

INSTITUTO TECNOLÓGICO DE AERONÁUTICA



Alexandre de Castro Fernandes

**ANÁLISE E PREVISÃO DE INDICADORES DE
PERFORMANCE DO TRÁFEGO AÉREO BRASILEIRO**

Trabalho de Graduação
2022

Curso de Engenharia Civil-Aeronáutica

Alexandre de Castro Fernandes

**ANÁLISE E PREVISÃO DE INDICADORES DE
PERFORMANCE DO TRÁFEGO AÉREO BRASILEIRO**

Orientador

Prof. Dr. Marcelo Xavier Guterres (ITA)

ENGENHARIA CIVIL-AERONÁUTICA

**SÃO JOSÉ DOS CAMPOS
INSTITUTO TECNOLÓGICO DE AERONÁUTICA**

Dados Internacionais de Catalogação-na-Publicação (CIP)
Divisão de Informação e Documentação

Fernandes, Alexandre de Castro
Análise e previsão de indicadores de performance do tráfego aéreo brasileiro / Alexandre de Castro Fernandes.
São José dos Campos, 2022.
66f.

Trabalho de Graduação – Curso de Engenharia Civil-Aeronáutica– Instituto Tecnológico de Aeronáutica, 2022. Orientador: Prof. Dr. Marcelo Xavier Guterres.

1. Aeroporto. 2. Desempenho. 3. Transporte de passageiros. 4. Transporte aéreo.
5. Transportes. I. Instituto Tecnológico de Aeronáutica. II. Título.

REFERÊNCIA BIBLIOGRÁFICA

FERNANDES, Alexandre de Castro. **Análise e previsão de indicadores de performance do tráfego aéreo brasileiro**. 2022. 66f. Trabalho de Conclusão de Curso (Graduação) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSÃO DE DIREITOS

NOME DO AUTOR: Alexandre de Castro Fernandes

TÍTULO DO TRABALHO: Análise e previsão de indicadores de performance do tráfego aéreo brasileiro.

TIPO DO TRABALHO/ANO: Trabalho de Conclusão de Curso (Graduação) / 2022

É concedida ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias deste trabalho de graduação e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho de graduação pode ser reproduzida sem a autorização do autor.

Alexandre de Castro Fernandes

Alexandre de Castro Fernandes
Rua H8A, 132
12.228-460 – São José dos Campos–SP

ANÁLISE E PREVISÃO DE INDICADORES DE PERFORMANCE DO TRÁFEGO AÉREO BRASILEIRO

Essa publicação foi aceita como Relatório Final de Trabalho de Graduação

Alexandre de Castro Fernandes

Alexandre de Castro Fernandes

Autor

Marcelo Xavier Guterres

Marcelo Xavier Guterres (ITA)

Orientador

Evandro José da Silva

Prof. Dr. Evandro José da Silva
Coordenador do Curso de Engenharia Civil-Aeronáutica

São José dos Campos, 17 de novembro de 2022.

Aos meus pais por me apoiarem durante toda a jornada escolar e minha namorada que foi fundamental por estar comigo e me incentivar do início ao fim da construção desse trabalho.

Agradecimentos

Primeiramente, gostaria de agradecer à Deus por me dar saúde, bênçãos e sabedoria durante toda a jornada acadêmica.

Aos meus pais, Marcilea e Gilson, por desde o início dos meus estudos me incentivarem a sempre buscar o aprendizado. Sem vocês eu não teria chegado a esse momento.

À minha namorada Carol, por me apoiar e fornecer suporte durante toda a confecção desse trabalho e sempre acreditar em mim.

Ao Prof. Marcelo Xavier Guterres, pela orientação e apoio durante a realização deste trabalho.

Aos meus colegas de turma, por durante toda a jornada universitária me apoiarem nos momentos de insegurança.

*“Quando tudo parecer dar errado na sua vida,
lembre-se que o avião decola contra o vento, e não a favor dele.”*

— HENRY FORD

Resumo

O transporte aéreo brasileiro tem apresentado crescimento em sua demanda de forma que a medição do desempenho da operação de gestores de aeroportos e companhias aéreas cumpre papel fundamental na manutenção da qualidade do serviço prestado aos passageiros. O presente estudo busca analisar indicadores de *performance* do setor aeroportuário de forma a obter os principais fatores que influenciam na eficiência da operação. A pesquisa é feita levando-se em consideração os voos com decolagem no território brasileiro e insere-se no estudo informações detalhadas do voo como condições meteorológicas do local no momento da partida e dados sobre demanda e capacidade da respectiva aeronave e aeroporto. Com essas informações primeiramente visualizou-se o indicador de pontualidade de partida (KPI 01) levando-se em consideração quatro categorias na classificação. A partir disso, utilizou-se o método de Aumento de Gradiente Extremo (XGBOOST) como algoritmo de aprendizado de máquina supervisionado com o objetivo de prever a pontualidade de partida de um voo. Os resultados obtidos mostram que a pressão local, a quantidade de decolagens na mesma hora que a do respectivo voo e o volume de bagagem franqueada são os principais fatores que influenciam na pontualidade de partida. Por fim, analisou-se a eficiência na gestão da demanda e capacidade horárias dos aeroportos.

Abstract

Brazil's air transportation has shown growth in its demand. Hence, the measurement of the operation performance of airport managers and airlines plays a crucial role in maintaining the quality of service provided to passengers. This study seeks to analyze performance indicators of the airport sector to obtain the main factors that influence the efficiency of the operation. The research takes into account the flights with takeoff in The Brazilian territory and inserts in the study detailed information of the flight as weather conditions of the site at the time of departure and data on demand and capacity of the respective aircraft and airport. With this information, the starting punctuality indicator (KPI 01) was first visualized, taking into account four categories in the classification. From this, the Extreme Gradient Boosting (xgboost) method was used as a supervised machine-learning algorithm to predict the punctuality of a flight's departure. The results obtained show that the local pressure, the number of take-offs at the same time as the respective flight, and the volume of franchised baggage are the main factors that influence the punctuality of departure. Finally, the efficiency in the management of the hourly demand and capacity of airports was analyzed.

Lista de Figuras

FIGURA 3.1 – Distribuição geográfica dos aeroportos em estudo	24
FIGURA 3.2 – Evolução dos modelos de predição usando árvore de decisão	28
FIGURA 3.3 – Matriz de Confusão	31
FIGURA 4.1 – Número de voos que decolaram por aeroporto	33
FIGURA 4.2 – Número de voos que decolaram por companhia aérea	34
FIGURA 4.3 – Pontualidade de partida dos voos por aeroporto	34
FIGURA 4.4 – Pontualidade de partida dos voos por companhia aérea	35
FIGURA 4.5 – Matriz de confusão (Amostra desbalanceada)	36
FIGURA 4.6 – Matriz de confusão (Amostra balanceada)	37
FIGURA 4.7 – Importância das variáveis de entrada	37
FIGURA 4.8 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão	38
FIGURA 4.9 – Influência do números de voos na hora na pontualidade de partida e na acurácia da previsão	39
FIGURA 4.10 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão	40
FIGURA 4.11 – Matriz de confusão - Aeroporto de Guarulhos	41
FIGURA 4.12 – Importância das variáveis de entrada - Guarulhos	42
FIGURA 4.13 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Guarulhos	42
FIGURA 4.14 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Guarulhos	43
FIGURA 4.15 – Influência do número de voos na hora na pontualidade de partida e na acurácia da previsão - Guarulhos	44

FIGURA 4.16 –Matriz de confusão - Aeroporto de Viracopos	45
FIGURA 4.17 –Importância das variáveis de entrada - Viracopos	45
FIGURA 4.18 –Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Viracopos	46
FIGURA 4.19 –Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Viracopos	46
FIGURA 4.20 –Influência do número de voos na hora na pontualidade de partida e na acurácia da previsão - Viracopos	47
FIGURA 4.21 –Matriz de confusão - Aeroporto de Confins	48
FIGURA 4.22 –Importância das variáveis de entrada - Confins	48
FIGURA 4.23 –Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Confins	49
FIGURA 4.24 –Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - Confins	49
FIGURA 4.25 –Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Confins	50
FIGURA 4.26 –Matriz de confusão - GOL	51
FIGURA 4.27 –Importância das variáveis de entrada - GOL	51
FIGURA 4.28 –Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - GOL	52
FIGURA 4.29 –Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - GOL	52
FIGURA 4.30 –Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - GOL	53
FIGURA 4.31 –Matriz de confusão - TAM	53
FIGURA 4.32 –Importância das variáveis de entrada - TAM	54
FIGURA 4.33 –Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - TAM	54
FIGURA 4.34 –Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - TAM	55
FIGURA 4.35 –Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - TAM	55

FIGURA 4.36 –Matriz de confusão - AZUL	56
FIGURA 4.37 –Importância das variáveis de entrada - AZUL	57
FIGURA 4.38 –Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - AZUL	57
FIGURA 4.39 –Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - AZUL	58
FIGURA 4.40 –Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - AZUL	58
FIGURA 4.41 –Capacidade horária de pousos por aeroporto	59
FIGURA 4.42 –Percentual de dias com demanda acima de 90% da capacidade de pousos	60
FIGURA 4.43 –Percentual de meses com demanda acima de 90% da capacidade de pousos	61
FIGURA 4.44 –95 ^o percentil do movimento de pousos por aeroporto	62
FIGURA 4.45 –Taxa pico e capacidade de pousos por aeroporto	62

Lista de Abreviaturas e Siglas

KPI	Key Performance Indicator
DECEA	Departamento de Controle do Espaço Aéreo
XGBOOST	Aumento de gradiente extremo
ANAC	Agência Nacional de Aviação Civil
REDEMET	Rede de Meteorologia do Comando da Aeronáutica
VRA	Voo Regular Ativo
CGNA	Centro de Gerenciamento da Navegação Aérea
SMOTE	Técnica de Sobreamostragem Minoritária Sintética

Sumário

1	INTRODUÇÃO	15
1.1	Pergunta de Pesquisa	16
1.2	Objetivo Geral	16
1.2.1	Objetivos Específicos	16
1.3	Justificativa	16
1.4	Limitação e condição de contorno	17
1.5	Método de pesquisa	17
1.6	Estrutura do texto	17
2	REVISÃO BIBLIOGRÁFICA	19
3	METODOLOGIA	21
3.1	Descrição das bases de dados	21
3.2	Pré-processamento dos dados	22
3.3	Cálculo dos indicadores de performance	25
3.3.1	KPI 01 - Pontualidade de Partida	25
3.3.2	KPI 09 - Capacidade de chegada no aeroporto	25
3.3.3	KPI 10 - Taxa pico de chegada no aeroporto	26
3.3.4	KPI 11 - Eficiência de processamento no aeroporto	26
3.4	Modelo preditivo de pontualidade	27
4	RESULTADOS E DISCUSSÃO	33
4.1	Análise exploratória da base de dados	33
4.2	Análise de comportamento do modelo preditivo da pontualidade de partida	35

4.3	Análise dos fatores de performance nos aeroportos mais movimentados	40
4.3.1	SBGR - Aeroporto Internacional de Guarulhos	40
4.3.2	SBKP - Aeroporto Internacional de Viracopos	44
4.3.3	SBCF - Aeroporto Internacional de Confins	47
4.4	Análise dos fatores de performance para as companhias aéreas com maior marketshare	50
4.4.1	GOL - LINHAS AÉREAS S.A.	50
4.4.2	TAM - LINHAS AÉREAS S.A.	53
4.4.3	AZUL - LINHAS AÉREAS S.A.	56
4.5	Eficiência de processamento dos aeroportos	58
4.6	Diagnóstico da taxa pico de chegada nos aeroportos	61
5	CONCLUSÃO	64
	REFERÊNCIAS	65

1 Introdução

O setor de transporte aéreo brasileiro apresenta grande relevância na economia do país, sendo responsável por 1,1% (18,8 bilhões de dólares) do PIB (Produto Interno Bruto) e por gerar 839.000 empregos em 2018. (IATA, 2019). Dentro do ecossistema desse setor observam-se importantes componentes, dentre eles os aeroportos e as companhias aéreas.

Por exemplo, o Aeroporto Internacional de Guarulhos, principal polo do transporte aéreo brasileiro, receberá um aporte de 100 milhões de reais com o objetivo de expandir seu terminal de cargas (AEROFLAP, 2022). Esse investimento reflete a modernização do principal sítio aeroportuário e o aumento da oferta de voos. Além disso, o número de passageiros aéreos cresceu 210% entre 2000 e 2014 (NUMERO, 2015) o que demonstra um aumento na demanda.

Diante do crescimento em grande escala desse setor é importante a análise da performance dos aeroportos e companhias aéreas. Essa pode ser estudada desde níveis macro indicadores de receitas e custos totais até nível micro em análises individuais, através da avaliação de indicadores relativos a uma operação. Esses parâmetros analisados são chamados de KPI (Key Performance Indicator), indicadores-chave de *performance* em português.

Criado em meados da década de 1990, o termo KPI foi introduzido com o objetivo de analisar como um todo uma empresa ou setor. Em décadas anteriores, comumente analisava-se somente a produtividade e os KPIs viabilizaram medir o desempenho estratégico de uma empresa e assim otimizar o alcance dos objetivos.

Essa temática já apresenta estudos acerca da obtenção dos valores dos KPIs descritos na MCA 100-22 (BRASIL., 2020), porém ainda tem-se uma carência na obtenção dos fatores determinantes de performance e predição da performance.

No documento da MCA 100-22 desenvolvido pelo Departamento de Controle do Espaço Aéreo (DECEA) tem-se descritos dezenove diferentes KPIs que medem a performance do tráfego aéreo como um todo. Com isso toma-se como fundamentos base, desse estudo, esses indicadores.

1.1 Pergunta de Pesquisa

A partir do que está descrito na seção anterior, visualiza-se que o estudo da performance de aeroportos e companhias aéreas possui grande relevância no serviço fornecido aos passageiros. Com isso surge o questionamento que essa pesquisa acadêmica busca responder:

“É possível obter os principais fatores que influenciam em uma maior performance da operação de aeroportos e companhias aéreas e prevê-los utilizando aprendizado de máquina?”

1.2 Objetivo Geral

Diante do questionamento base da pesquisa, chega-se em um foco central do estudo:

“Desenvolvimento de estudo através da construção de modelos descritivos e preditivos para medição da performance de aeroportos e companhias aéreas e obtenção dos principais fatores relevantes para os indicadores de estudo”

1.2.1 Objetivos Específicos

De forma a alcançar o objetivo descrito, foram selecionados os objetivos específicos:

- 1: *Selecionar e calcular os KPIs para cada aeroporto e companhia aérea;*
- 2: *Identificar o algoritmo de aprendizado de máquina para obter os fatores de maior relevância na performance;*
- 3: *Validar o modelo construído para prever a performance.*

1.3 Justificativa

A presente pesquisa apresenta pertinência dado o impacto sobre as diferentes áreas do setor de transporte aéreo: (i) sucesso do cliente, (ii) financeiro, (iii) mercado

(i) sucesso do cliente: a partir da obtenção dos principais fatores que interferem na performance do setor, tem-se maior visibilidade sobre as áreas mais críticas que necessitam de maior investimento e atenção para fornecer um serviço de maior qualidade. Com isso, a métrica de NPS (Net Promotion Score), por exemplo, sofrerá interferência direta dos insights resultantes do estudo de maneira que a satisfação dos usuários, com o serviço oferecido, aumente.

(ii) *financeiro*: um aspecto de grande impacto em qualquer operação é a área do financeiro. A partir da predição de fatores que influenciam na performance, o setor financeiro possuirá uma assertividade maior na determinação dos departamentos a receberem maiores investimentos.

(iii) *mercado*: a partir da obtenção de fatores e predição dos mesmos, os aeroportos e companhias aéreas têm maior capacidade de evoluir a qualidade de seus serviços e com isso alargar uma maior quota de mercado, marketshare, o que leva à necessidade de empresas concorrentes intensificarem também melhorias em seus serviços.

1.4 Limitação e condição de contorno

As análises estão diretamente relacionadas a agentes externos à operação de um aeroporto ou companhia aérea, como por exemplo, a pandemia do COVID-19. Dado o período de análise estar correlacionado aos anos mais recentes, tem-se uma influência externa da pandemia sobre o mercado como um todo.

1.5 Método de pesquisa

A classificação da presente pesquisa é baseada nos critérios de Gil (GIL, 2002), os quais são pautados nos objetivos gerais e procedimentos técnicos aplicados.

Classificação em relação aos objetivos gerais: a presente pesquisa apresenta análises descritivas em que, baseada em dados do setor de transporte aéreo, busca calcular os indicadores de performance em diversos níveis. Além disso, têm-se análises utilizando métodos de aprendizado de máquina com o objetivo de prever fatores relevantes de tal forma que a pesquisa possa ser inserida como um estudo exploratório.

Classificação em relação aos procedimentos técnicos: como trata-se de uma pesquisa em que se utiliza bases de dados que podem não ter sido previamente tratadas de forma analítica, pode ser classificada como pesquisa documental.

1.6 Estrutura do texto

A presente monografia está dividida em seis partes:

- 1 - *introdução teórica demonstrando a relevância e os objetivos principais da pesquisa;*
- 2 - *revisão teórica com o objetivo de dissertar sobre os conceitos teóricos aplicados à pesquisa e relacionar com estudos já realizados;*

3 - *descrição da metodologia da pesquisa através da dissertação dos materiais e métodos utilizados;*

4 - *descrição dos resultados obtidos e posterior análise desses de forma a alcançar os objetivos e responder o questionamento da pesquisa;*

5 - *considerações finais e possíveis sugestões de melhoria da pesquisa desenvolvida;*

6 - *referências bibliográficas utilizadas na construção da monografia.*

2 Revisão Bibliográfica

O crescimento do setor aeroportuário no âmbito do número de passageiros no Brasil entre os anos de 2003 e 2014 foi em média de 10,3 por cento ao ano segundo dados do estudo realizado pelo BNDES (YOSIMOTO, 2016). Em 2003 totalizavam 71 milhões de passageiros e em 2014 já eram 209 milhões. Diante disso, a operação das atividades do aeroporto demanda planejamento, organização e controle de forma a fornecer um serviço de qualidade aos passageiros.

Fayol (1916) no estudo denominado 'Administração industrial e geral' descreve as funções da administração denominadas por ele como prever, organizar, ordenar, conectar e controlar. A função de prever, atualizada atualmente com a nomenclatura de planejamento, compreende processos de decisão a respeito de ações a serem implementadas e métodos de implementação no futuro.

A função de organização compreende a constituição de uma dupla: organização material e social de forma a definir os responsáveis pelas atividades estabelecidas no planejamento e os recursos necessários para execução com o objetivo de alcançar um bom desempenho. Pode-se também destacar a função de controle que objetiva garantir que a operação ocorra de acordo com normas internas e legislações previstas além de ordens dadas. O controle apresenta grande relação com os objetivos da empresa de forma que contribui com o acompanhamento e atingimento das metas estipuladas.

De acordo com Tezza (2010) a medição de desempenho em atividades através de metas possuem foco em aspectos financeiros, bem como na qualidade e produtividade. Além disso, essas podem ser divididas em estratégicas com visão a longo prazo, em táticas com objetivo de monitorar processos e pessoas. Por fim as metas operacionais que são medidas de rotina e ligadas a feedback em tempo real.

Na operação de aeroportos estão presentes diversos setores como: controle de tráfego aéreo, atendimento ao cliente no check-in e salas de embarque que necessitam estarem coordenados para obter um bom desempenho e fornecer um serviço de qualidade. De acordo com Neves (2009) a medição de *performance* em nível tático é importante de ser feita usando-se KPIs de forma que serão bastante úteis no setor de *Business Intelligence* (*BI*) da empresa.

Em 2018 Schultz (SCHULTZ, 2018) fez um estudo de forma a relacionar as condições meteorológicas à performance de um aeroporto através do algoritmo *Air Traffic Management Airport Performance (ATMAP)*. Nesse estudo o índice calculado pelo algoritmo é maior para condições mais extremas de meteorologia e teve-se como resultado que o número de voos com atrasos ou cancelamentos aumenta a partir de maiores valores do índice.

No estudo realizado por Bubalo (2020) foi construída uma relação entre a demanda e capacidade do sítio aeroportuário com a pontualidade dos voos. Obteve-se que o planejamento da distribuição dos horários de partida dos voos possui grande importância para obter maior performance na operação do aeroporto.

Sternberg (2017) indica que o atraso em voos possuem importância literária, pois tem-se impactos econômicos e ambientais. Para então entender possíveis causas e importância dos fatores que geram o atraso foram, em primeira análise, descritas as variáveis advindas de diferentes áreas como planejamento da distribuição dos voos, variáveis temporais como hora e dia, condições meteorológicas, variáveis geográficas como cidade e região, fatores operacionais como demanda e capacidade dos aeroportos, características da aeronave como capacidade e modelo além de variáveis de condições operacionais.

A partir da seleção das variáveis Sternberg (2017) definiu diferentes métodos que contribuem para identificação dos principais fatores e além disso possibilitam a previsão da pontualidade dos voos. Dentre os métodos destaca-se algoritmos de Machine Learning, simulações, análises estatísticas e modelos probabilísticos. Com o estudo, destaca-se que entre 1997 e 2017 ocorreu um grande aumento de publicações de estudos relacionados à temática e classificando-as em relação ao método utilizado destaca-se que Machine Learning tem sido o modelo mais utilizado.

A Sociedade Brasileira de Computação (FREITAS, 2019) apresenta o aprendizado de máquina (Machine Learning) como uma sub-área da Inteligência Artificial que tem crescido em número e abrangência nas aplicações. Esse tipo de modelagem apresenta diversas vantagens ao ser aplicada como por exemplo a automatização de processos e reconhecimento de padrões de forma que obtem-se grande economia de tempo, além de possibilitar previsões com alto índice de acurácia e confiabilidade.

Dalmau (2019) utiliza o aprendizado de máquina para aumentar a eficácia da previsão da pontualidade de decolagem. Os algoritmos construídos foram: Árvore de Decisão (Gradient Boosting Decision Trees - GBDT) e Redes Neurais (Artificial Neural Network). Ambos modelos alcançaram o resultado esperado com aumento da acurácia da previsão e além disso obteve as dez variáveis com maior importância na previsão, ou seja, que melhor se ajustam e preveem a pontualidade.

3 Metodologia

3.1 Descrição das bases de dados

Os dados utilizados no estudo advém de três bancos de dados no período entre 06/01/2020. e 04/06/2022, listados a seguir:

- METAR - REDEMET
- VRA - ANAC
- Microdados - ANAC

A primeira base de dados foi obtida através da Rede de Meteorologia do Comando da Aeronáutica (REDEMET, 2022). As informações obtidas são mensagens metrificadas a cada hora que indicam índices de condições meteorológicas no respectivo aeroporto. Essas mensagens consistem do produto METAR do Departamento de Controle do Espaço Aéreo (DECEA).As medidas utilizadas no estudo advindas do METAR são:

- *temp.ar* Temperatura do ar (°C)
- *temp.orvalho* Temperatura de orvalho (°C)
- *vento* Velocidade do vento (nós)
- *visibilidade* Visibilidade (m)
- *pressao* Pressão (hPa)

A segunda base de dados utilizada é a de Voos Regulares Ativos (VRA) que foi obtida através da Agência Nacional de Aviação Civil (ANAC, 2022). As informações utilizadas no estudo presentes na base VRA são:

- *num.assentos* Número de assentos
- *partida.prevista* Partida Prevista

- *partida.real* Partida Real
- *chegada.prevista* Chegada Prevista
- *chegada.real* Chegada Real
- *aeroporto.origem* Aeroporto de Origem
- *aeroporto.destino* Aeroporto de Destino

A terceira base de dados empregada é a de Microdados administrada pelo Governo Federal. Nessa base, dentre as informações destaca-se:

- *companhia.aerea* Companhia Aérea
- *num.passag.pagos* Número de passageiros pagos
- *num.passag.gratis* Número de passageiros grátis
- *kg.bagagem.livre* Quantidade de bagagem livre
- *kg.bagagem.excesso* Quantidade de bagagem excesso
- *kg.carga.paga* Quantidade de carga paga
- *kg.carga.gratis* Quantidade de carga grátis
- *aeroporto.origem* Aeroporto de Origem
- *aeroporto.destino* Aeroporto de Destino

3.2 Pré-processamento dos dados

De forma a estudar de forma conjunta os dados, fez-se primeiramente a união entre as bases considerando a interseção entre a base VRA e a Microdados como sendo o número do voo e o dia em que esse ocorreu. Partindo-se da base obtida conjunta, a junção com a base advinda do METAR foi realizada considerando como interseção o aeroporto e o horário previsto de partida do voo. Assim na base consolidada, as informações meteorológicas consistem na condição climática no momento da decolagem.

O estudo foi realizado considerando-se 24 aeroportos localizados no território brasileiro, sendo eles:

- *SBAR* Aeroporto Internacional de Aracaju

- *SBBE* Aeroporto Internacional de Belém
- *SBCF* Aeroporto Internacional de Confins
- *SBCG* Aeroporto Internacional de Campo Grande
- *SBCT* Aeroporto Internacional Afonso Pena
- *SBCY* Aeroporto Internacional de Cuiabá
- *SBEG* Aeroporto Internacional de Manaus
- *SBFI* Aeroporto Internacional de Foz do Iguaçu
- *SBFL* Aeroporto Internacional de Florianópolis
- *SBFZ* Aeroporto Internacional de Fortaleza
- *SBGL* Aeroporto Internacional Tom Jobim
- *SBGO* Aeroporto Internacional de Goiânia
- *SBGR* Aeroporto Internacional de São Paulo/Guarulhos
- *SBJP* Aeroporto Internacional de João Pessoa
- *SBKP* Aeroporto Internacional de Viracopos
- *SBMO* Aeroporto Internacional de Maceió
- *SBNF* Aeroporto Internacional de Navegantes
- *SBPA* Aeroporto Internacional de Porto Alegre
- *SBPS* Aeroporto Internacional de Porto Seguro
- *SBSG* Aeroporto Internacional de Natal
- *SBSL* Aeroporto Internacional de São Luís
- *Sbsp* Aeroporto de São Paulo/Congonhas
- *SBSV* Aeroporto Internacional de Salvador
- *SBVT* Aeroporto Internacional de Vitória



FIGURA 3.1 – Distribuição geográfica dos aeroportos em estudo

De forma a analisar um grande espectro de fatores calcula-se métricas relativas ao movimento nos aeroportos no momento do respectivo voo. Assim acrescentou-se duas variáveis:

- *num.voos* Número de voos programados na hora do voo no aeroporto
- *maximo.voos* Máximo de voos que o aeroporto já operou em uma hora

Acrescenta-se também fatores relativos ao tempo de forma a entender se o momento da decolagem do voo apresenta impacto sobre a eficiência do aeroporto e companhias aéreas. Os fatores inseridos relativos ao tempo são:

- *hora.do.dia* Horário programado para partida do voo (número inteiro entre 00 e 23)
- *periodo.dia* Período do dia (0 - Madrugada (00 às 05h) / 1 - Manhã (06 às 11h) / 2 - Tarde (12 às 17h) / 3 - Noite (18 às 23h))
- *dia.semana* Dia da semana (0 - Segunda-feira / 1 - Terça-feira / 2 - Quarta-feira / 3 - Quinta-feira / 4 - Sexta-feira / 5 - Sábado / 6 - Domingo)
- *fim.semana* Dia da semana / Fim de semana (0 - Dia da semana / 1 - Fim de semana)

3.3 Cálculo dos indicadores de performance

Esse estudo de performance dos aeroportos e companhias aéreas baseia-se nos KPIs (Key Performance Indicator) presentes na MCA 100-22. (BRASIL., 2020) e o cálculo de cada indicador é descrito nesse referido documento.

3.3.1 KPI 01 - Pontualidade de Partida

Inicia-se através da obtenção do KPI - 01 que corresponde à pontualidade de partida dos voos e é descrito como sendo a porcentagem de voos que saem do *gate* no horário programado. De acordo com Resolução da ANAC (Agência Nacional de Aviação Civil) um voo é considerado como atrasado, quando a diferença entre sua decolagem real e a partida prevista supera 15 minutos. Assim, a partir das informações de dia e horário de decolagem prevista e real pode-se calcular a diferença abaixo:

$$\text{delta.partida} = \text{partida.real} - \text{partida.prevista} \quad (3.1)$$

Portanto, no caso de *delta.partida* ser superior à 900 segundos, considera-se que o voo atrasou. O cálculo do indicador pode ser feito através da divisão entre o número total de voos pontuais sobre o montante total de voos no período analisado:

$$KPI_{01} = \frac{\sum(\text{voos que decolaram pontualmente})}{\sum(\text{voos que decolaram})} \times 100\% [\% \text{ de voos}] \quad (3.2)$$

3.3.2 KPI 09 - Capacidade de chegada no aeroporto

De forma a otimizar os estudos relacionados às possíveis expansões e investimentos nos sítios aeroportuários, a análise da capacidade de pousos é um importante fator e está descrito no KPI - 09. Para cada aeroporto tem-se um valor declarado obtido pelo aeroporto através de uma padronização dada pela CGNA (Centro de Gerenciamento da Navegação Aérea) e equivale a capacidade de pousos em um aeroporto em uma hora.

Esse cálculo é realizado levando-se em consideração fatores de planejamento e relativos à operações de pouso.

Discretizando-se os fatores de planejamento tem-se:

- condições ideais de coordenação do tráfego aéreo
- equipes operacionais consideradas com mesma capacitação

Em relação aos fatores operacionais, leva-se em consideração:

- tempos médios de ocupação de pista
- mix de aeronaves
- percentual de utilização das cabeceiras
- comprimento do segmento de aproximação final
- separação mínima regulamentar de aeronaves
- configuração das pistas de pouso e de táxi
- velocidade de aproximação final
- utilização ou não de equipamento de radar

No capítulo 4 está descrito por aeroporto presente no estudo sua capacidade horária de pousos.

3.3.3 KPI 10 - Taxa pico de chegada no aeroporto

O KPI - 10 corresponde a uma métrica de análise de demanda do aeroporto e leva em consideração o volume de pousos no aeroporto em cada hora de sua operação.

Com o histórico horário obtido, classifica-se em ordem crescente pelo número de pousos e o indicador é obtido como sendo o 95^o percentil dessa ordenação, ou seja, representa o número tal que em somente 5% do tempo há demanda superior a ele.

Para cada aeroporto do estudo, calculou-se esse indicador levando-se em consideração todo o período de análise e apresenta-se os resultados no capítulo 4.

3.3.4 KPI 11 - Eficiência de processamento no aeroporto

O processamento da operação de pousos no aeroporto corresponde em visão macro a uma relação entre demanda e capacidade e fornece indicadores que possibilitam analisar o quanto esse aeroporto atinge sua capacidade.

O cálculo é feito usando-se os resultados do KPI - 09 que correspondem à capacidade de pousos do sítio aeroportuário e a demanda horária de pousos.

Com isso faz-se a divisão entre a demanda e a capacidade pela equação:

$$KPI_{11} = \frac{(\textit{numero de pousos na hora})}{(\textit{capacidade horaria de pousos})} \times 100\% \quad (3.3)$$

3.4 Modelo preditivo de pontualidade

A partir da classificação de cada voo entre pontual e não pontual usando o critério presente em 3.3.1 adicionou-se na base de dados esse resultado com o objetivo de elaborar um modelo preditivo usando aprendizado de máquina.

O algoritmo desenvolvido trata-se de um método de aprendizado de máquina supervisionado em que utiliza-se variáveis de entrada e um resultado de saída previamente conhecido com o objetivo de conectá-los e a partir disso, apresentar a capacidade de prever essa saída a partir de novos conjuntos de dados.

Nesse estudo utilizou-se a estrutura do método XGBOOST (Aumento de gradiente extremo) (SILVA, 2020), o qual é baseado em árvores de decisão. A escolha desse método advém majoritariamente devido ao seu elevado desempenho em contextos semelhantes e além disso alta velocidade de treinamento e fácil interpretabilidade.

Os algoritmos baseados em árvore de decisão (METODOS, 2020) são amplamente utilizados para dados tabulados e receberam evoluções de suas formas mais simples até atingir o modelo XGBOOST (GOMES, 2019).

Primeiramente, a árvore de decisão é um modelo que tem por objetivo prever o valor de uma variável, que nesse estudo é a pontualidade de partida do voo, através do aprendizado feito por decisões, ou seja, condições impostas às variáveis (uma variável por decisão) que dividem a base de dados.

A expansão desse modelo é o *Bagging* que pode ser aplicado nos âmbitos de regressão ou classificação. No presente estudo usa-se XGBOOST de classificação e nesse caso o *bagging* consiste em uma simulação de várias árvores de decisão e a classe prevista é a que obtiver maioria dentre todas as árvores. Isso faz com que tenha-se menor variabilidade na predição.

Com a evolução seguinte chega-se nas Florestas Aleatórias (*RandomForest*) que utilizam do método *Bagging* ao realizar a simulação de várias árvore de decisão, porém apresenta como principal vantagem a seleção de somente parte das variáveis de entrada para determinação da classe prevista. Dessa forma, mesmo em casos em que uma das variáveis é muito influente sobre o resultado, o modelo terá alta acurácia, dado que essa variável não será selecionada em todas as árvores e assim não haverá limitação no *Bagging*.

Como melhoria do método de Florestas Aleatórias tem-se o *Boosting* dado que nas Florestas Aleatórias e nos métodos já citados o resultado de uma árvore de decisão não influencia as demais. No método *Boosting* utiliza-se que as árvores são construídas sequencialmente, ou seja, as informações resultantes de uma árvore influencia a árvore seguinte. Dessa forma faz-se um contínuo aprimoramento do modelo dado que nos casos em que o modelo não soube classificar corretamente faz-se uma nova iteração de forma a

aprimorá-los e alcançar o resultado correto.

Na sequência progride-se para o método *Gradient Boosting* em que faz-se uma melhoria do *Boosting* dado que além de utilizar o modelo anterior como suporte para o atual, ainda objetiva-se que o atual minimize o erro do anterior usando-se uma função de perda. Com isso o resíduo final, ou seja, a diferença entre o valor previsto e o real diminui e o modelo obtém maior acurácia.

Por fim, chega-se ao XGBOOST que realiza a otimização de sistemas e faz aprimoramentos dos algoritmos em relação ao *Gradient Boosting* e por isso apresenta maior acurácia.

Na Figura 3.2 mostra-se visualmente a evolução do algoritmo utilizado para predição através de árvores de decisão até o XGBOOST.

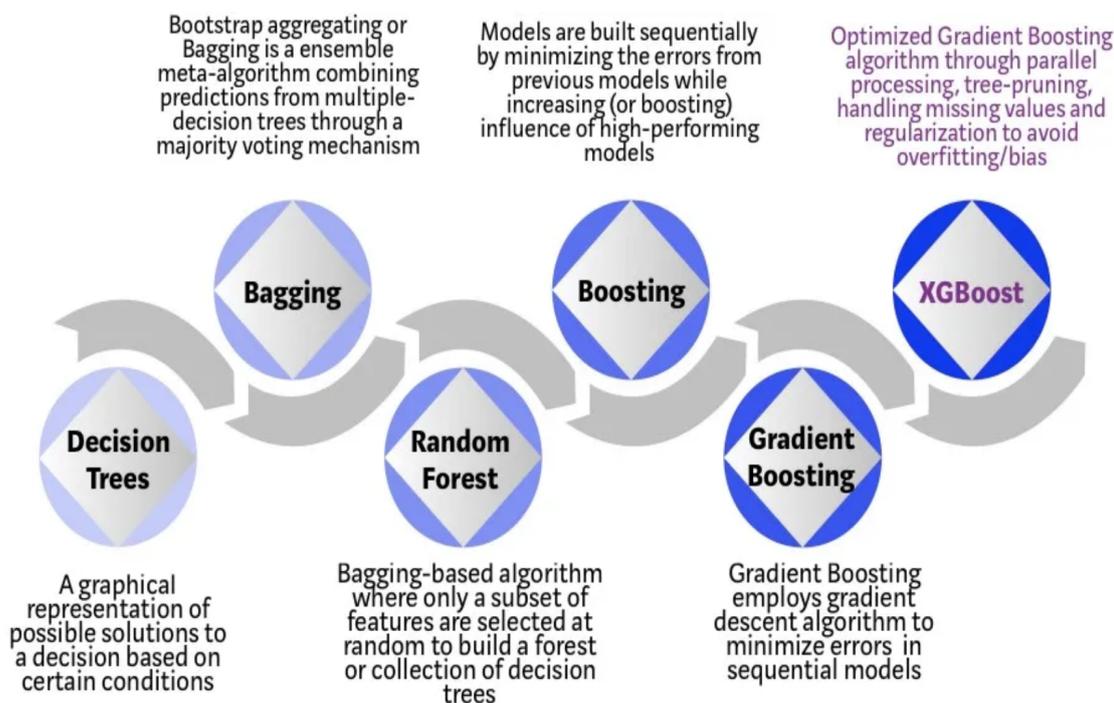


FIGURA 3.2 – Evolução dos modelos de predição usando árvore de decisão

Fonte: Data Geeks

Na biblioteca XGBoost da linguagem de programação Python está presente o pacote *XGBClassifier* o qual é utilizado como algoritmo de predição da pontualidade.

No presente estudo utiliza-se como variáveis de entrada informações meteorológicas locais na hora da decolagem, operacionais de demanda, capacidade da aeronave e aeroporto, além de dados temporais e informações específicas de geolocalização características do voo.

As variáveis de entrada meteorológicas usadas são:

- *temperatura do ar (Nr)* Temperatura do ar
- *temperatura de orvalho (Nr)* Temperatura de orvalho
- *vento (Nr)* Velocidade do vento
- *visibilidade (Nr)* Visibilidade
- *pressao (Nr)* Pressão

As informações de demanda e capacidade da aeronave e aeroporto são:

- *nr – passag – pagos* Número de passageiros pagos
- *kg – bagagem – livre* Quantidade de bagagem livre
- *kg – carga – paga* Quantidade de carga paga
- *N de voos na hora* Número de decolagens na hora do voo
- *% ocupacao da aeronave* Porcentagem de ocupação da aeronave
- *Numero de Assentos* Total de assentos
- *% operacao aeroporto* Porcentagem de operação do aeroporto
- *Maximo de voos aeroporto* Máximo de decolagens em uma hora no aeroporto

Os dados temporais usados são:

- *Dia da semana / Fim de semana* Booleana separando fim de semana e dia de semana
- *Dia da semana* Correspondência numérica referente a cada dia da semana
- *Periodo do dia* Correspondência numérica em relação do período do dia (madrugada, manhã, tarde e noite)
- *Mes* Mês do voo
- *nr – dia – referencia* Dia do mês do voo

As informações de geolocalização características do voo são:

- *Nacional / Internacional* Booleana separando voos com destino nacional e internacional

Iniciou-se o processamento do modelo supervisionado através da seleção aleatória de 70% dos dados para compor a base de treinamento do algoritmo que apresenta a função de criar o modelo e os 30% restantes compõem a base de teste que verifica a qualidade e eficácia do modelo construído.

Após a partição da base configurou-se o modelo através de seus parâmetros (XGBOOST, 2022a):

- *n – estimators* : 650 Quantidade de árvores
- *max – depth* : 10 Tamanho máximo da árvore (profundidade)
- *learning – rate* : 0.01 Taxa de aprendizagem (controle de ponderação de novas árvores)

Com o objetivo de obter alta qualidade do modelo priorizou-se um menor valor de taxa de aprendizagem (0.01) que tem maior probabilidade de gerar um modelo com excelente ajuste aos dados de treino. No caso do tamanho máximo da árvore usa-se uma profundidade intermediária dado que valores superiores a 10 geram alta complexidade do modelo e maior tempo de processamento e no cenário de pequeno valor de profundidade tem-se mais simplicidade que gera menor capacidade de aprendizado e ajuste aos dados. Por fim, de forma a diminuir a variabilidade inerente à modelos baseados em árvore de decisão, usou-se uma quantidade alta de árvores (650) para que a avaliação democrática da predição das classes seja mais eficiente.

A partir do processamento do algoritmo objetiva-se medir a qualidade desse utilizando-se um indicador que seja capaz de diferenciar resultados eficientes e não eficientes. Diante disso, adotou-se a acurácia como a métrica de performance.

A acurácia é calculada comparando as classificações de pontualidade de partida resultantes do modelo e os valores reais presentes na base de dados. Para voos em que o modelo tenha previsto corretamente assinala-se 1 e no caso de erro, marca-se 0.

Assim, a acurácia pode ser obtida fazendo-se a média aritmética entre esse valores assinalados e na sequência transformando o indicador para porcentagem, ou seja:

$$Acuracia = \frac{(\textit{numero de voos previstos corretamente})}{(\textit{total de voos})} X 100\% \quad (3.4)$$

Com a obtenção do indicador de qualidade do modelo, pode-se analisar a distribuição de acertos e erros de forma a entender se existe um enviesamento para uma das classes. Faz-se isso através de uma matriz de confusão (DATA, 2020) conforme a Figura 3.3.

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

FIGURA 3.3 – Matriz de Confusão

- Verdadeiros Positivos: os valores previstos e reais são os mesmos e além disso a estimativa do modelo é positiva.
- Verdadeiros Negativos: os valores previstos e reais são os mesmos e além disso a estimativa do modelo é negativa.
- Falso Positivo: é o erro do tipo I em que os valores previstos e reais não são os mesmos e além disso a estimativa do modelo é positiva.
- Falso Negativo: é o erro do tipo II em que os valores previstos e reais não são os mesmos e além disso a estimativa do modelo é negativa.

A partir disso, a acurácia pode também ser calculada através dos valores presentes na matriz:

$$Acuracia = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\% \quad (3.5)$$

Diante disso, destaca-se dois possíveis cenários para um modelo com alto valor de acurácia.

O primeiro cenário é de ambas as classes possuírem total de acertos em ordem de grandeza parecido ou pelo menos em ordem de grandeza acima da quantidade de erros. Nesse cenário o modelo é validado e pode-se retirar percepções a partir dele.

Em contrapartida, existe o cenário de apesar de se obter alto valor de acurácia (próximo à 100%), a ordem de grandeza do número de acertos de uma das classes é bastante superior aos outros números. Isso mostra que se tem um enviesamento do modelo para prever essa classe e isso pode ser causado por um desbalanceamento entre as classes.

Nessa conjuntura torna-se necessário um ajuste na base de dados de forma a alcançar o primeiro cenário e com isso obter um resultado plausível de ser analisado.

Um procedimento aplicável para contornar esse cenário é fazer uma sobreamostra (*oversample*) da classe com menor frequência de forma a equilibrar as classes em quantidade de dados. Esse aumento no número de registros da menor classe pode ser realizada pelo método SMOTE (Técnica de Sobreamostragem Minoritária Sintética) que sintetiza novas informações com base nos registros já existentes fazendo com que não seja necessário duplicar dados o que poderia enviesar e prejudicar o estudo. A inserção de novas informações é realizada através da seleção aleatória de amostra da classe com menor representatividade e leva em consideração vizinhos dessa amostra. Na sequência a partir de uma interpolação desses vizinhos gera-se uma nova amostra.

A partir da obtenção do indicador de qualidade do modelo e validá-lo, pode-se retirar do método a relevância de cada variável (ABU-RMILEH, 2019) na eficácia de predição do resultado. O objetivo dessa análise é entender quais fatores possuem maior influência no evento de decolagem do voo no que concerne à pontualidade de partida.

No pacote de algoritmos utilizado, *XGBClassifier*, tem-se o método de análise da importância das variáveis de entrada de um modelo de aprendizado de máquina supervisionado chamado de *get - booster().get - score()* (XGBOOST, 2022b) que traz como retorno um número para cada variável de entrada. O tipo de importância das variáveis utilizado é chamado de *weight* (peso) em que contabiliza-se para cada variável o número de vezes em que ela foi utilizada como decisora em uma partição das árvores de decisão.

Assim, variáveis de entrada que tiverem mais aparições considerando-se todas as árvores de decisão geradas pelo algoritmo apresentarão maior importância na predição com eficácia da pontualidade de partida dos voos.

4 Resultados e discussão

4.1 Análise exploratória da base de dados

Inicialmente, foi feita uma análise detalhada sobre os dados depois de processados de forma a fazer comparações entre os aeroportos e companhias aéreas. A métrica de número de voos que decolaram no período analisado é importante de ser visualizada de forma a entender como o volume tráfego aéreo é distribuído espacialmente no território brasileiro. A Figura 4.1 mostra essa distribuição.

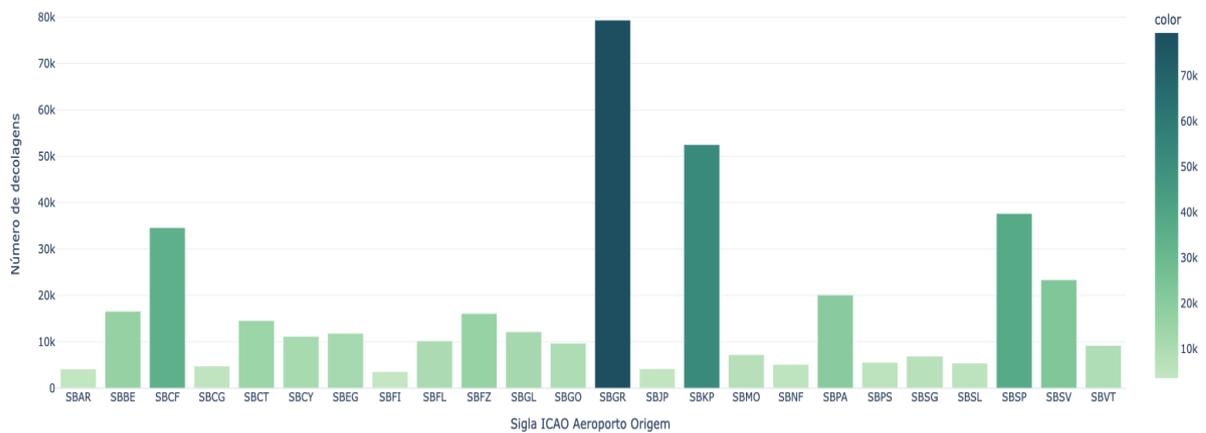


FIGURA 4.1 – Número de voos que decolaram por aeroporto

Em relação às companhias aéreas, visualiza-se na Figura 4.2 que no território nacional apresentam-se três empresas com *marketshare* (porcentagem do mercado) combinado de 97,5%.

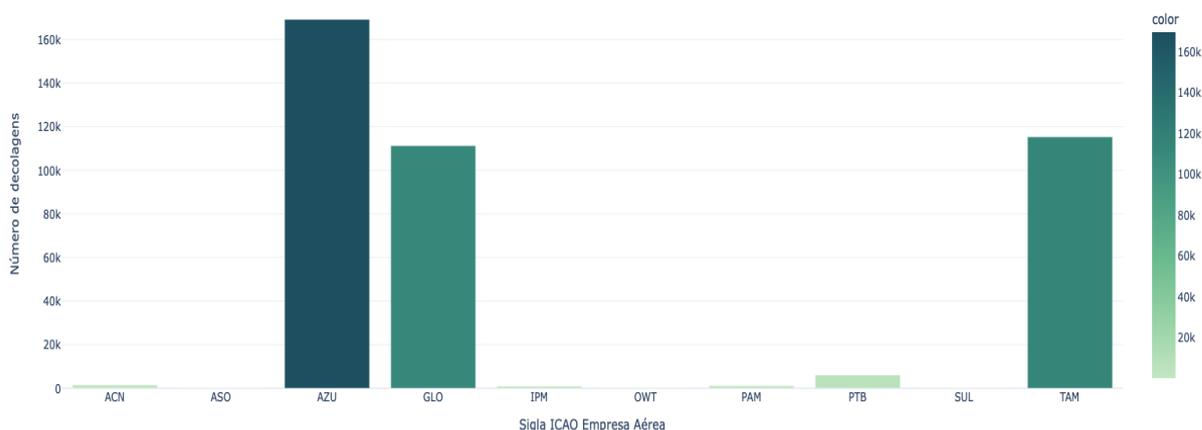


FIGURA 4.2 – Número de voos que decolaram por companhia aérea

A Figura 4.3 mostra a performance de cada aeroporto com relação à pontualidade de partida dos voos. Nessa análise cada voo foi classificado dentre quatro categorias:

- Pontual (atraso até 15 minutos)
- Não pontual (atraso superior a 15 minutos e inferior a 30 minutos)
- Não pontual (atraso superior a 30 minutos e inferior a 60 minutos)
- Não pontual (atraso superior a 60 minutos)

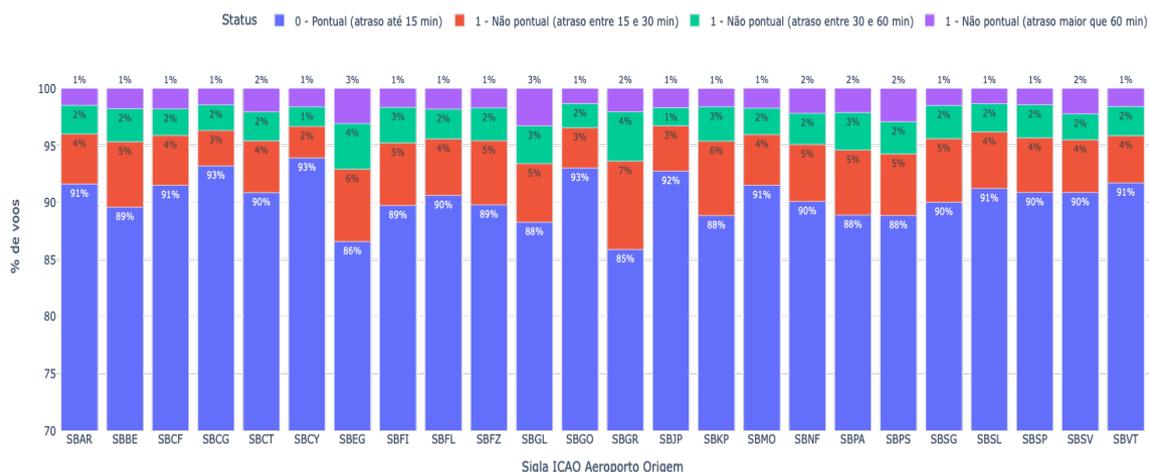


FIGURA 4.3 – Pontualidade de partida dos voos por aeroporto

Os resultados indicam que os principais aeroportos do país possuem pontualidade de partida de seus voos superior a 85%. Porém visualiza-se que o Aeroporto Internacional

de Guarulhos, mais movimentado do país, possui a menor pontualidade, o que ressalta a possibilidade de o alto fluxo de passageiros e a complexidade de operação refletirem em uma menor performance.

Na Figura 4.4 estão presentes as porcentagens de pontualidade de partida referente às companhias aéreas em operação no Brasil. Visualiza-se que as três empresas majoritárias no mercado possuem pontualidade acima de 86%.

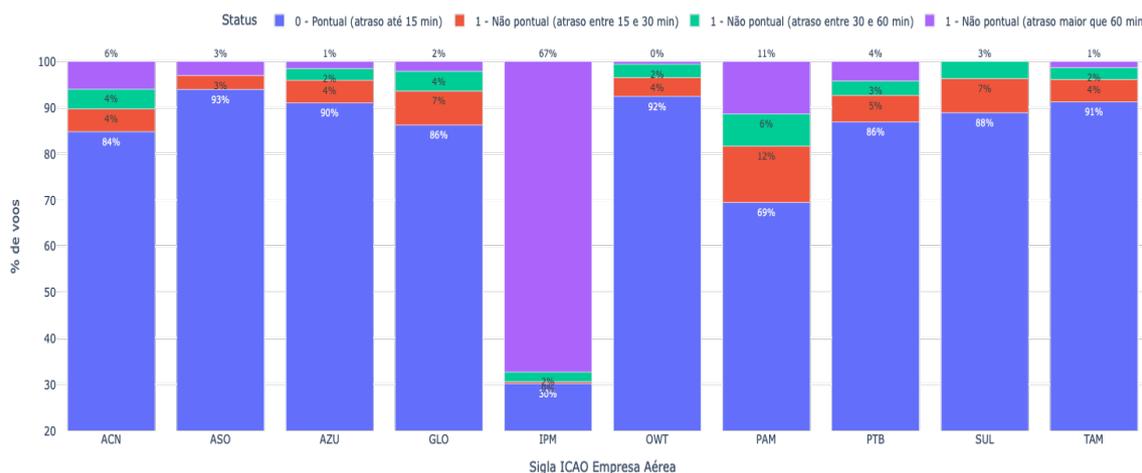


FIGURA 4.4 – Pontualidade de partida dos voos por companhia aérea

4.2 Análise de comportamento do modelo preditivo da pontualidade de partida

Através da aplicação do método de Aumento Extremo de Gradiente (XGBoost) obteve-se um modelo preditivo da pontualidade de partida dos voos. No estudo inicial utilizando-se da base completa de voos obteve-se acurácia de 89,92% e a matriz de confusão presente na Figura 4.5.

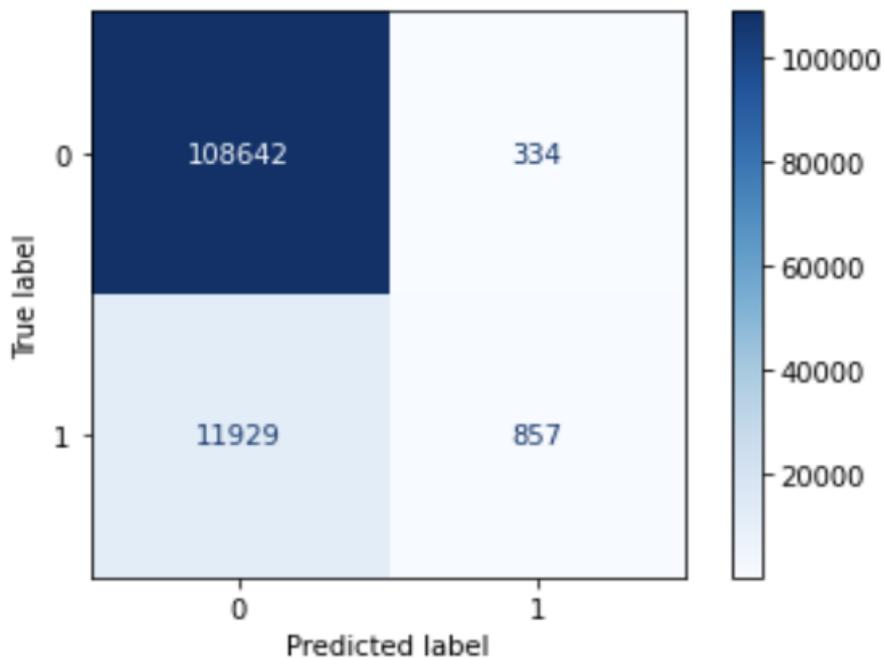


FIGURA 4.5 – Matriz de confusão (Amostra desbalanceada)

Verifica-se que a base de dados está desbalanceada, dado que o volume de voos na base de teste que foram pontuais foi de 108.976 e no caso o de não pontuais foi 12.786. Portanto 89,5% dos voos são pontuais.

Com isso, utilizou-se da técnica de balanceando dos dados por sobreamostra de forma a avaliar em igual peso os voos pontuais e não pontuais. Aplicando-se novamente o algoritmo de classificação XGBOOST com amostragem de 30% para teste obteve-se acurácia de 92,96% e a respectiva matriz de confusão está na Figura 4.6

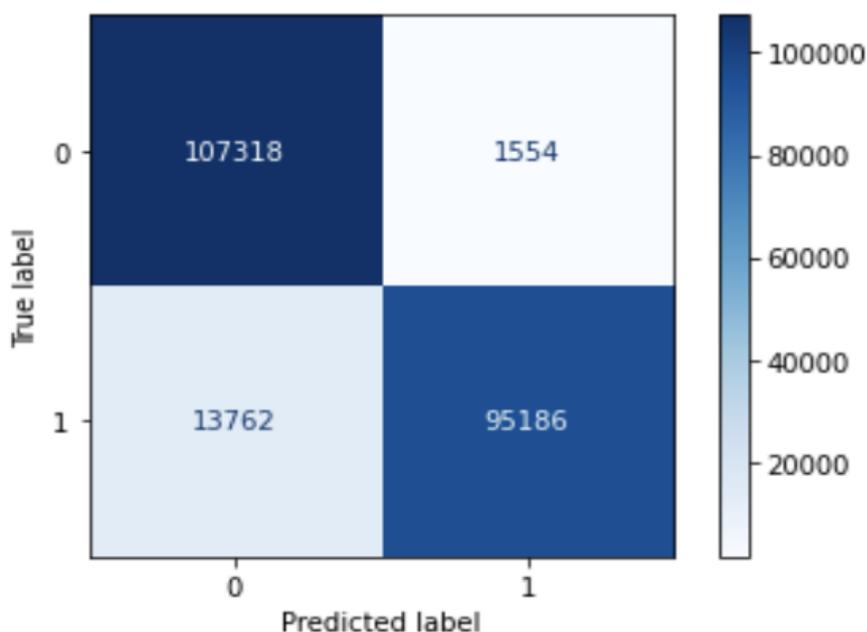


FIGURA 4.6 – Matriz de confusão (Amostra balanceada)

Com os resultados obtidos pôde-se alcançar os fatores que tem maior influência sobre a pontualidade de partida dos voos. Na Figura 4.7 eles estão ordenados de forma decrescente em importância.

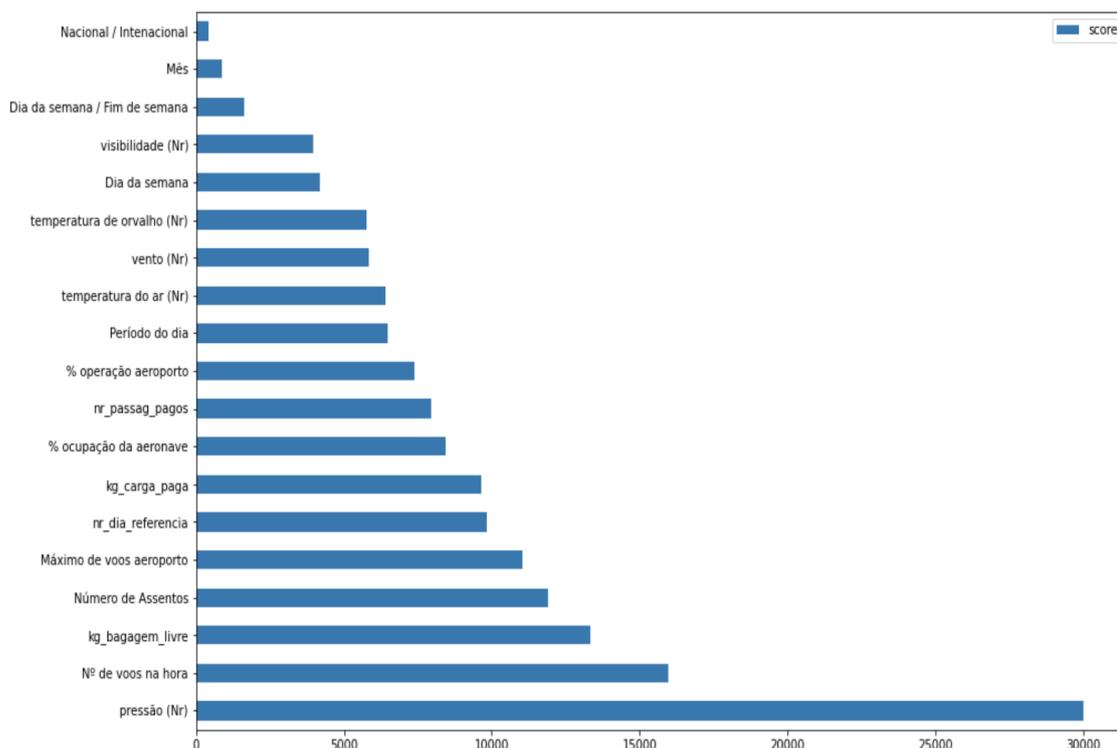


FIGURA 4.7 – Importância das variáveis de entrada

Com o objetivo de investigar em mais detalhes sobre o impacto dos fatores em estudo,

inicia-se a análise das variáveis partindo-se do resultado obtido na Figura 4.7.

O primeiro diagnóstico realizado foi em relação à pressão no momento da decolagem do voo. O histograma presente na Figura 4.8 mostra para cada valor de pressão a pontualidade de partida e visualiza-se que essa é maior a medida que aumenta-se a pressão, bem como a acurácia do modelo.

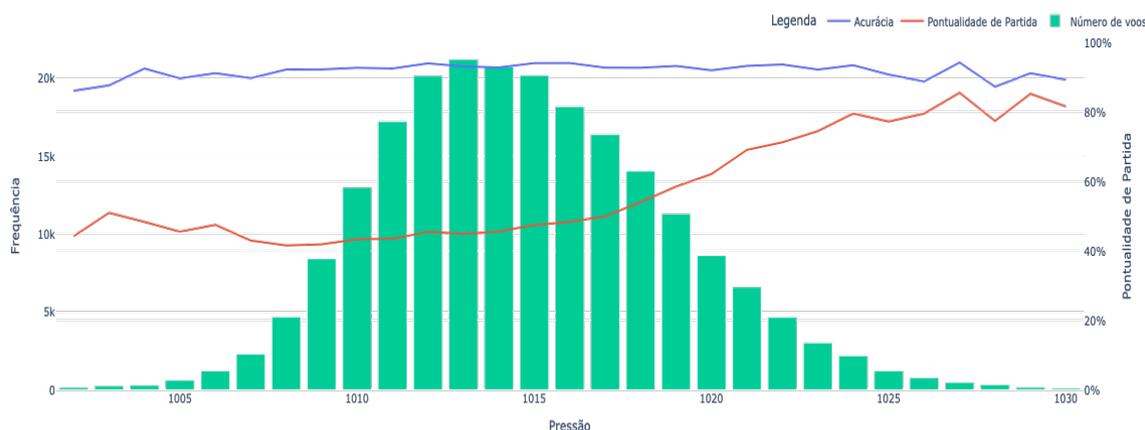


FIGURA 4.8 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão

Além disso, vê-se que a acurácia do modelo obtida está próxima a 90% para todos os intervalos de pressão.

Torna-se nítido na Figura 4.8 que a pressão apresenta grande importância na previsibilidade de pontualidade diante de correlação positiva entre pressão e pontualidade. A partir de maiores pressões tem-se maior tempo de ocupação das pistas o que ocasiona maior congestionamento de aeronaves e por fim aumenta a probabilidade de atrasos.

O fator que possui a segunda maior importância é o número de voos com decolagem na mesma hora em cada aeroporto. Os resultados obtidos para essa variável estão na Figura 4.9.

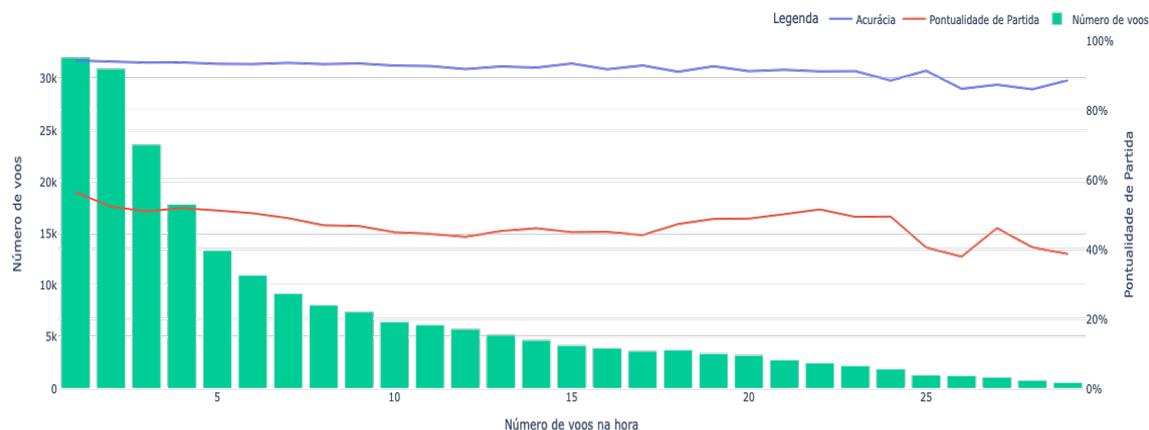


FIGURA 4.9 – Influência do números de voos na hora na pontualidade de partida e na acurácia da previsão

Diante do histograma obtido, visualiza-se que a acurácia do modelo mostrou-se alta em todos os cenários de demanda de voos. Além disso, observa-se que a pontualidade de partida diminui em cenários de maior volume de decolagens no intervalo de uma hora nos aeroportos brasileiros.

Com isso, tem-se uma correlação negativa entre o volume de decolagens por hora e pontualidade de partida.

Na sequência analisa-se o fator que possui a 3^a maior importância dentre todas as variáveis do estudo que é o volume de bagagem livre correspondente à bagagem franqueada pelos passageiros.

Na Figura 4.10 tem-se a análise da pontualidade de partida e acurácia para diversos intervalos (de 10 em 10 kilogramas) de volume total de bagagem livre.

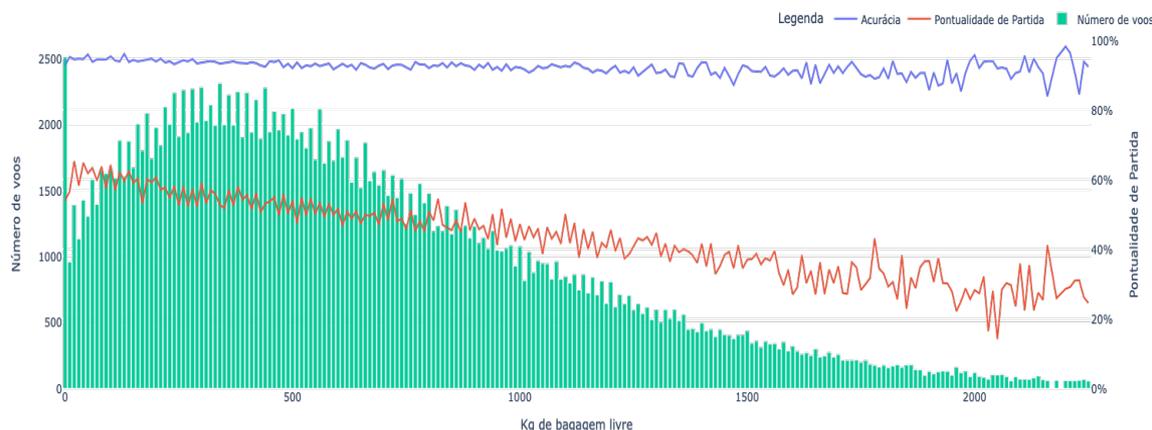


FIGURA 4.10 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão

Novamente destaca-se que a acurácia do modelo está próxima a 90% para todos os intervalos analisados e além disso visualiza-se que a pontualidade de partida diminui consideravelmente a medida que o volume total de bagagem aumenta, o que se tinha como esperado dado que a logística de despacho e armazenamento das bagagens torna-se mais complexa com o aumento do volume..

4.3 Análise dos fatores de performance nos aeroportos mais movimentados

De acordo com a Figura 4.1 tem-se que dentre os aeroportos em estudo, alguns dos mais movimentados em número de decolagens são:

- *SBGR* Aeroporto Internacional de São Paulo/Guarulhos
- *SBKP* Aeroporto Internacional de Viracopos
- *SBCF* Aeroporto Internacional de Confins

Faz-se o diagnóstico para cada um desses aeroportos com relação à pontualidade de partida dos voos.

4.3.1 SBGR - Aeroporto Internacional de Guarulhos

O Aeroporto de Guarulhos possui o maior volume de voos dentre todos os sítios aeroportuários do país com um total de 79.357 voos no período analisado.

O modelo de aprendizado de máquina construído pode ser analisado levando-se em consideração especificamente o Aeroporto de Guarulhos e após processamento considerando a base de dados constituída somente desse aeroporto, chega-se a acurácia também bastante alta com valor de 90,2%.

A matriz de confusão obtida está na Figura 4.11:

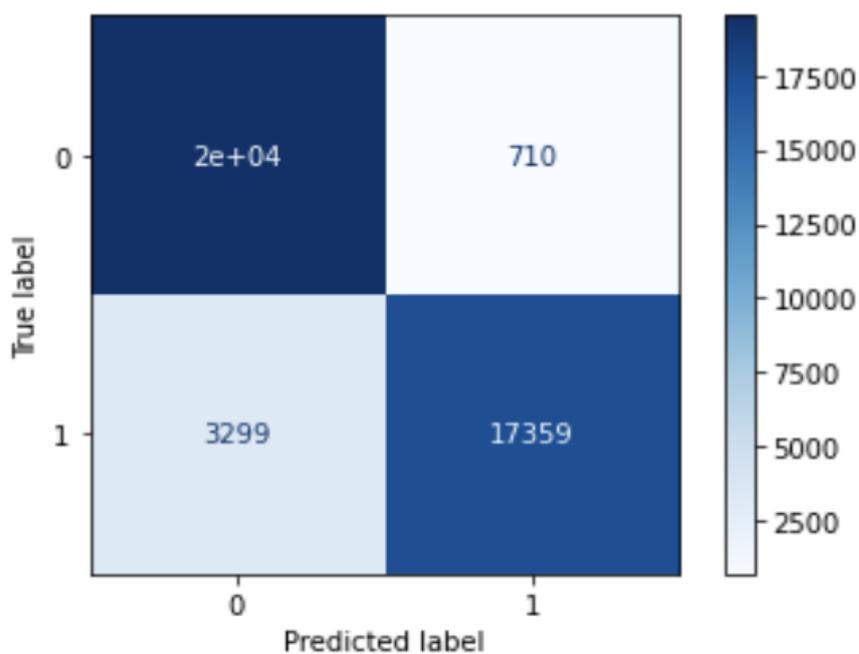


FIGURA 4.11 – Matriz de confusão - Aeroporto de Guarulhos

Com o objetivo de analisar os principais fatores que influenciam na pontualidade dos voos obteve-se a ordem crescente de importância das variáveis e os resultados obtidos estão presented na Figura 4.12.

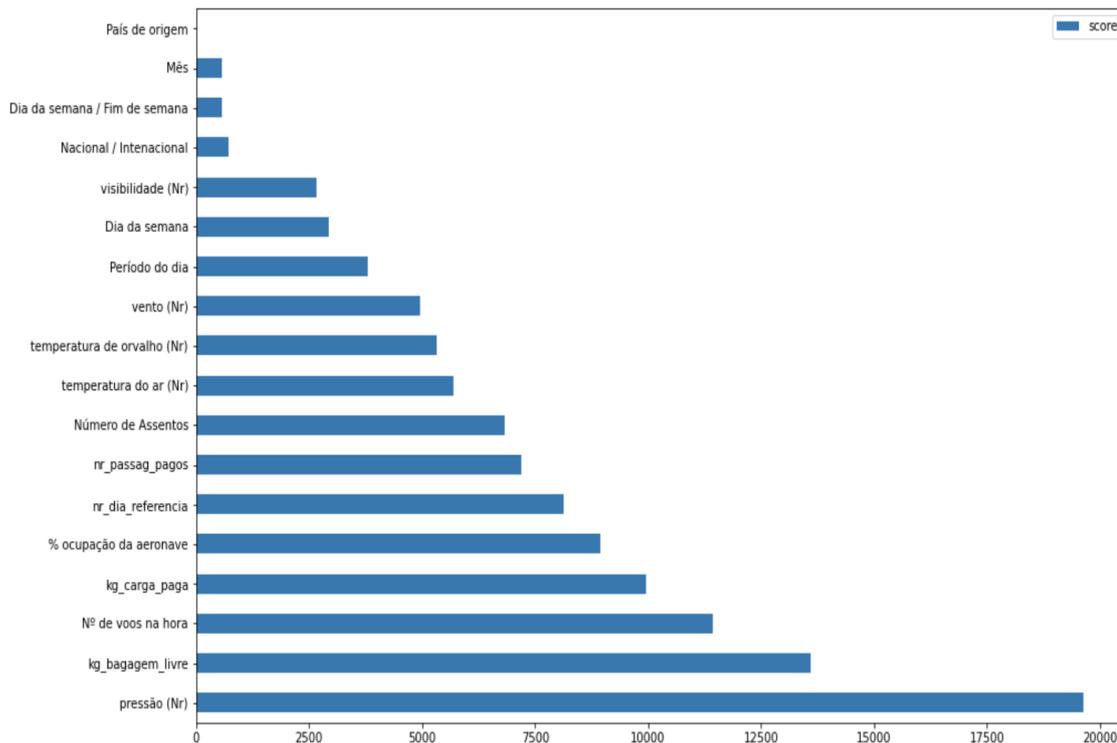


FIGURA 4.12 – Importância das variáveis de entrada - Guarulhos

Portanto, os fatores com maior importância são respectivamente a pressão, o volume de bagagem franqueada e o número de voos no intervalo de uma hora. Ressalta-se que esses fatores são os mesmos obtidos na análise conjunta de todos os aeroportos, porém nessa investigação tem-se a inversão entre o volume de bagagem e o número de voos na hora.

Investiga-se o comportamento de cada uma dessas variáveis graficamente iniciando-se pela pressão. O resultado obtido está presente na Figura 4.13.

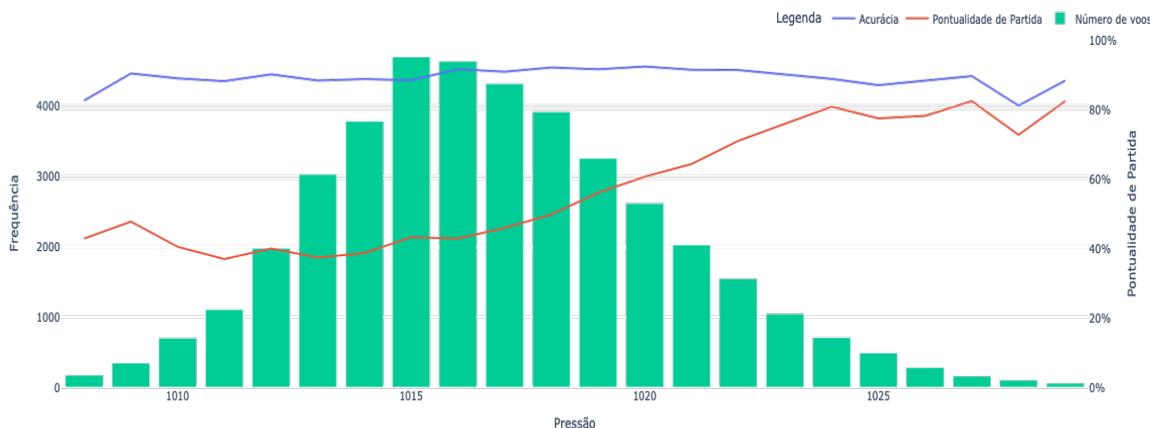


FIGURA 4.13 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Guarulhos

Retira-se do gráfico que para maiores valores de pressão no momento da decolagem tem-se que a porcentagem de voos pontuais também é alta. Com isso, existe uma correlação positiva entre a pressão local e a pontualidade de partida dos voos em Guarulhos.

Na sequência, analisa-se o volume de bagagem franqueada e sua correlação com a pontualidade de partida. Essa relação está presente na Figura 4.14

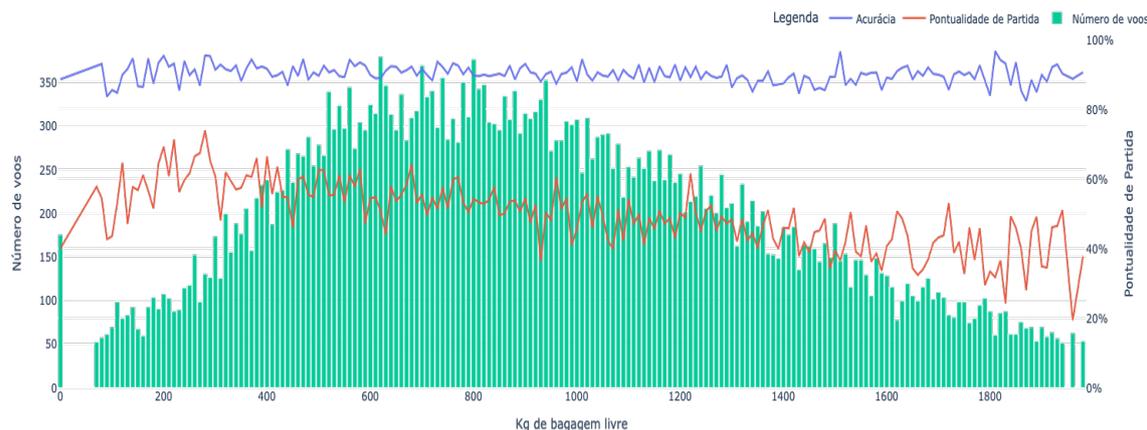


FIGURA 4.14 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Guarulhos

O gráfico mostra que existe uma maior variabilidade da pontualidade de partida nos intervalos de volume de bagagem franqueada, porém em uma visão macro percebe-se que para um volume maior de bagagem tem-se menores pontualidades de partida, o que corrobora com o esperado dada a logística de coleta e armazenamento das bagagens se tornar novamente mais complexa com o aumento do volume.

Por fim, estuda-se o comportamento da variável referente ao número de voos que decolaram no aeroporto no intervalo de hora de cada voo da base de dados. Obteve-se então a Figura 4.15.

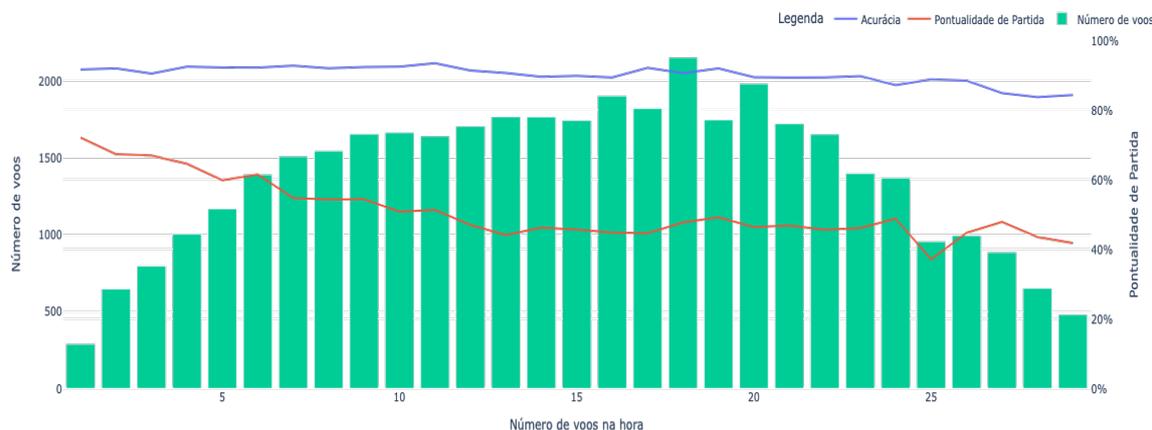


FIGURA 4.15 – Influência do número de voos na hora na pontualidade de partida e na acurácia da previsão - Guarulhos

Similar ao resultados obtido considerando-se todos os aeroportos, tem-se na Figura 4.15 que no aeroporto de Guarulhos, com o aumento do número de voos que decolaram na mesma hora que o respectivo voo tem-se uma queda na pontualidade de partida. Destaca-se que isso corrobora o resultado esperado dado que com o aumento da demanda, chega-se mais próximo da capacidade de operação do sítio aeroportuário e atrasos em áreas como check-in, inspeção de segurança e deslocamento da aeronave tornam-se mais prováveis.

4.3.2 SBKP - Aeroporto Internacional de Viracopos

O Aeroporto de Viracopos possui o segundo maior volume de voos dentre do aeroportos brasileiros com um total de 52.507 voos.

Fez-se o processamento do algoritmo preditivo da pontualidade filtrando-se esse aeroporto e obteve-se a matriz de confusão da Figura 4.16 e acurácia de 93,3%.

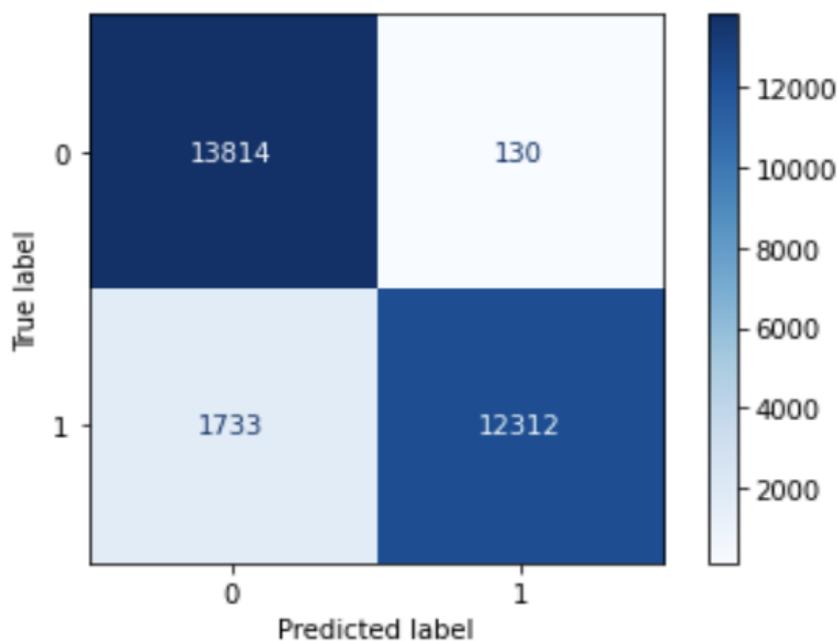


FIGURA 4.16 – Matriz de confusão - Aeroporto de Viracopos

Para analisar os principais fatores influenciadores obteve-se a importância das variáveis de entrada e o resultado está presente na Figura 4.17.

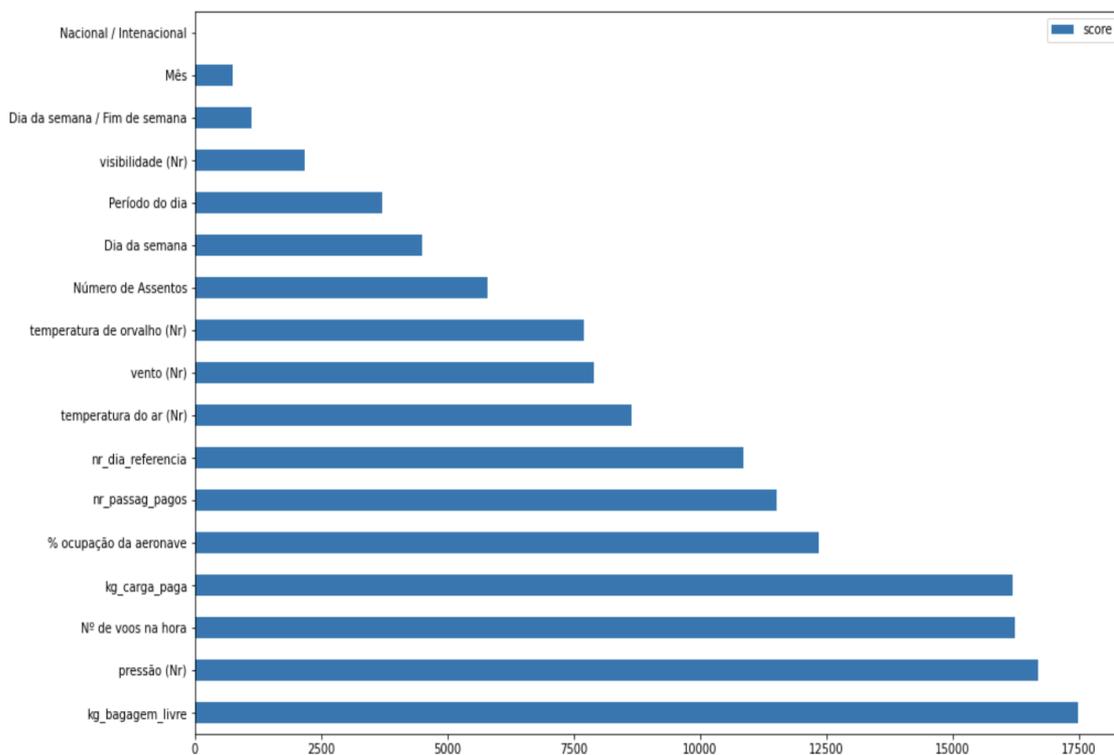


FIGURA 4.17 – Importância das variáveis de entrada - Viracopos

No cenário de Viracopos tem-se novamente as variáveis com maior relevância como sendo a pressão, o volume de bagagens e o número de voos simultâneos no aeroporto.

O fator de maior revelância é o volume de bagagens franqueada e sua relação com a pontualidade pode ser visualizado na Figura 4.18.

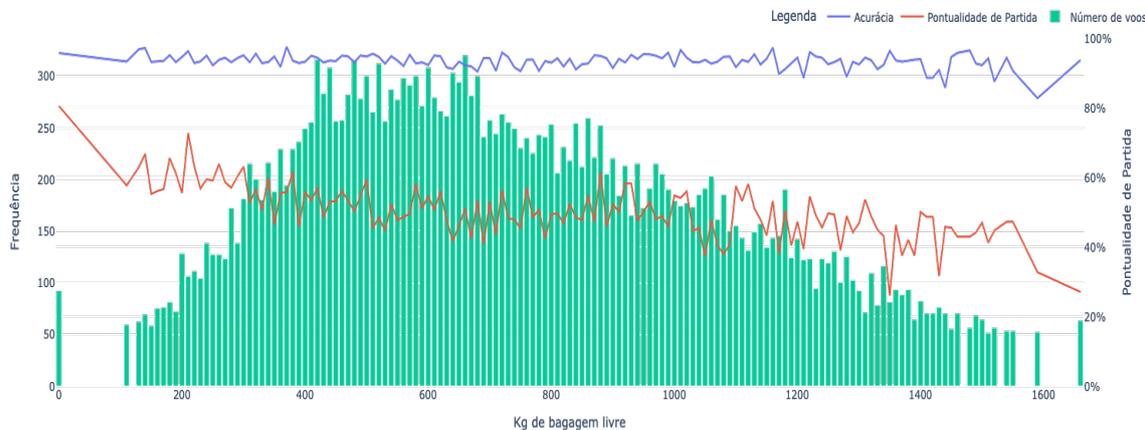


FIGURA 4.18 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Viracopos

Observa-se que o aumento no volume de bagagens franqueadas ocasiona uma maior probabilidade de atraso na partida dos voos em uma visão macro dos resultados.

Na Figura 4.19 tem a relação entre a pressão, a pontualidade de partida e a acurácia do modelo.

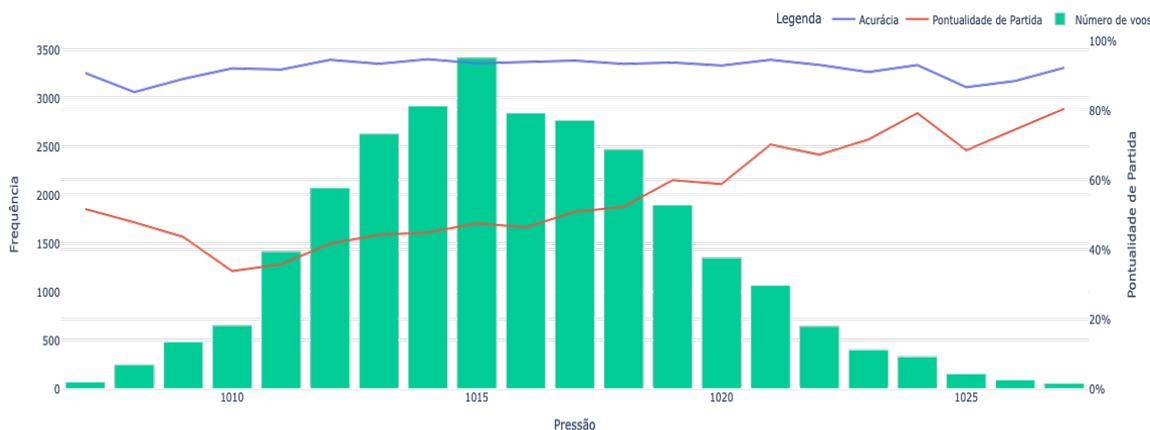


FIGURA 4.19 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Viracopos

Visualiza-se que com o aumento do valor da pressão no momento da decolagem, tem-se maior probabilidade de o voo ser classificado como pontual. Assim tem-se uma correlação positiva entre essas variáveis.

No quesito número de voos na hora da decolagem tem-se na Figura 4.20 os resultados para o aeroporto de Viracopos.

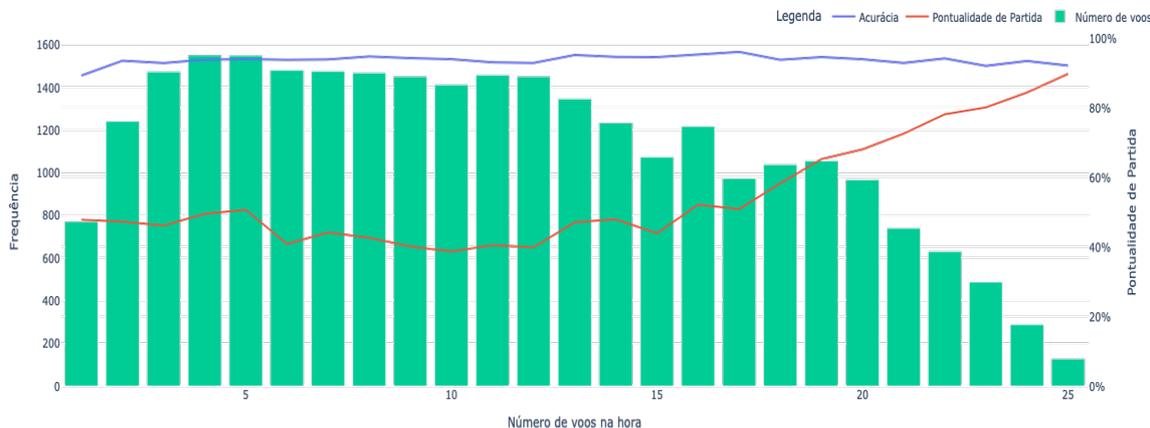


FIGURA 4.20 – Influência do número de voos na hora na pontualidade de partida e na acurácia da previsão - Viracopos

Diante do resultado, verifica-se que quando o aeroporto está mais movimentado, ou seja, com um maior de decolagens no intervalo de uma hora, a pontualidade aumenta. Isso se contrapõe ao resultado visualizado no Aeroporto de Guarulhos e na análise nacional e demonstra que o aeroporto de Viracopos possui alta performance sob condições de grande demanda.

4.3.3 SBCF - Aeroporto Internacional de Confins

O Aeroporto de Confins, localizado em Belo Horizonte é o principal aeroporto de Minas Gerais, segundo estado mais populoso do país. No período de análise teve um total de 34.612 voos.

O algoritmo construído obteve acurácia de 95,2% quando processado somente para Confins. A respectiva matriz de confusão está presente na Figura 4.21

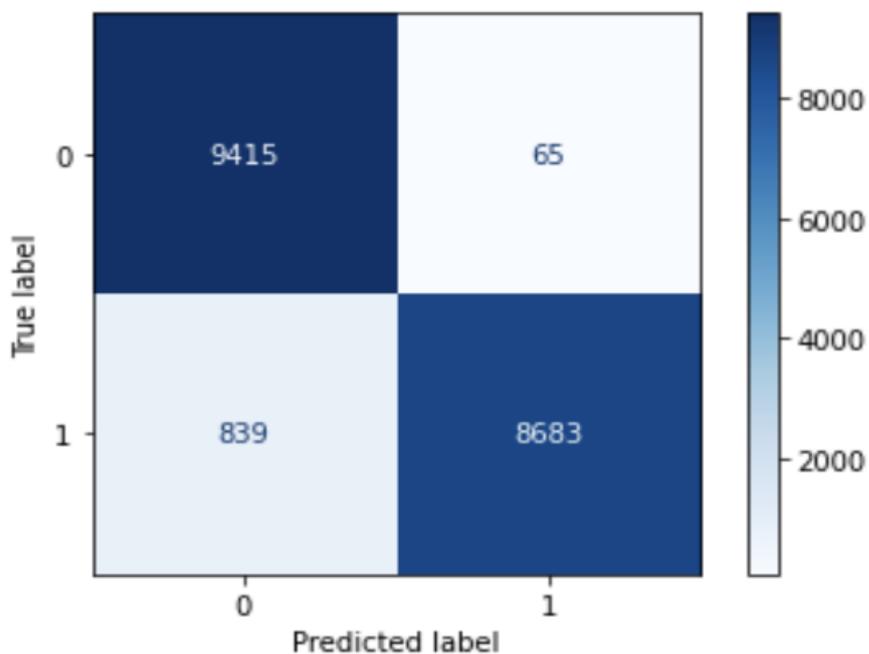


FIGURA 4.21 – Matriz de confusão - Aeroporto de Confins

Fazendo-se a análise de relevância entre as variáveis de entrada, tem-se a ordem crescente presente na Figura 4.22.

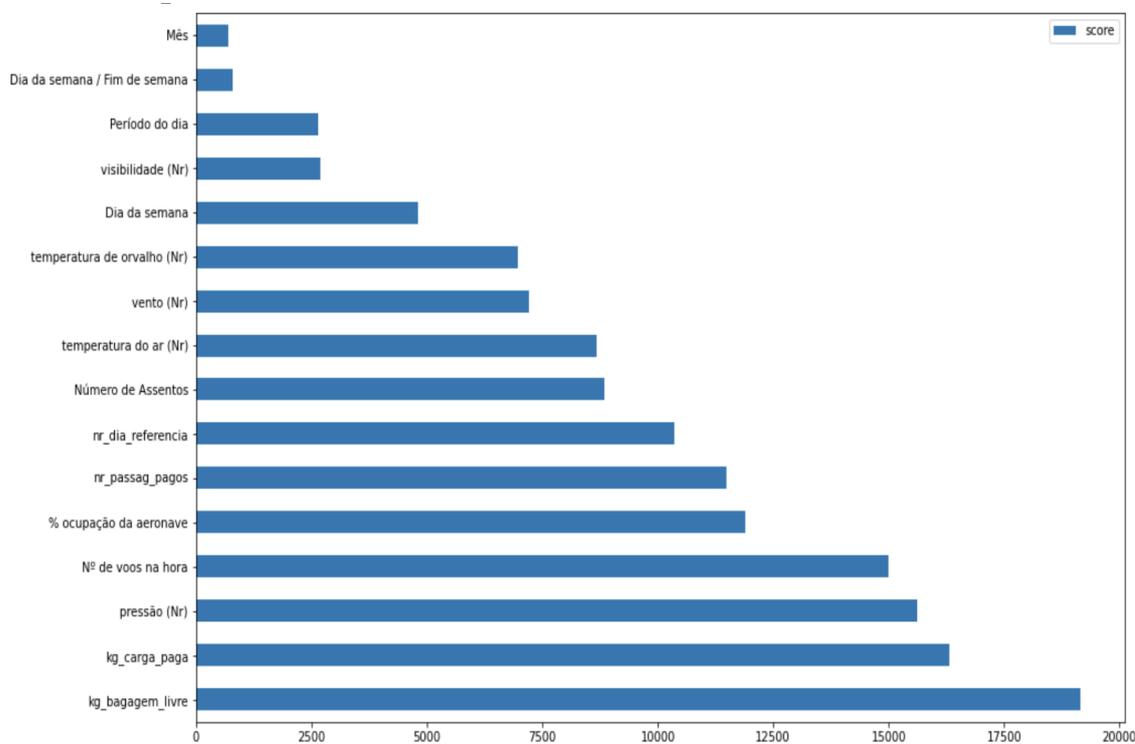


FIGURA 4.22 – Importância das variáveis de entrada - Confins

Tem-se nessa análise que os fatores de maior importância são o volume de bagagens e de carga franqueada, além da pressão no local da decolagem.

Inicia-se pela investigação sobre a relação entre o volume de bagagens despachada e a pontualidade de partida dos voos em Confins que está presente na Figura 4.23.

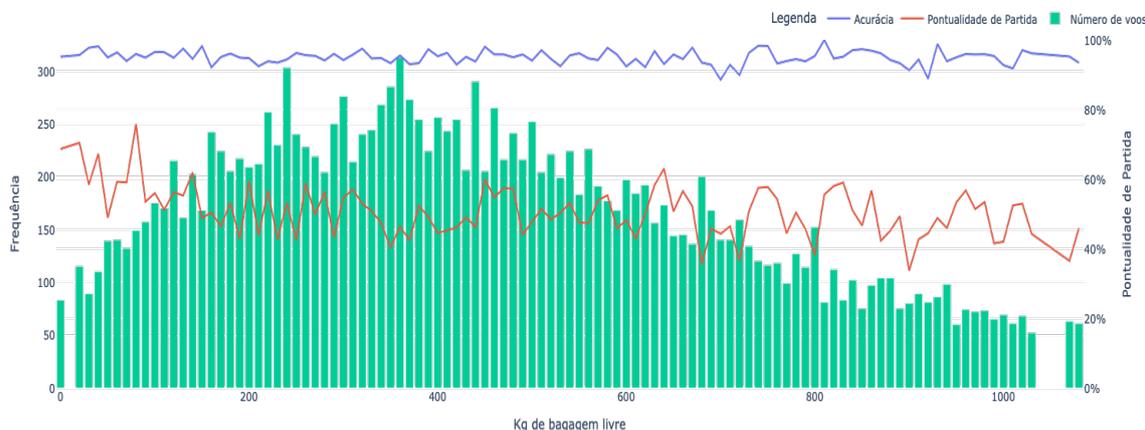


FIGURA 4.23 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - Confins

Diante da figura, visualiza-se que o aumento do volume de bagagens ocasiona maior probabilidade de o voo decolar após 15 minutos do respectivo horário agendado para decolagem.

Acerca da variável referente à carga paga presente no avião, tem-se presente na Figura 4.24 os resultados obtidos.

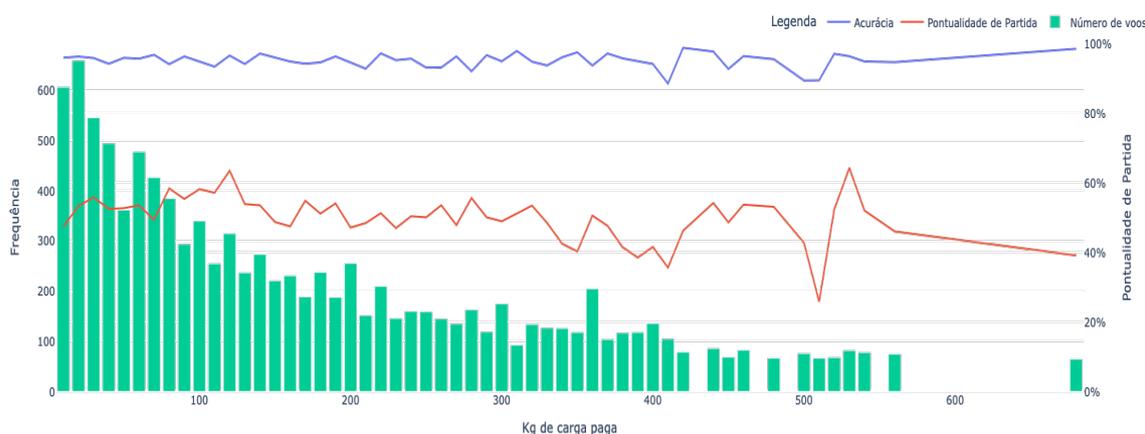


FIGURA 4.24 – Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - Confins

Observa-se que para voos com maior volume de carga paga na aeronave, tem-se maior probabilidade de atraso na decolagem e além disso tem-se acurácia próxima à 100%.

No que tange à pressão local do Aeroporto dw Confins e sua relação com a pontualidade de partida, tem-se na Figura 4.25 os resultados.

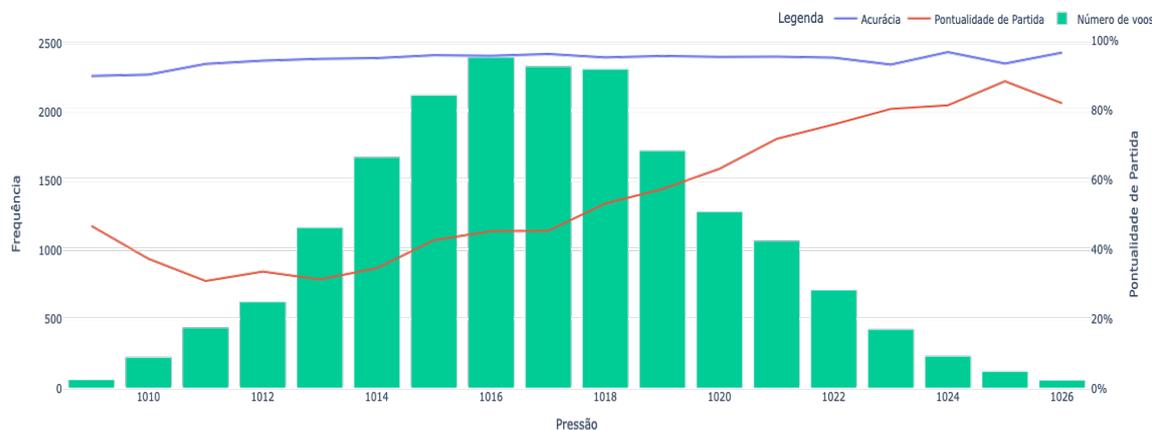


FIGURA 4.25 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - Confins

Evidência-se que há grande influência sobre a pontualidade por parte da pressão local. A pontualidade aproxima-se de 90% para voos com as maiores pressões.

4.4 Análise dos fatores de performance para as companhias aéreas com maior marketshare

As companhias que possuem maiores quotas de mercado nos aeroportos de estudo em relação ao volumes de decolagens são:

- *GOL* - LINHAS AÉREAS S.A.
- *TAM* - LINHAS AÉREAS S.A.
- *AZUL* - LINHAS AÉREAS S.A.

4.4.1 GOL - LINHAS AÉREAS S.A.

A GOL possui 27,4% do total de decolagens no período analisado e obteve-se 91,6% de acurácia em relação ao modelo construído para especificamente para essa empresa conforme matriz de confusão da Figura 4.26.

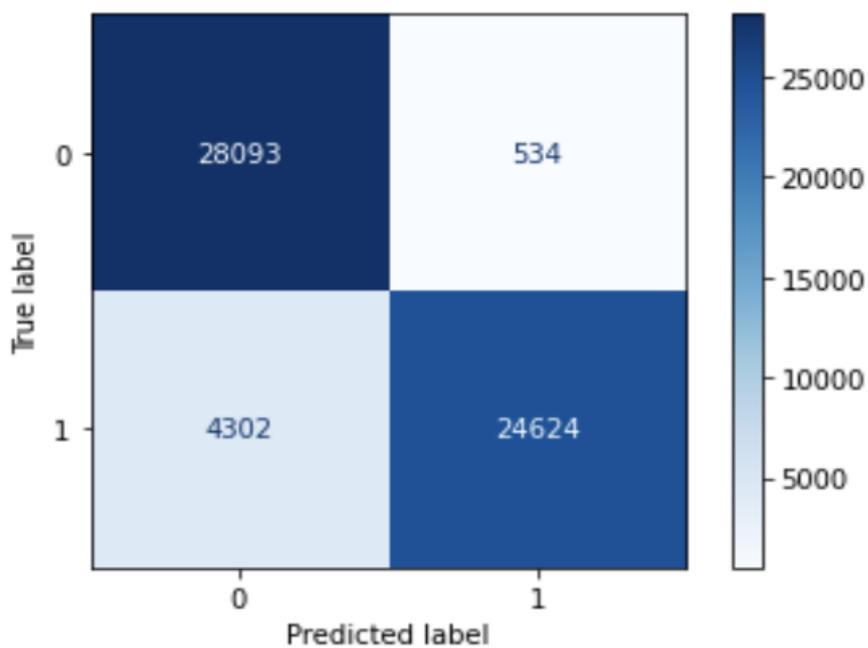


FIGURA 4.26 – Matriz de confusão - GOL

Acerca da importância das variáveis atingiu-se o resultado presente na Figura 4.27.

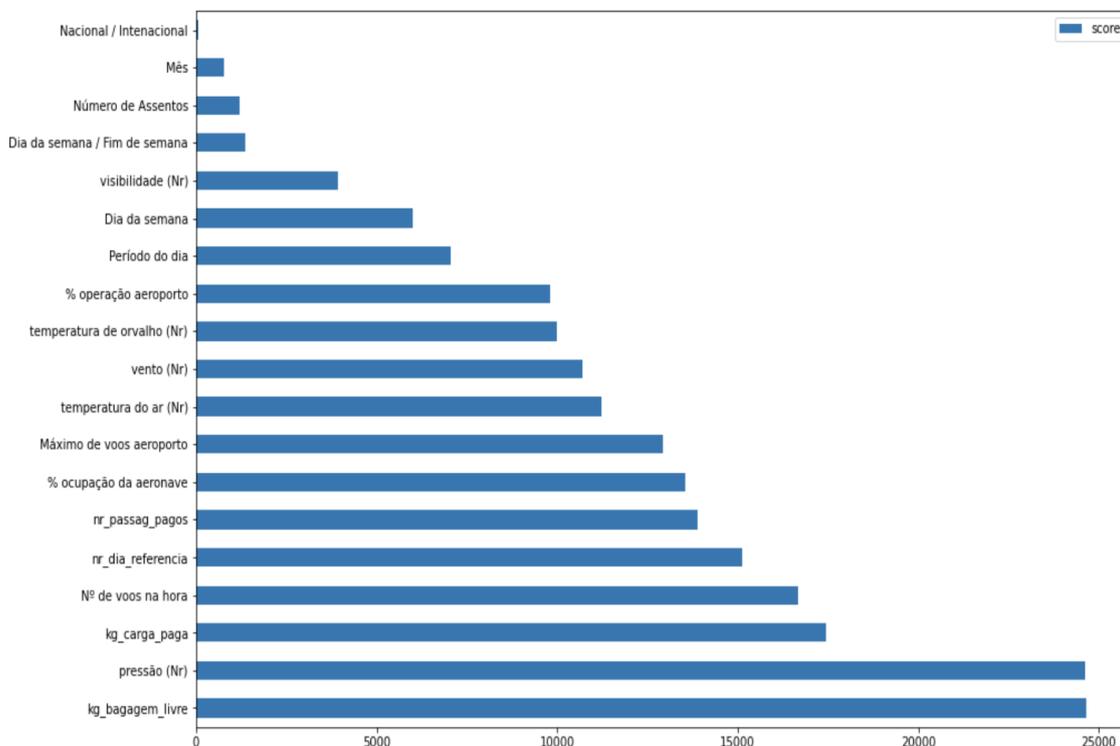


FIGURA 4.27 – Importância das variáveis de entrada - GOL

Destaca-se que os fatores que mais influenciam sobre a performance de decolagem das aeronaves da GOL são o volume de bagagens franqueada, pressão e volume de carga paga.

Analisa-se a variável relacionada ao volume de bagagens dos passageiros na Figura 4.28 e visualiza-se que o aumento do volume bagagem ocasiona maior probabilidade de atraso do voo para todos os trechos de volume de bagagem.

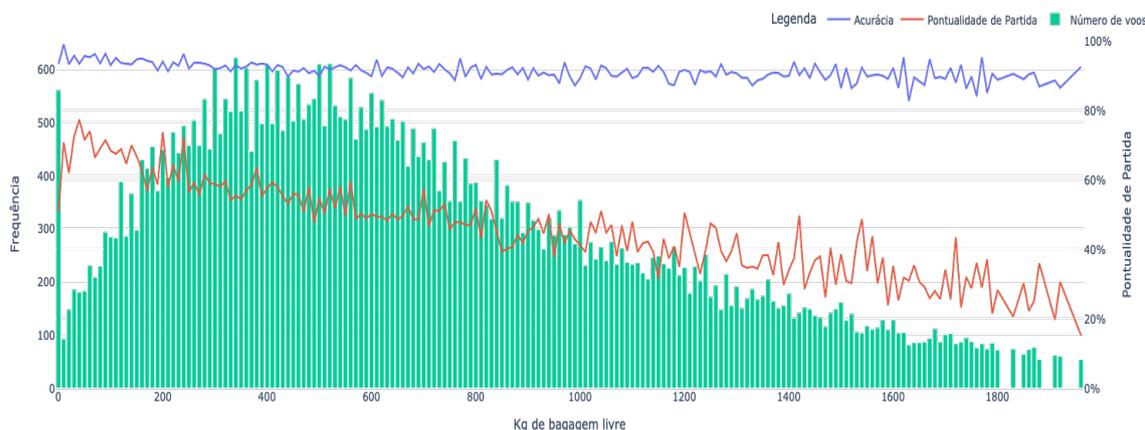


FIGURA 4.28 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - GOL

Na sequência constrói-se análise similar para a pressão local na Figura 4.29. e observa-se que para maiores valores de pressão, novamente tem-se maior probabilidade de o voo ser pontual.

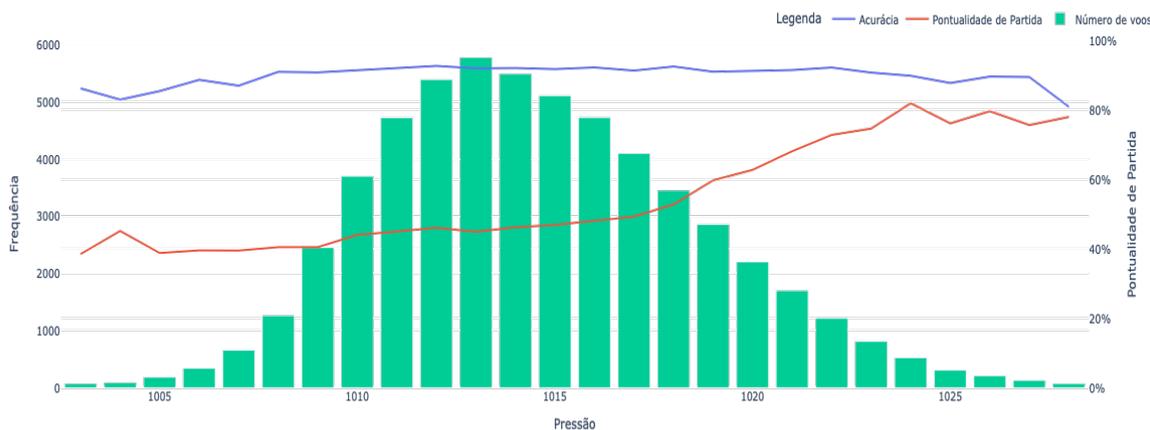


FIGURA 4.29 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - GOL

Por fim, em relação ao volume de carga paga tem-se a análise na Figura 4.30 e visualiza-se que principalmente para voos com grande quantidade de carga paga tem-se menor índice de pontualidade de partida dos voos.

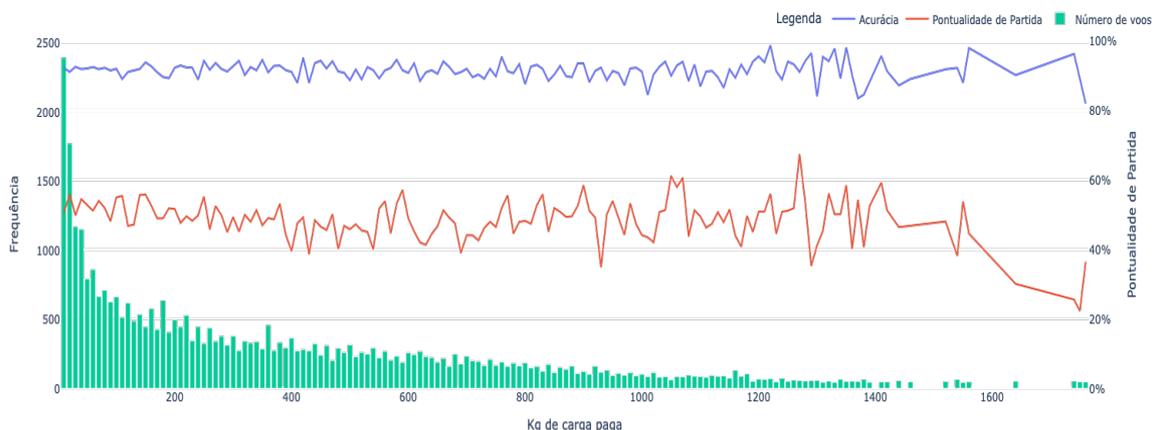


FIGURA 4.30 – Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - GOL

4.4.2 TAM - LINHAS AÉREAS S.A.

Acerca da companhia aérea TAM obteve-se que ela possui 28,4% do número de voos no período analisado e o modelo de predição após processado atingiu acurácia de 93,9%.

Nas Figuras 4.31 e 4.32 estão respectivamente a matriz de confusão obtida nesse modelo e a relevância dos fatores de entrada na pontualidade de partida.

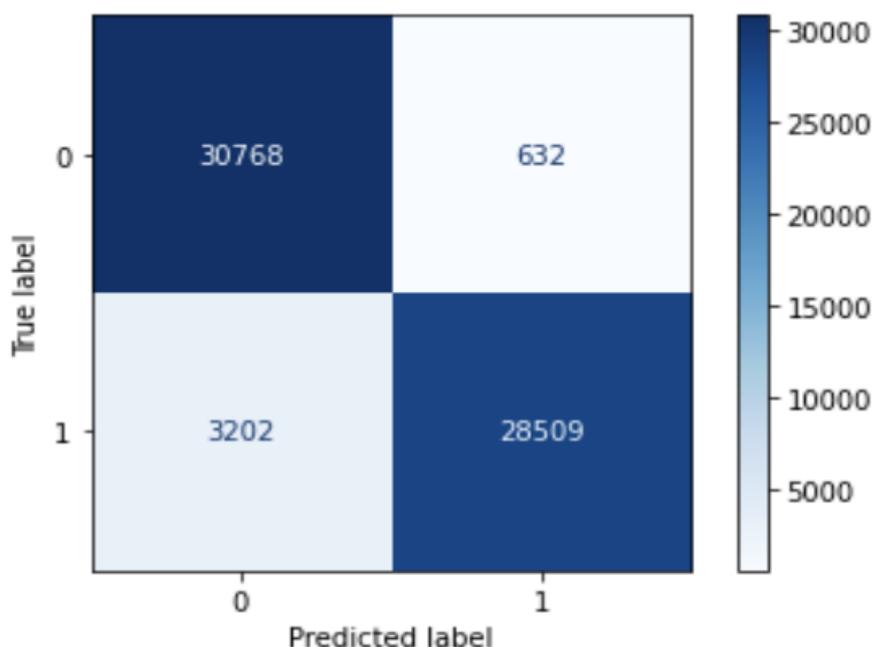


FIGURA 4.31 – Matriz de confusão - TAM

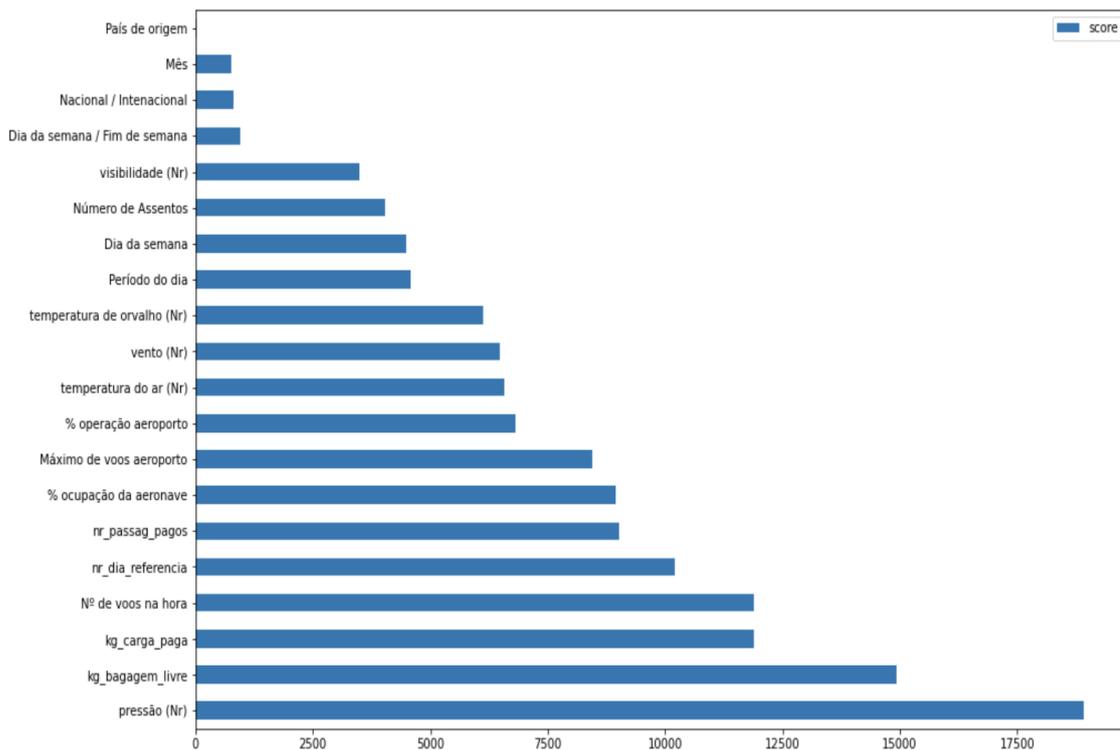


FIGURA 4.32 – Importância das variáveis de entrada - TAM

Destaca-se então que as variáveis pressão local, volume de bagagem franqueada e volume de carga paga possuem as maiores importâncias dentre todos os fatores analisados.

Acerca da pressão tem-se na Figura 4.33 para cada valor inteiro da pressão a pontualidade de partida e com isso visualiza-se que o aumento da pontualidade ocorre ao ser aumentada a pressão.

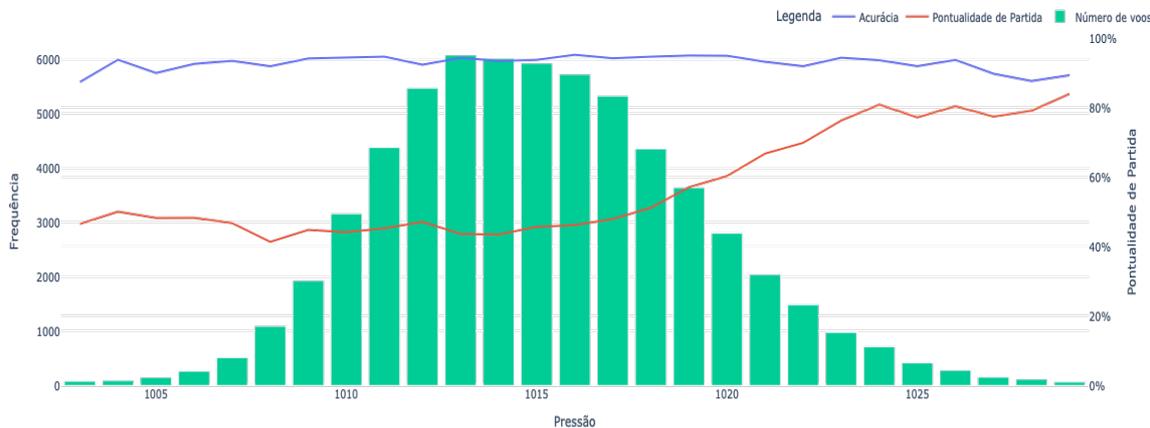


FIGURA 4.33 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - TAM

A seguir analisa-se a relação do volume de bagagem franqueada e a pontualidade e

visualiza-se na Figura 4.34 que para voos com menor volume de bagagens tem-se maior probabilidade de a partida acontecer pontualmente.

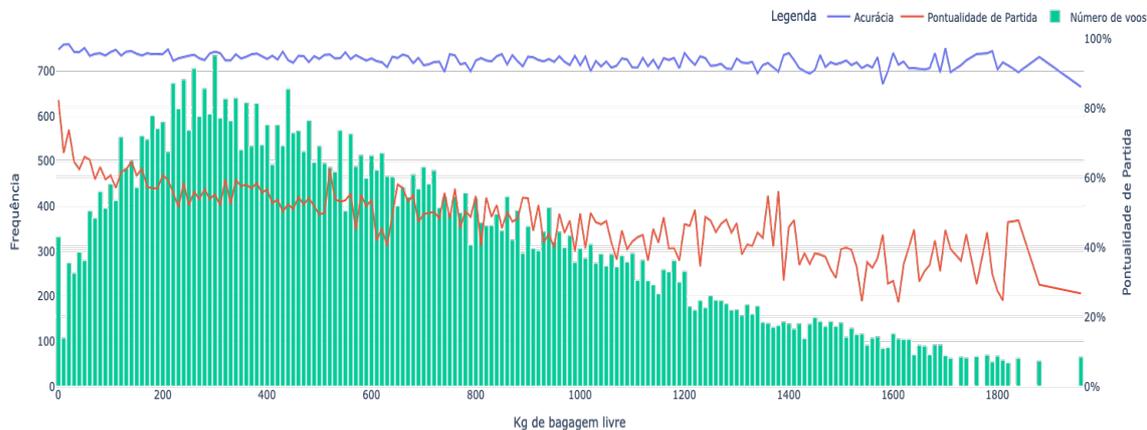


FIGURA 4.34 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - TAM

Sobre o volume de carga paga tem-se na Figura 4.35 que voos com grande quantidade de carga tendem à apresentar maior probabilidade de atraso na decolagem. Isso é corroborado pela complexidade da logística de gerenciamento e armazenamento dessa carga.



FIGURA 4.35 – Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - TAM

4.4.3 AZUL - LINHAS AÉREAS S.A.

A Azul possui 41,7% do volume de voos que decolaram dos aeroportos em análise, caracterizando como a 1^a companhia aérea com mais voos.

Aplicando-se o modelo de predição para essa companhia, obteve-se acurácia de 94,0%.

Nas Figuras 4.36 e 4.37 tem-se respectivamente a matriz de confusão do modelo e a relevância das variáveis de entrada.

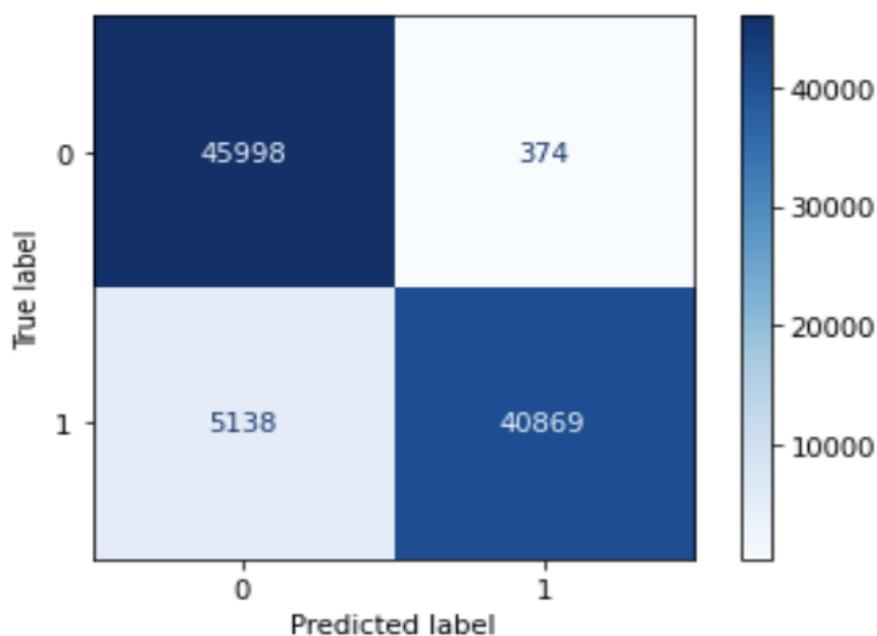


FIGURA 4.36 – Matriz de confusão - AZUL

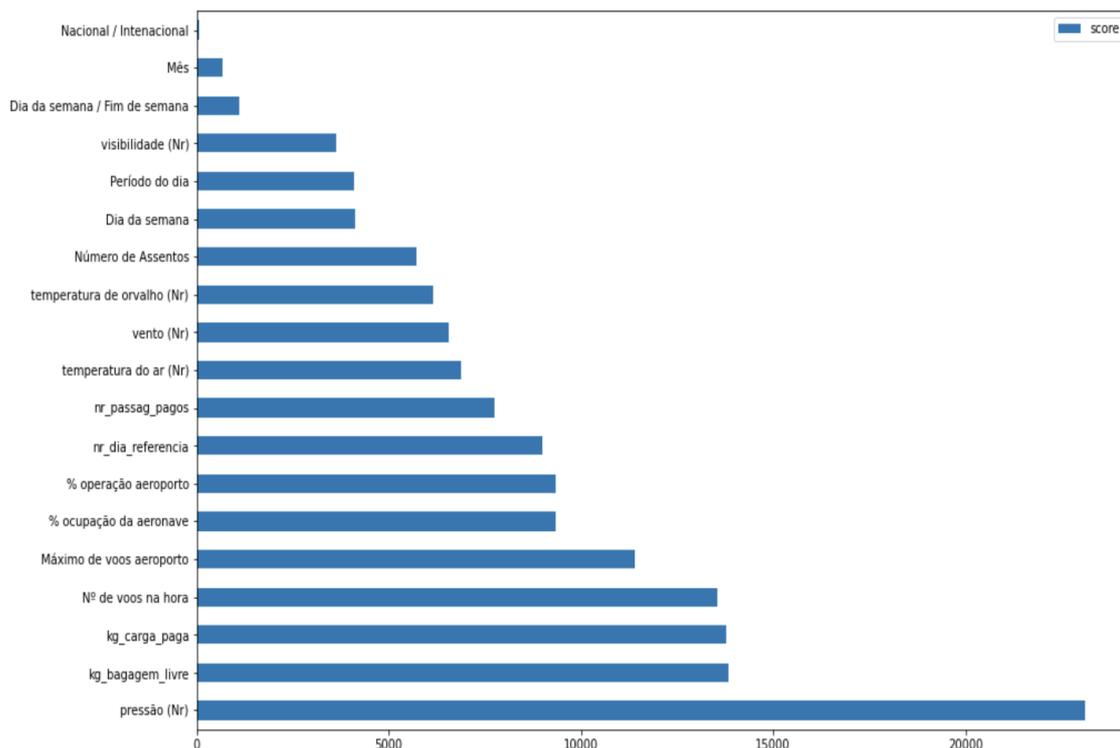


FIGURA 4.37 – Importância das variáveis de entrada - AZUL

Destaca-se que as variáveis mais significativas na predição da pontualidade são a pressão local, o volume de bagagens franqueada e o volume de carga paga.

Na Figura 4.38 mostra-se a pontualidade de partida para cada valor inteiro de pressão e observa-se que para maiores pressões os voos tendem à decolar pontualmente.

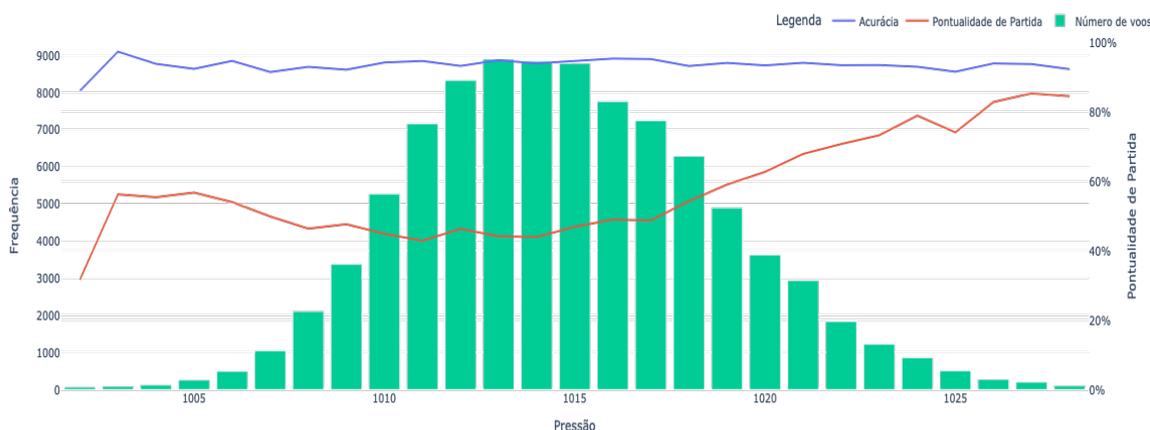


FIGURA 4.38 – Influência da pressão atmosférica na pontualidade de partida e na acurácia da previsão - AZUL

Em relação ao volume de bagagem franqueada tem-se na Figura 4.39 a relação com a pontualidade de partida e visualiza-se que para voos com maior volume de bagagens a

decolagem tem maior probabilidade de acontecer com atraso.

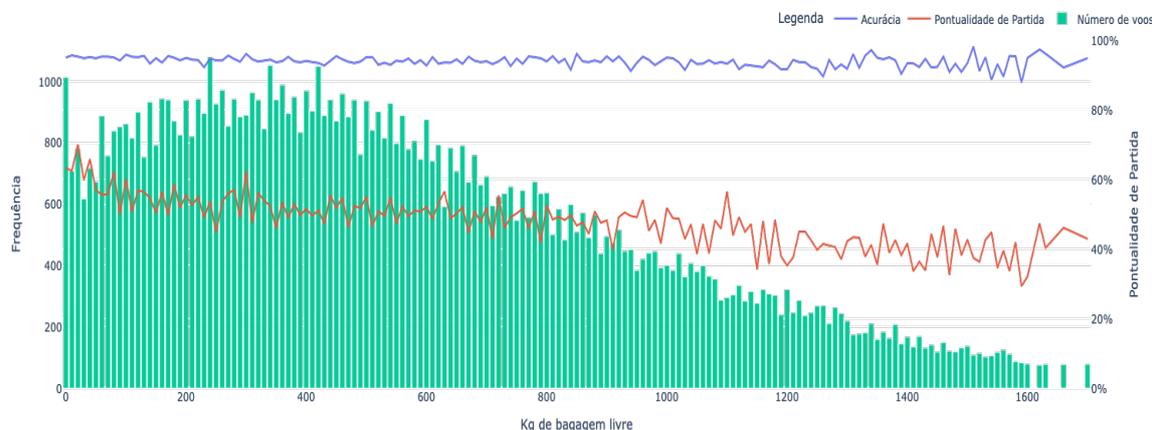


FIGURA 4.39 – Influência do volume de bagagem franqueada na pontualidade de partida e na acurácia da previsão - AZUL

Por fim, no que concerne ao volume de carga paga, visualiza-se na Figura 4.40 que em aeronaves com maior carga paga tem-se maior probabilidade de atraso na decolagem do voo.

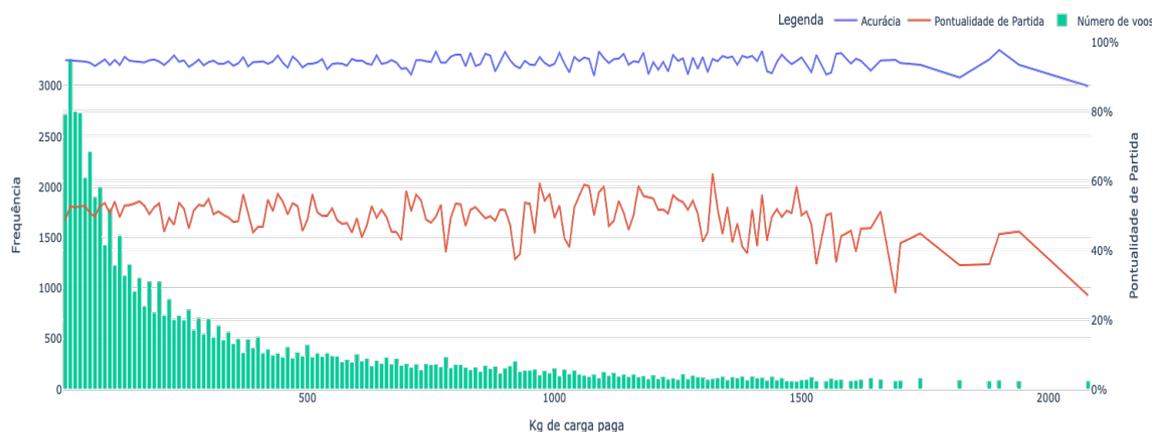


FIGURA 4.40 – Influência do volume de carga paga na pontualidade de partida e na acurácia da previsão - AZUL

4.5 Eficiência de processamento dos aeroportos

De forma a analisar a gestão, de capacidade e demanda, por parte dos gestores dos sítios aeroportuários de forma macro, analisa-se o KPI 11 também presente na MCA 100-22.

Primeiramente observa-se na Figura 4.41 a capacidade de pousos por hora em cada aeroporto do estudo.

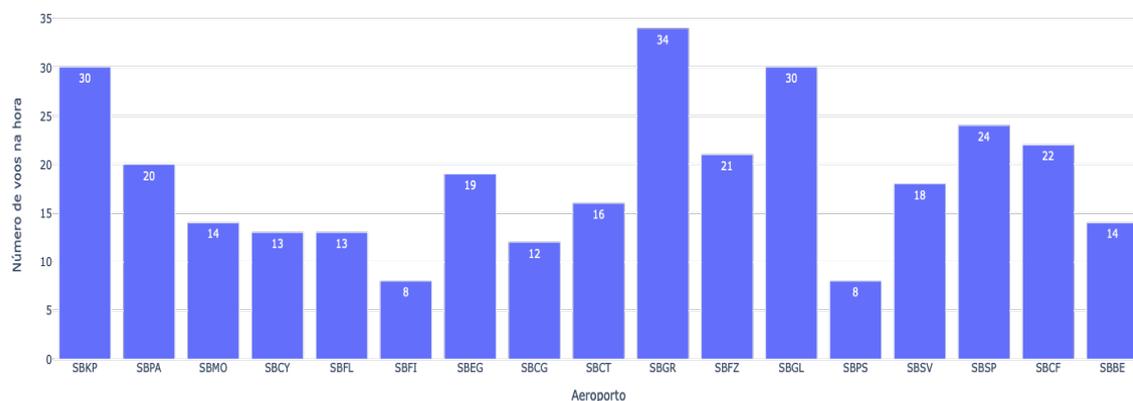


FIGURA 4.41 – Capacidade horária de pousos por aeroporto

Fonte: ANAC

Na Figura 4.42 estão presentes por aeroporto o percentual de dias em que a demanda de pousos é superior a 90% da capacidade horária definida para o respectivo aeroporto.

Destaca-se que essa capacidade está intrinsecamente conectada ao volume de voos no respectivo aeroporto. Cita-se por exemplo os aeroportos situados no estado de São Paulo (Guarulhos, Viracopos e Congonhas) que dentre os sítios em estudo apresentam as maiores demandas no tempo analisado e estão entre os quatro aeroportos com maior capacidade de pouso juntamente com o Aeroporto Internacional Tom Jobim situado no estado do Rio de Janeiro.

O uso da capacidade declarada por cada aeroporto pode ser analisada levando-se em consideração a porcentagem de dias em que aconteceu de existir uma hora tal que a demanda de pousos foi superior a 90% da capacidade.

Tal análise está presente na Figura 4.42 abaixo:

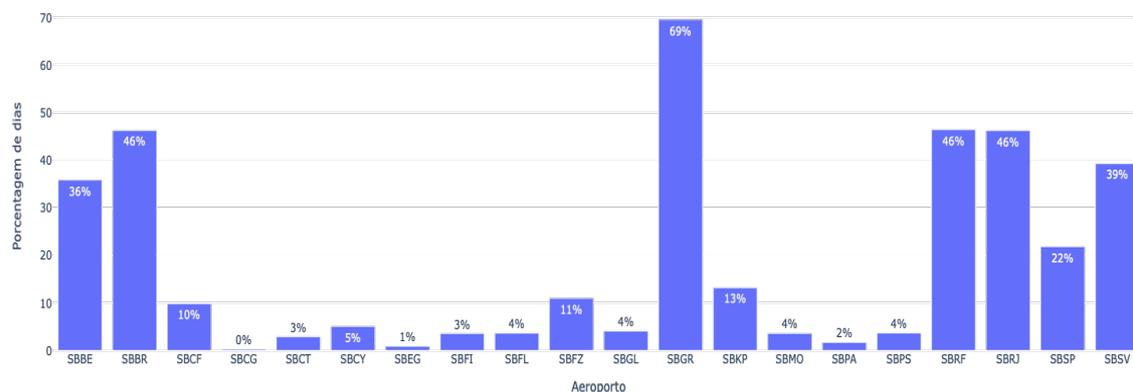


FIGURA 4.42 – Percentual de dias com demanda acima de 90% da capacidade de pousos

Destaca-se que os aeroportos com maior eficiência diária no âmbito da gestão são:

- *SBBE* Aeroporto Internacional de Belém
- *SBBR* Aeroporto Internacional de Brasília
- *SBGR* Aeroporto Internacional de São Paulo/Guarulhos
- *SBRF* Aeroporto Internacional do Recife
- *SBRJ* Aeroporto Santos Dumont
- *SBSV* Aeroporto Internacional de Salvador

Analogamente faz-se a análise da porcentagem de meses que houve uma hora em que a demanda foi superior a 90% da capacidade de pouso para cada aeroporto. Os resultados estão presentes na Figura 4.43.

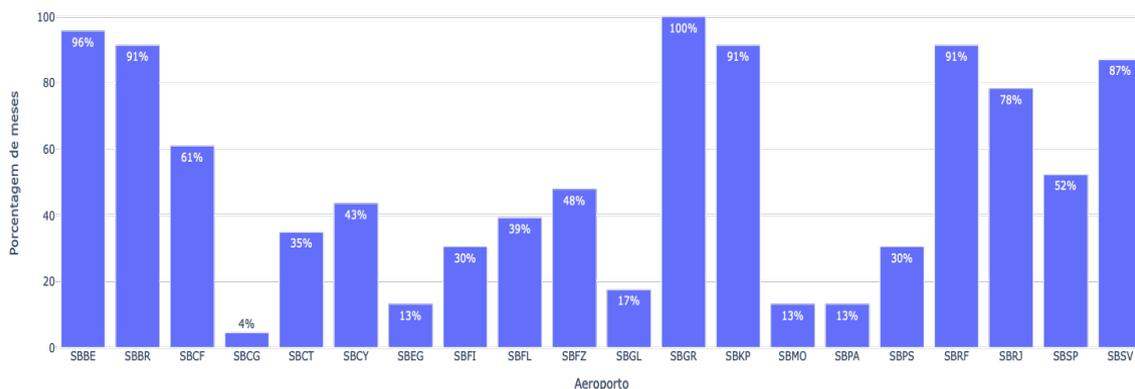


FIGURA 4.43 – Percentual de meses com demanda acima de 90% da capacidade de pousos

Diante desses resultados visualiza-se que alguns sítios aeroportuários apresentam demanda muito abaixo de sua capacidade. Pode-se observar os aeroportos abaixo:

- *SBCG* Aeroporto Internacional de Campo Grande
- *SBEG* Aeroporto Internacional de Manaus
- *SBGL* Aeroporto Internacional de Tom Jobim
- *SBMO* Aeroporto Internacional de Maceió
- *SBPA* Aeroporto Internacional de Porto Alegre

4.6 Diagnóstico da taxa pico de chegada nos aeroportos

Através da metodologia descrita no KPI 10 presente na MCA 100-22 computou-se para cada aeroporto de estudo o número de pousos por hora e após ordenação e seleção do 95º percentil chega-se aos resultados abaixo:

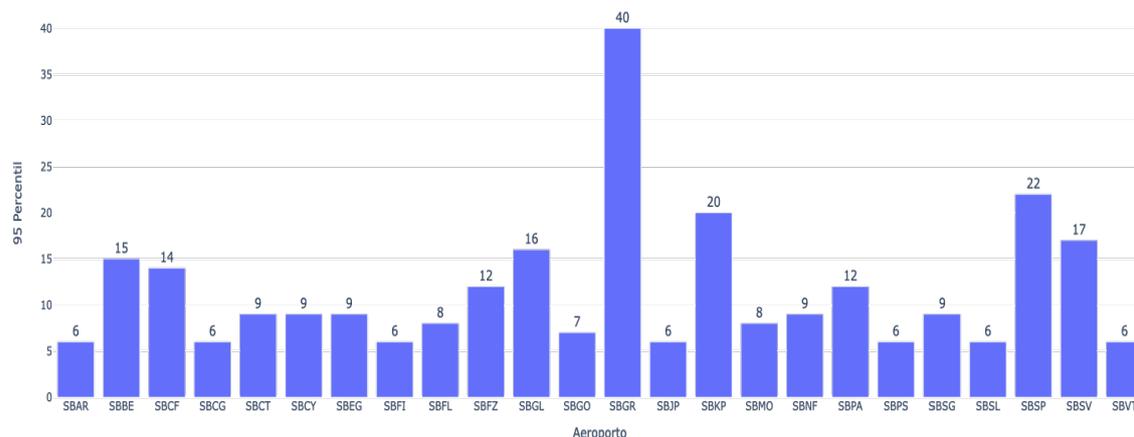


FIGURA 4.44 – 95º percentil do movimento de pousos por aeroporto

A partir disso, pode-se comparar a capacidade de chegada declarada com a taxa pico de forma a entender a qualidade da gestão de demanda de cada aeroporto e previsibilizar a necessidade de futuros investimentos.

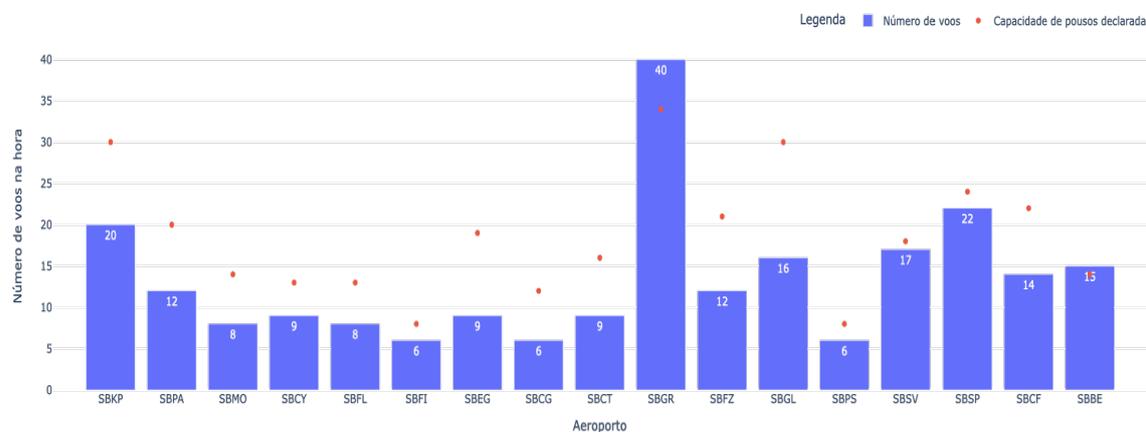


FIGURA 4.45 – Taxa pico e capacidade de pousos por aeroporto

Despreende-se da Figura 4.45 que alguns aeroportos possuem um volume de voos por hora bastante próximos ou até mesmo acima da capacidade de chegada. Isso mostra que em pelo menos 5% do tempo o Aeroporto Internacional de Guarulhos (SBGR) e o Aeroporto Internacional de Belém (SBBE) estão operando acima de 90% de sua capacidade.

Isso reforça que investimentos nas áreas de operação para otimizar a logística de distribuição de voos e financeira de forma a potencializar novos investimentos em infraestrutura para comportar com maior eficiência os volumes de voos são importantes para manutenção do fornecimento de um serviço de qualidade.

Por outro lado, visualiza-se que alguns sítios aeroportuários possuem capacidade bastante superior à demanda na hora de pico como por exemplo o Aeroporto Internacional de Manaus (SBEG) e Aeroporto Internacional Tom Jobim (SBGL). Esses aeroportos na hora de pico possuem próximo à 50% de demanda em relação à capacidade declarada de pousos.

Esse fato demonstra que por exemplo a área de relacionamento externo com companhias aéreas pode concentrar seus esforços em estabelecer novas parcerias de forma à disponibilizar maior volume de voos. Além disso, estabelecer maior número de parcerias com outros sítios aeroportuários regionais próximos e sítios com alta demanda espalhados pelo Brasil fará com que o número de pousos aumente.

5 Conclusão

O trabalho desenvolvido teve por principal objetivo estudar o desempenho da operação de aeroportos brasileiros através de indicadores de performance. Foram levados em consideração fatores meteorológicos medidos pela REDEMET e informações de demanda e capacidade das aeronaves e aeroportos disponibilizados pela ANAC entre os dias 06/01/2020 e 04/06/2022.

Iniciou-se o estudo calculando o indicador de pontualidade de partida descrito pelo DECEA e analisando para aeroportos e companhias aéreas com decolagens no território brasileiro. Na sequência através do método Aumento de Gradiente Extremo construiu-se um modelo de aprendizado de máquina supervisionado com o intuito de prever a pontualidade de partida voos e obteve-se acurácia média de 92,96%.

Após o processamento do modelo, analisou-se a relevância das variáveis de entrada com o objetivo de selecionar os fatores com maior importância sobre a previsão da pontualidade e teve-se como resultado a pressão local, o número de voos que decolaram na mesma hora e o volume de bagagens despachada pelos passageiros como principais fatores. Diante disso, ressalta-se que tanto fatores meteorológicos como operacionais influenciam na performance de decolagem da aeronave.

Como sugestão para amplificar o estudo em trabalhos futuros, recomenda-se a inserção de novos indicadores de performance descritos pelo DECEA de forma a analisar de forma completa diferentes áreas dos sítios aeroportuários. Além disso, no modelo de aprendizado de máquina sugere-se realizar o estudo fazendo-se o balanceamento da base de dados através de método sob amostra (*under – sample*). Ademais, sugere-se inserir no estudo aeroportos de outros países de forma a comparar a eficiência operacional dos aeroportos brasileiros com os de outros países.

Referências

- ABU-RMILEH, A. The multiple faces of ‘feature importance’ in xgboost. 2019. Available at: <https://xgboost.readthedocs.io/en/stable/parameter.html>. Accessed on: 01 ago. 2022.
- AEROFLAP. Gru airport recebe investimento de mais de r\$ 100 milhões para expandir seu terminal de cargas. **AEROFLAP**, 2022. Available at: <https://www.aeroflap.com.br/gru-airport-recebe-investimento-de-mais-de-r-100-milhoes-para-expandir-seu-terminal-de-cargas/>. Accessed on: 18 mai. 2022.
- ANAC. 2022. Available at: <https://sas.anac.gov.br/sas/bav/view/frmConsultaVRA>. Accessed on: 25 set. 2022.
- BRASIL., M. da Defesa. Comando da Aeronáutica. Departamento de Controle do E. A. Mca 100-22: Metodologia de indicadores atm do sisceab. **ICAO**, 2020. Available at: <https://www.icao.int/SAM/Documents/2021-RLA06901-TallerVOLIII-1-/BRASIL\%20\%20MCA\%20100-22.pdf>. Accessed on: 19 mai. 2022.
- BUBALO, B. Airport capacity and performance in europe: a study of transport economics, service quality and sustainability. Dissertation (Doctor rerum oeconomicarum Universitat Hamburg. Hamburg, p. 98, 2020. Available at: <https://arxiv.org/pdf/2102.02379.pdf>. Accessed on: 25 set. 2022.
- DALMAU et. al. Improving the predictability of take-off times with machine learning. 2019. Available at: https://upcommons.upc.edu/bitstream/handle/2117/181173/SIDs_2019_paper_36.pdf. Accessed on: 25 set. 2022.
- DATA. 2020. Available at: <https://datascience.eu/pt/aprendizado-de-maquina/matriz-de-confusao/>. Accessed on: 01 ago. 2022.
- FAYOL, H. Administration industrielle et gÉnÉrale. Paris: Luc Marco et Edi-Gestion,, 1916. Available at: <https://ihpm.hypotheses.org/files/2016/09/Fayol-1916.pdf>. Accessed on: 24 set. 2022.
- FREITAS, O. S. A. L. Machine learning: Desafios para um brasil competitivo. Revista da Sociedade Brasileira de Computação, v.39, n.1, p. 7-10, 2019. Available at: https://www.sbc.org.br/images/flippingbook/computacaobrasil/computa_39/pdf-/CompBrasil_39_180.pdf. Accessed on: 25 set. 2022.

- GIL, A. C. Como elaborar projetos de pesquisa. são paulo, atlas, 2002. 2002. Available at: https://files.cercomp.ufg.br/weby/up/150/o-/Anexo_C1_como_elaborar_projeto_de_pesquisa_-_antonio_carlos_gil.pdf. Accessed on: 18 mai. 2022.
- GOMES, P. C. T. Conheça o algoritmo xgboost. 2019. Available at: <https://www.datageeks.com.br/xgboost/>. Accessed on: 01 ago. 2022.
- IATA. O valor do transporte aéreo no brasil. **IATA Economics**, p. 1–6, 2019. Available at: <https://www.iata.org/contentassets/bc041f5b6b96476a80db109f220f8904/brazil-o-valor-do-transporte-aereo.pdf>. Accessed on: 18 mai. 2022.
- METODOS. Metodos baseados em árvores. 2020. Available at: <http://cursos.leg.ufpr.br/ML4all/apoio/Arvores.html>. Accessed on: 01 ago. 2022.
- NEVES, M. A. O. Tudo sobre indicadores de desempenho em logística. *In: _____*. Maringá: Revista Mundo Logística, set/out, 2009. p. 31–45. Accessed on: 01 mai. 2022.
- NUMERO. Numero de passageiros aéreos cresceu 210% de 2000 a 2014, diz cnt. **Correio Braziliense**, 2015. Available at: https://www.correiobraziliense.com.br/app/noticia-/turismo/2015/11/11/interna_turismo,506012/numero-de-passageiros-aereos-cresceu-210-de-2000-a-2014-diz-cnt.shtml. Accessed on: 18 mai. 2022.
- REDEMET. 2022. Available at: <https://www.redemet.aer.mil.br>. Accessed on: 25 set. 2022.
- SCHULTZ, M. et.al. Weather impact on airport performance. *Aerospace*, v.5, n.4, article number 109, 2018. Available at: <https://www.mdpi.com/2226-4310/5/4/109/htm>. Accessed on: 25 set. 2022.
- SILVA, J. 2020. Available at: <https://medium.com/equals-lab/uma-breve-introdução-ao-algoritmo-de-machine-learning-gradient-boosting-utilizando-a-biblioteca-311285783099>. Accessed on: 01 ago. 2022.
- STERNBERG, A. et. al. A review on flight delay prediction. 2017. Available at: <https://arxiv.org/abs/1703.06118>. Accessed on: 25 set. 2022.
- TEZZA, R. et. al. Sistemas de medição de desempenho: uma revisão e classificação da literatura. *Gest. Prod., São Carlos*, v. 17, n.1, p. 75-93, 2010. Available at: https://www.researchgate.net/publication-/250984346_Sistemas_de_medicao_de_desempenho_uma_revisao_e_classificacao_da_literatura. Accessed on: 24 set. 2022.
- XGBOOST. 2022. Available at: <https://xgboost.readthedocs.io/en/stable/parameter.html>. Accessed on: 01 ago. 2022.
- XGBOOST. 2022. Available at: https://xgboost.readthedocs.io/en/latest/python/python_api.html#xgboost.Booster.get_score. Accessed on: 01 ago. 2022.
- YOSIMOTO, V. et. al. A lógica atual do setor aeroportuário brasileiro. 2016. Available at: https://web.bndes.gov.br/bib/jspui/bitstream/1408/9394/2/7%20-%20A%20lógica%20atual%20do%20setor%20aeroportuário%20brasileiro_P.pdf. Accessed on: 01 jun. 2022.

FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO <p style="text-align: center;">TC</p>	2. DATA <p style="text-align: center;">25 de novembro de 2022</p>	3. DOCUMENTO Nº <p style="text-align: center;">DCTA/ITA/TC-114/2022</p>	4. Nº DE PÁGINAS <p style="text-align: center;">66</p>
5. TÍTULO E SUBTÍTULO: Análise e previsão de indicadores de performance do tráfego aéreo brasileiro			
6. AUTOR(ES): Alexandre de Castro Fernandes			
7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): Instituto Tecnológico de Aeronáutica – ITA			
8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: Aeroporto; KPI; Performance			
9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: Aeroporto; Desempenho; Transporte de Passageiros; Transporte Aéreo; Transportes			
10. APRESENTAÇÃO: <input checked="" type="checkbox"/> Nacional <input type="checkbox"/> Internacional ITA, São José dos Campos. Curso de Graduação em Engenharia Civil-Aeronáutica. Orientador: Marcelo Xavier Guterres. Publicado em: 2022			
11. RESUMO: <p> O transporte aéreo brasileiro tem apresentado crescimento em sua demanda de forma que a medição do desempenho da operação de gestores de aeroportos e companhias aéreas cumpre papel fundamental na manutenção da qualidade do serviço prestado aos passageiros. O presente estudo busca analisar indicadores de <i>performance</i> do setor aeroportuário de forma a obter os principais fatores que influenciam na eficiência da operação. A pesquisa é feita levando-se em consideração os voos com decolagem no território brasileiro e insere-se no estudo informações detalhadas do voo como condições meteorológicas do local no momento da partida e dados sobre demanda e capacidade da respectiva aeronave e aeroporto. Com essas informações primeiramente visualizou-se o indicador de pontualidade de partida (KPI 01) levando-se em consideração quatro categorias na classificação. A partir disso, utilizou-se o método de Aumento de Gradiente Extremo (XGBOOST) como algoritmo de aprendizado de máquina supervisionado com o objetivo de prever a pontualidade de partida de um voo. Os resultados obtidos mostram que a pressão local, a quantidade de decolagens na mesma hora que a do respectivo voo e o volume de bagagem franqueada são os principais fatores que influenciam na pontualidade de partida. Por fim, analisou-se a eficiência na gestão da demanda e capacidade horárias dos aeroportos. </p>			
12. GRAU DE SIGILO: <p style="text-align: center;"> <input checked="" type="checkbox"/> OSTENSIVO <input type="checkbox"/> RESERVADO <input type="checkbox"/> SECRETO </p>			