "mix"

(continua)	
Classe	Classificação
0	"Mix High"
1	
2	
3	

Tabela 7. Grupos de "mix" (conclusão)

4	
5	
6	"Mix Mid"
7	
8	
9	
10	
11	"Mix Low"
12	

Fonte: Base de dados originais

Tabela 8. Acurácia agrupada - Base 1 (domínio 1)

Modelo	Grupos de "mix"		
	"High"	"Mid"	"Low"
Árvore de Decisão	0,898	0,826	0,901
"Random Forest"	0,915	0,853	0,919
"XGBoost"	0,914	0,858	0,925

Fonte: resultados originais da pesquisa

Tabela 9. Acurácia agrupada - Base 2 (domínios 1 e 2)

Modelo	Grupos de "mix"		
	"High"	"Mid"	"Low"
Árvore de Decisão	0,922	0,834	0,903
"Random Forest"	0,939	0,882	0,938
"XGBoost"	0,934	0,889	0,951

Fonte: resultados originais da pesquisa

A inclusão do segundo domínio na Base 2 aumentou a variabilidade dos dados, levando a uma maior especialização dos modelos em captar padrões mais amplos de classificação, como mostrado na Tabela 9. Em outras palavras, o modelo treinado na Base 2 consegue apresentar uma performance ligeiramente mais satisfatória na acurácia por agrupamento. Isso contrasta com os menores índices de precisão, sensibilidade e "F1-score", que refletem uma redução na performance global devido à complexidade adicional introduzida pela variabilidade entre os dois domínios.

Portanto, a acurácia dentre os grupos selecionados de classes de tarifa indica uma "acurácia ajustada" que melhor ilustra o desempenho dos modelos de maneira típica ao negócio por essa perspectiva em especial.

Outro fator importante para o negócio é avaliar se a variável de competitividade (CMT) sofreu alterações significativas, dado que o objetivo da gestão de demanda de curto prazo é alinhar os voos com as variáveis internas e com a concorrência de maneira paritária. Considera-se competitivo o voo cuja classe aplicada seja igual ou inferior à da concorrência.

Nesse aspecto, os modelos “XGBoost” e “Random Forest” tiveram um desempenho similar ao índice original das bases testadas, conforme apresentado na Tabela 10. Já a Árvore de Decisão apresentou um desempenho abaixo do esperado no cenário de maior complexidade dos dados (Breiman, 2001), o que impacta os resultados em um cenário real.

Tabela 10. Índice de Competitividade - Geral

Referência	Índice original	Modelos		
		“XGBoost”	“Random Forest”	Árvore de Decisão
Base 1 (domínio 1)	82,0%	82,3%	82,2%	61,2%
Base 2 (domínios 1 e 2)	79,3%	79,6%	79,6%	55,6%

Fonte: resultados originais da pesquisa

Conclusão

Este estudo mostrou que os modelos “ensemble”, em especial o “XGBoost”, apresentaram desempenho promissor em vários cenários. No entanto, como é comum em análises de “machine learning”, há potencial para melhorias adicionais por meio de ajustes de hiperparâmetros, juntamente com uma ampliação da base de dados, sobretudo no modelo “Random Forest”, visando alcançar um desempenho ainda mais próximo ao do “XGBoost”. A Árvore de Decisão, enquanto modelo simplificado, não atendeu aos critérios esperados, especialmente quanto ao índice de competitividade, destacando-se como menos adequada para a complexidade da base de dados utilizada.

Os modelos apresentaram uma tendência ao “overfitting” quando foram treinados em um único domínio (Base 1), o que foi mitigado pela combinação de um segundo domínio ao primeiro nos testes (Base 2). Esse cenário destaca uma escolha importante para o negócio: optar por um modelo mais generalista, que sacrifica um pouco de precisão e sensibilidade para atender a diferentes mercados, ou eleger modelos especializados, adaptados às características exclusivas de certos segmentos, como elasticidade de demanda e sazonalidade.

A análise de acurácia por grupos revelou que os modelos foram eficazes ao classificar entre baixas, médias e altas classes tarifárias. Isso sugere que, em uma situação real, a utilização de classes adjacentes para alcançar objetivos semelhantes é viável, minimizando a percepção de “erros adjacentes” e favorecendo a flexibilidade nas decisões.

Assim, os modelos “ensemble” demonstram ser ferramentas úteis para automatizar a gestão de demanda em companhias aéreas, com uma leve vantagem do “XGBoost” sobre o “Random Forest” nos cenários testados.

Referências

ANDERSON, Chris K.; BELL, David B.; KAISER, Kurt. Revenue management: An overview and research agenda. *Production and Operations Management*, v. 12, n. 4, p. 423-431, 2003.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.

Breiman, L., (1996). Bagging predictors. *Machine Learning* 26 (2), 123–140.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>

Dietterich, T.G., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine Learning* 40 (2), 139–157.

Fávero LP, Belfiore P. 2024. *Manual de Análise de Dados: Estatística e Machine Learning com Excel, SPSS, STATA, R e Python*. 2a ed. Rio de Janeiro: LTC Livros Técnicos e Científicos Ltda.

Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119-139.

Ganin, Y. et al. (2016). Domain-Adversarial Training of Neural Networks. *The Journal of Machine Learning Research*.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Edition. New York: Springer.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. New York: Springer.

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. New York: Springer.

McAfee, R. P., & te Velde, V. (2007). Dynamic Pricing in the Airline Industry. *Handbook on Economics and Information Systems*, 1, 317–358.

Philips LR. *Pricing and Revenue Optimization*. Stanford, California: Stanford University Press, 2011.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.

Sartoris Neto, A. *Estatística e Introdução à Econometria*. 2ª ed. São Paulo: Saraiva, 2015.

Schapire, R. E. (1990). The Strength of Weak Learnability. *Machine Learning*, 5(2), 197-227.

Talluri, K. T., & Van Ryzin, G. J. (2004). "The Theory and Practice of Revenue Management."

Tzeng, E. et al. (2015). Simultaneous Deep Transfer Across Domains and Tasks. Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV).