

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS - CAMPUS BAURU
DEPARTAMENTO DE COMPUTAÇÃO
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

ANA BEATRIZ MARTINS RIBEIRO

**PORTOVISION – SISTEMA DE ANÁLISE E PREDIÇÃO PARA OTIMIZAÇÃO DO
TRANSPORTE MARÍTIMO BRASILEIRO BASEADO EM DADOS DA ANTAQ**

Bauru
2025

UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”

FACULDADE DE CIÊNCIAS - CAMPUS BAURU

DEPARTAMENTO DE COMPUTAÇÃO

BACHARELADO EM SISTEMAS DE INFORMAÇÃO

ANA BEATRIZ MARTINS RIBEIRO

**PORTOVISION – SISTEMA DE ANÁLISE E PREDIÇÃO PARA OTIMIZAÇÃO DO
TRANSPORTE MARÍTIMO BRASILEIRO BASEADO EM DADOS DA ANTAQ**

Trabalho apresentado como exigência para a
conclusão do Curso de Bacharelado em Sistemas
de Informação da Faculdade de Ciências –
UNESP Campus Bauru.

Orientador: Prof. Dr. Higor Amario de Souza

Bauru

2025

R484p	RIBEIRO, ANA B. M. PORTOVISION – SISTEMA DE ANÁLISE E PREDIÇÃO PARA OTIMIZAÇÃO DO TRANSPORTE MARÍTIMO BRASILEIRO / ANA B. M. RIBEIRO. -- Bauru, 2025 59 p. : tabs., fotos Trabalho de conclusão de curso (Bacharelado - Sistemas de Informação) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências, Bauru Orientador: Higor Amario de Souza 1. logística marítima. 2. sistema de apoio à decisão. 3. aprendizado de máquina. 4. gestão portuária. 5. análise de dados. I. Título.
-------	--

ANA BEATRIZ MARTINS RIBEIRO

**PORTOVISION – SISTEMA DE ANÁLISE E PREDIÇÃO PARA OTIMIZAÇÃO DO
TRANSPORTE MARÍTIMO BRASILEIRO BASEADO EM DADOS DA ANTAQ**

Trabalho de Conclusão de Curso do Curso
apresentada à Universidade Estadual Paulista
(UNESP), Faculdade de Ciências, Bauru, para
obtenção do título de Bacharel em Sistemas de
Informação.

Banca Examinadora:

Prof. Dr. Higor Amario de Souza
USP - Escola Politécnica - Cidade Universitária

Prof. Dr. João Pedro Albino
UNESP - Faculdade de Ciências - Campus de Bauru

Prof. Dr. José Remo Ferreira Brega
UNESP - Faculdade de Ciências - Campus de Bauru

Bauru, _____ de _____ de _____

Dedico este trabalho à minha família, pois é graças ao esforço deles que hoje posso concluir o meu curso.

AGRADECIMENTOS

Primeiramente, agradeço aos meus pais e à minha família por todo o amor, apoio, paciência e confiança. Sem o apoio incondicional de vocês, nada disso seria possível.

Agradeço também ao meu orientador, pelo direcionamento e disponibilidade durante o desenvolvimento deste trabalho, sua ajuda foi essencial para a definição do tema. Estendo meus agradecimentos a todos os professores da faculdade, que ao longo destes anos compartilharam seus conhecimentos, experiências e ensinamentos.

Ao meu namorado, agradeço imensamente pelo companheirismo, paciência, incentivo constante e, principalmente, por acreditar em mim, mesmo nos momentos mais desafiadores.

Por fim, agradeço a todas as pessoas que, de alguma forma, cruzaram o meu caminho durante todos esses anos de universidade. Em especial, aos meus amigos, por estarem sempre ao meu lado, oferecendo apoio, amizade e incentivo, tornando essa jornada mais leve e cheia de boas lembranças.

RESUMO

O modal marítimo desempenha papel importante no transporte de cargas em grandes distâncias, sendo crucial para a economia global. No entanto, fatores como congestionamentos portuários, variações na produtividade de berços, paralisações e incertezas operacionais impactam diretamente os custos e prazos de entrega. Com o avanço da tecnologia, análise de dados e inteligência artificial, surgem oportunidades para otimizar processos, realizar análises e gerar previsões interessantes que apoiem decisões mais assertivas e melhorem o planejamento logístico. Este trabalho apresenta o desenvolvimento de um sistema de apoio à decisão voltado à logística marítima portuária brasileira, embasado em dados da ANTAQ. Foram realizadas análises exploratórias que possibilitaram a identificação de padrões, tendências e sazonalidades, contribuindo para uma melhor compreensão do comportamento das operações portuárias ao longo dos anos. O objetivo principal é transformar dados brutos em insights estratégicos, melhorando a eficiência, reduzindo incertezas, orientando decisões e investimentos logísticos em países como o Brasil. A partir das análises realizadas, bem como do dashboard e do modelo preditivo desenvolvidos, espera-se que o trabalho contribua para o aprimoramento da gestão e do planejamento portuário, fortalecendo a competitividade do transporte marítimo nacional.

Palavras-chave: logística marítima; sistema de apoio à decisão; aprendizado de máquina; gestão portuária; análise de dados.

ABSTRACT

Maritime transport plays an important role in the transport of cargo over long distances and is crucial to the global economy. However, factors such as port congestion, variations in berth productivity, stoppages, and operational uncertainties directly impact costs and delivery times. With advances in technology, data analysis, and artificial intelligence, opportunities arise to optimize processes, perform analyses, and generate interesting forecasts that support more assertive decisions and improve logistics planning. This work presents the development of a decision support system focused on Brazilian maritime port logistics, based on data from ANTAQ. Exploratory analyses were carried out that enabled the identification of patterns, trends, and seasonality, contributing to a better understanding of the behavior of port operations over the years. The main objective is to transform raw data into strategic insights, improving efficiency, reducing uncertainties, and guiding logistics decisions and investments in countries such as Brazil. Based on the analyses performed, as well as the dashboard and predictive model developed, it is expected that this work will contribute to the improvement of port management and planning, strengthening the competitiveness of national maritime transport.

Keywords: maritime logistics; decision support system; machine learning; port management; data analysis.

LISTA DE FIGURAS

Figura 1 - Tipos de tarefas de aprendizado de máquina	6
Figura 2 - Visualização da plataforma MarineTraffic	10
Figura 3 - Visualização da plataforma seaexplorer	10
Figura 4 - Visualização do painel estatístico aquaviário	11
Figura 5 – Diagrama de sequência do sistema de análise e predição portuária	12
Figura 6 - Diagrama de bloco do aprendizado de máquina	13
Figura 7 - Modelo de dados ANTAQ	14
Figura 8 - Interface PortoVision	23
Figura 9 - Evolução anual de atracação nos portos brasileiros	25
Figura 10 - Mapa de calor de atracações por mês e ano.	27
Figura 11 - Evolução dos top 10 portos mais movimentados	29
Figura 12 - Tela dashboard PortoVision	32
Figura 13 - Tela predição PortoVision	34
Figura 14 - Tela inicial PortoVision	37
Figura 15 - Tela inicial PortoVision Mobile	37

LISTA DE TABELAS

Tabela 1 - Distribuição de registros de atracções por ano.	25
Tabela 2 - Distribuição de atracções por região geográfica (2010-2024)	28
Tabela 3 - Distribuição de atracções por tipo de operação (2010-2024)	30
Tabela 4 - Distribuição das atracções por tipo de navegação (2010-2024)	31

LISTA DE ABREVIATURAS E SIGLAS

ANTAQ	Agência Nacional de Transportes Aquaviários
API	Application Programming Interface (Interface de Programação de Aplicações)
BNDES	Banco Nacional de Desenvolvimento Econômico e Social
CENTRONAVE	Centro Nacional de Navegação Transatlântica
DSS	Decision Support Systems
EAD	Análise Exploratória de Dados
EDA	Exploratory Data Analysis
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
KPI	Key Performance Indicator (Indicador-Chave de Desempenho)
LGPD	Lei Geral de Proteção de Dados Pessoais
MAE	Mean Absolute Error (Erro Médio Absoluto)
MAPE	Mean Absolute Percentage Error (Erro Absoluto Percentual Médio)
ML	Machine Learning (Aprendizado de Máquina)
PAC	Programa de Aceleração do Crescimento
PIB	Produto Interno Bruto
RMSE	Root Mean Squared Error (Raiz do Erro Quadrático Médio)
R^2	R-quadrado ou Coeficiente de Determinação
SAD	Sistema de Apoio à Decisão
TUPs	Terminais de Uso Privado
UNCTAD	Conferência das Nações Unidas sobre Comércio e Desenvolvimento

SUMÁRIO

1 INTRODUÇÃO	1
1.1 OBJETIVOS	2
1.1.1 OBJETIVO GERAL	2
1.1.2 OBJETIVOS ESPECÍFICOS	2
1.2 MOTIVAÇÃO	3
2 FUNDAMENTAÇÃO TEÓRICA	4
2.1 LOGÍSTICA MARÍTIMA E TRANSPORTE PORTUÁRIO	4
2.2 SISTEMAS DE APOIO À DECISÃO	5
2.3 APRENDIZADO DE MÁQUINA APLICADO À LOGÍSTICA	5
2.4 ANÁLISE DE DADOS PORTUÁRIOS	7
2.4.1 PRINCIPAIS INDICADORES DE DESEMPENHO PORTUÁRIO	7
2.4.2 RELEVÂNCIA DA ANÁLISE DE DADOS PARA A GESTÃO PORTUÁRIA	8
2.5 BASE DE DADOS DA ANTAQ	8
2.5.1 RELEVÂNCIA PARA O DESENVOLVIMENTO DO TRABALHO	8
2.6 SOLUÇÕES EXISTENTES	9
3 METODOLOGIA E DESENVOLVIMENTO	12
3.1 ARQUITETURA DO SISTEMA	12
3.2 FERRAMENTAS E TECNOLOGIAS UTILIZADAS	15
3.2.1 COLETA E PRÉ-PROCESSAMENTO DOS DADOS	15
3.2.2 ANÁLISE EXPLORATÓRIA DOS DADOS	16
3.2.3 DESENVOLVIMENTO DO MODELO PREDITIVO	18
3.2.4 DESENVOLVIMENTO DO DASHBOARD	19
3.2.5 DESENVOLVIMENTO DO WEBSITE	20
3.4 PROCESSOS DE VALIDAÇÃO	21
4 RESULTADOS	23
4.1 FUNCIONALIDADES DO SISTEMA	23
4.2 ANÁLISE EXPLORATÓRIA	23
4.2.1 PADRÕES E TENDÊNCIAS IDENTIFICADOS	24
4.3 DASHBOARD DESENVOLVIDO	32
4.4 MODELO PREDITIVO	33
4.4.1 TREINAMENTO E VALIDAÇÃO	34
4.4.2 MÉTRICAS DE DESEMPENHO	34
4.5 WEBSITE	36
5 CONCLUSÃO	38
5.1 TRABALHOS FUTUROS	39
REFERÊNCIAS	40

1 INTRODUÇÃO

O comércio internacional é um dos pilares centrais da economia global, movimentando todos os dias milhões de toneladas de mercadorias pelos oceanos. Dentro desse cenário, o transporte marítimo exerce papel estratégico, responsável por mais de 80% do volume total de comércio no mundo, segundo dados da Conferência das Nações Unidas sobre Comércio e Desenvolvimento (UNCTAD, 2023). Mais do que simplesmente levar cargas de um ponto a outro, o setor conecta mercados, impulsiona cadeias produtivas e influencia diretamente os custos logísticos. Além do comércio entre países, também tem grande relevância dentro do próprio território nacional, especialmente por meio da cabotagem e da navegação interior.

No Brasil, o setor portuário é um dos principais e mais relevantes modais, sendo responsáveis por grande parte da movimentação de produtos, considerando importação e exportação e, conseqüentemente, tendo impacto na economia do país. Em 2024, os portos brasileiros movimentaram 1,32 bilhão de toneladas, de acordo com dados da ANTAQ (MINISTÉRIO DOS PORTOS E AEROPORTOS, 2025), estabelecendo o maior volume da série histórica e sinalizando uma tendência de crescimento contínuo para os próximos anos.

Essa expansão portuária, associada a investimentos expressivos, como os R\$54,9 bilhões previstos até 2030 pelo novo PAC para o setor portuário brasileiro, exige soluções modernas e eficientes para lidar com desafios estruturais, operacionais e ambientais que ainda limitam a competitividade do setor. Tais desafios envolvem desde a infraestrutura física deficiente, passando pelo atraso nos processos logísticos, até a baixa integração de sistemas inteligentes de gestão portuária.

Nesse cenário, a adoção de novas tecnologias tem se mostrado fundamental para impulsionar a transformação do setor portuário. O uso de ferramentas de análise e predição de dados ajuda a identificar gargalos, prever períodos de maior demanda e orientar decisões importantes sobre movimentação de cargas, uso de berços e gestão da cadeia logística. A aplicação de técnicas de aprendizado de máquina e inteligência artificial representa, portanto, uma oportunidade de modernização, capaz de reduzir custos operacionais, diminuir incertezas e fortalecer a competitividade dos portos nacionais no comércio internacional.

Diante desse panorama, este trabalho apresenta o desenvolvimento de um sistema de análise e predição voltado à otimização do transporte marítimo brasileiro, fundamentado em dados estatísticos disponibilizados pela ANTAQ. Inicialmente, realizou-se uma análise exploratória abrangente, que possibilitou a identificação de padrões, tendências e sazonalidades nas operações portuárias ao longo dos anos, contribuindo para uma compreensão mais profunda do comportamento do setor. Em seguida, desenvolveu-se um modelo preditivo e um dashboard interativo, que permite a visualização de indicadores operacionais e a previsão de tempos de estadia portuária. A proposta visa explorar o potencial analítico dos dados do setor aquaviário para fornecer insights estratégicos que apoiem a gestão portuária.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Desenvolver um sistema web de apoio à decisão para o setor portuário brasileiro que, utilizando técnicas de machine learning e análise de dados da ANTAQ, permita prever o tempo de estadia de embarcações e oferecer através de um dashboard indicadores mais voltados a performance, a fim de otimizar processos logísticos e subsidiar decisões estratégicas no transporte marítimo.

1.1.2 OBJETIVOS ESPECÍFICOS

1. Aplicar algoritmos de aprendizado supervisionado para prever o tempo de estadia de embarcações com base em dados operacionais portuários;
2. Analisar variáveis como tempos operacionais, ocupação de berços, tipos de carga e características das embarcações;
3. Desenvolver uma aplicação web interativa com visualizações gráficas e formulários para simulação de predições; e
4. Promover o uso de dados abertos e inteligência artificial como ferramentas de modernização e eficiência no setor portuário brasileiro.

1.2 MOTIVAÇÃO

O setor portuário enfrenta desafios estruturais e operacionais que comprometem sua eficiência, como gargalos logísticos, baixa produtividade, processos burocráticos e infraestrutura limitada (MOLITZAS, 2024). A crescente disponibilidade de dados públicos fornecidos pela ANTAQ abre novas oportunidades para o desenvolvimento de ferramentas analíticas e modelos preditivos capazes de apoiar a tomada de decisão. A transformação desses dados em informações estratégicas pode contribuir significativamente para reduzir tempos de espera, melhorar o planejamento das atracções e aumentar a eficiência operacional dos portos brasileiros.

A motivação para o desenvolvimento deste trabalho surge, portanto, da necessidade de integrar ciência de dados e gestão portuária, oferecendo uma solução que permite uma análise exploratória, visualização interativa e predição de tempo de estadia de embarcações. O sistema proposto, denominado **PortoVision**, busca não apenas demonstrar o potencial do uso de dados abertos da ANTAQ, mas também fomentar uma cultura de inovação e inteligência operacional no setor portuário brasileiro.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais referenciais teóricos que embasam o desenvolvimento do estudo, reunindo conceitos fundamentais sobre logística marítima, sistemas de apoio à decisão, aprendizado de máquina e análise de dados portuários.

2.1 LOGÍSTICA MARÍTIMA E TRANSPORTE PORTUÁRIO

Segundo Stopford (2017), o transporte marítimo distingue-se pela capacidade de movimentar grandes volumes de carga a custos relativamente baixos por unidade transportada, tornando-se indispensável para o funcionamento das cadeias de suprimentos globalizadas. No contexto brasileiro, o modal aquaviário assume relevância estratégica devido à extensa costa marítima de mais de 7.000 km e à participação expressiva no escoamento da produção nacional.

Os portos desempenham um papel importante dentro das redes logísticas, atuando como pontos de conexão entre o transporte marítimo e o terrestre. Conforme destacam Notteboom e Rodrigue (2014), a eficiência das operações portuárias tem influência direta sobre a competitividade dos países no comércio internacional, já que problemas como atrasos, congestionamentos e altos custos acabam afetando toda a cadeia logística. A gestão portuária vai muito além da infraestrutura física, ela depende também de processos bem estruturados e do uso de sistemas de informação integrados.

Os principais desafios enfrentados pelo setor portuário incluem: (i) congestionamentos e filas de espera para atracação; (ii) baixa produtividade na movimentação de cargas; (iii) infraestrutura física deficiente ou saturada; (iv) processos burocráticos complexos; e (v) limitada integração entre os diversos atores da cadeia logística (QUINTELLA, 2024). Esses fatores contribuem para o aumento dos custos logísticos totais e comprometem a competitividade econômica.

A superação desses desafios exige investimentos consistentes em infraestrutura, modernização tecnológica e a adoção de modelos de gestão mais inteligentes.

2.2 SISTEMAS DE APOIO À DECISÃO

Os Sistemas de Apoio à Decisão (SAD), do inglês *Decision Support Systems* (DSS), são ferramentas computacionais desenvolvidas para auxiliar gestores e tomadores de decisão na análise de problemas complexos e na escolha de alternativas estratégicas (POWER, 2002).

Na logística portuária, os SADs têm ganhado cada vez mais espaço como aliados na gestão e no planejamento das operações. Eles ajudam em tarefas como a alocação de berços, a previsão de demanda, o controle de estoques, o planejamento de dragagens, o dimensionamento de frotas e a otimização de rotas. Diante da complexidade das atividades portuárias, que envolvem muitos fatores interligados e um alto nível de incerteza, esses sistemas se tornam ferramentas essenciais para apoiar tanto o planejamento estratégico quanto às decisões do dia a dia.

Para Druzdzel e Flynn (2002), os SADs não substituem o julgamento humano, mas o ampliam ao fornecer informações relevantes, análises fundamentadas e projeções de cenários alternativos. Essa característica é particularmente importante em ambientes dinâmicos como os portos, onde cada decisão precisa levar em conta diversos fatores ao mesmo tempo e pode impactar diretamente o andamento das operações.

2.3 APRENDIZADO DE MÁQUINA APLICADO À LOGÍSTICA

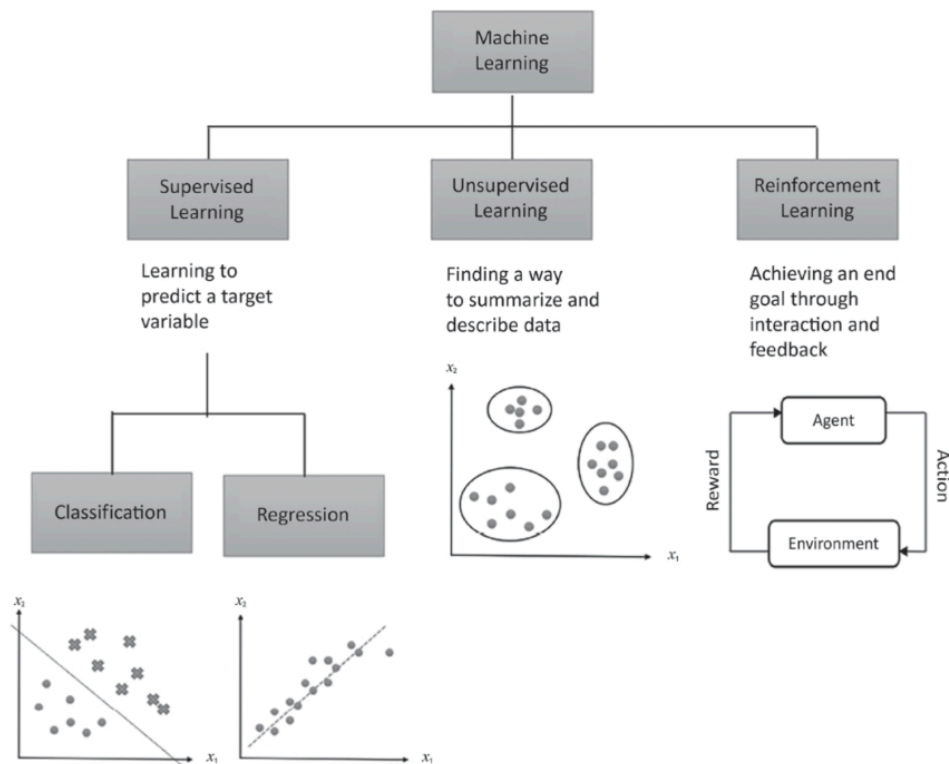
O aprendizado de máquina, ou machine learning, reúne um conjunto de técnicas que permitem aos sistemas identificar padrões e aprimorar seu desempenho com base nos dados que recebem, sem a necessidade de serem programados diretamente para cada tarefa (MITCHELL, 1997). Trata-se de uma área da inteligência artificial voltada ao desenvolvimento de algoritmos capazes de aprender com a experiência, tornando-se mais eficientes à medida que são expostos a novas informações.

De acordo com Liu (2025), os algoritmos de aprendizado de máquina podem ser divididos em três grandes categorias: (i) **aprendizado supervisionado**, em que o modelo é treinado com dados rotulados, ou seja, entradas acompanhadas de saídas conhecidas, para aprender a fazer previsões ou classificações em novos conjuntos de dados; (ii) **aprendizado não supervisionado**, que trabalha com dados sem rótulos e busca identificar padrões, agrupar informações semelhantes e revelar estruturas ocultas nos dados; e (iii) **aprendizado por**

reforço no qual o sistema aprende por meio de tentativa e erro, recebendo recompensas ou penalidades, até atingir o melhor desempenho possível.

Essas três categorias são ilustradas na Figura 1, que apresenta os principais tipos de tarefas de aprendizado de máquina.

Figura 1 - Tipos de tarefas de aprendizado de máquina.



Fonte: LIU (2025).

O avanço das tecnologias de informação e o crescimento exponencial na geração de dados têm impulsionado a transformação digital em diversos setores da economia, e a logística se destaca entre as áreas que mais se beneficiam desse movimento. Com cadeias de suprimentos cada vez mais complexas, dinâmicas e interconectadas, as aplicações de aprendizado de máquina surgem como uma ferramentas estratégicas para otimizar processos, reduzir custos e aprimorar a tomada de decisão em tempo real.

2.4 ANÁLISE DE DADOS PORTUÁRIOS

2.4.1 PRINCIPAIS INDICADORES DE DESEMPENHO PORTUÁRIO

Os indicadores de desempenho portuário são ferramentas essenciais para avaliar a eficiência das operações, identificar gargalos e orientar na tomada de decisões. A Conferência das Nações Unidas sobre Comércio e Desenvolvimento (UNCTAD, 2023) ressalta que o monitoramento desses indicadores permite compreender o desempenho global de um porto, combinando dimensões operacionais, financeiras, ambientais e institucionais.

Entre os principais indicadores operacionais tradicionalmente utilizados, destacam-se:

1. **Prancha Média (Taxa de Transferência):** Considerado um indicador relevante, pois mede a velocidade de movimentação de carga no terminal, expressa em toneladas por hora ou contêineres por hora. Valores elevados indicam maior produtividade e eficiência operacional.
2. **Tempo de Espera para Atracação:** Indica quanto tempo uma embarcação fica na fila aguardando vaga no berço. Tempos elevados sinalizam problemas na infraestrutura ou na forma como a capacidade está sendo gerenciada.
3. **Tempo de Atendimento (Operação):** Corresponde ao período em que o navio fica atracado realizando as operações de carga e descarga. Operações mais rápidas significam maior giro de berços e melhor uso da estrutura disponível.
4. **Tempo de Estadia Total:** Representa a soma do tempo de espera para atracação e do tempo de operação, indicando o período total de permanência da embarcação na área portuária.
5. **Nível de Serviço:** Definido pela relação entre o tempo de espera e o tempo de operação, se refere a percepção de qualidade experimentada pelos clientes do terminal portuário.
6. **Taxa de Ocupação de Berços:** Mede o percentual de tempo em que o berço está ocupado em relação ao tempo total disponível. Taxas elevadas podem indicar saturação da infraestrutura, enquanto taxas muito baixas sugerem ociosidade e subutilização de recursos.

Esses indicadores formam a base da avaliação operacional tradicional e continuam sendo amplamente utilizados por autoridades portuárias e órgãos reguladores, incluindo a

ANTAQ, no monitoramento dos terminais brasileiros. Nos relatórios mais recentes da UNCTAD (2023), é proposto também uma visão mais ampla do desempenho portuário, ultrapassando a dimensão puramente operacional. Essa abordagem incorpora sete categorias de análise: governança, recursos humanos, finanças, operações de embarcações, resiliência, operações de carga e sustentabilidade ambiental.

2.4.2 RELEVÂNCIA DA ANÁLISE DE DADOS PARA A GESTÃO PORTUÁRIA

A análise de dados portuários permite não apenas a identificação de ineficiências operacionais, mas também a formulação de estratégias baseadas em evidências para a melhoria contínua. Nesse sentido, a exploração analítica dos dados da ANTAQ representa uma oportunidade valiosa para gerar insights estratégicos que apoiem tanto o setor público quanto o privado na busca por maior eficiência operacional e sustentabilidade econômica.

2.5 BASE DE DADOS DA ANTAQ

A Agência Nacional de Transportes Aquaviários, criada pela Lei nº 10.233/2001, desempenha papel fundamental como agência reguladora do setor aquaviário e portuário brasileiro. Vinculada ao Ministério de Portos e Aeroportos, é responsável pela regulação, supervisão e fiscalização das atividades relacionadas à prestação de serviços de transporte aquaviário e à exploração da infraestrutura portuária. Dentre suas atribuições, destaca-se como uma das principais provedoras de dados públicos sobre o desempenho do setor, disponibilizando informações essenciais para análises estratégicas e planejamento logístico (ANTAQ, 2020).

2.5.1 RELEVÂNCIA PARA O DESENVOLVIMENTO DO TRABALHO

Embora a ANTAQ seja uma fonte oficial e confiável, os dados que ela disponibiliza nem sempre chegam prontos para uso. Como ocorre em grandes bases operacionais, é comum encontrar informações incompletas, registros duplicados, campos em branco e falta de padronização entre diferentes períodos ou portos.

Segundo Dasu e Johnson (2003, apud WICKHAM, 2014), o processo de preparação de dados pode consumir até 80% do tempo total de um projeto de análise, especialmente quando se trabalha com dados reais provenientes de múltiplas fontes e sistemas. No caso do setor portuário, esse processo se torna ainda mais desafiador por causa da diversidade de formatos entre os períodos, das mudanças na estrutura das tabelas ao longo dos anos e da necessidade de cruzar informações de naturezas distintas para que as análises ganhem profundidade e contexto.

Mesmo com os desafios mencionados, os dados disponibilizados pela ANTAQ representam um recurso valioso para quem busca entender sobre logística portuária. A amplitude e o nível de detalhe das informações permitem observar padrões de operação, acompanhar tendências ao longo do tempo, identificar variações sazonais e construir modelos preditivos capazes de antecipar comportamentos do sistema portuário, oferecendo suporte concreto para tomada de decisões.

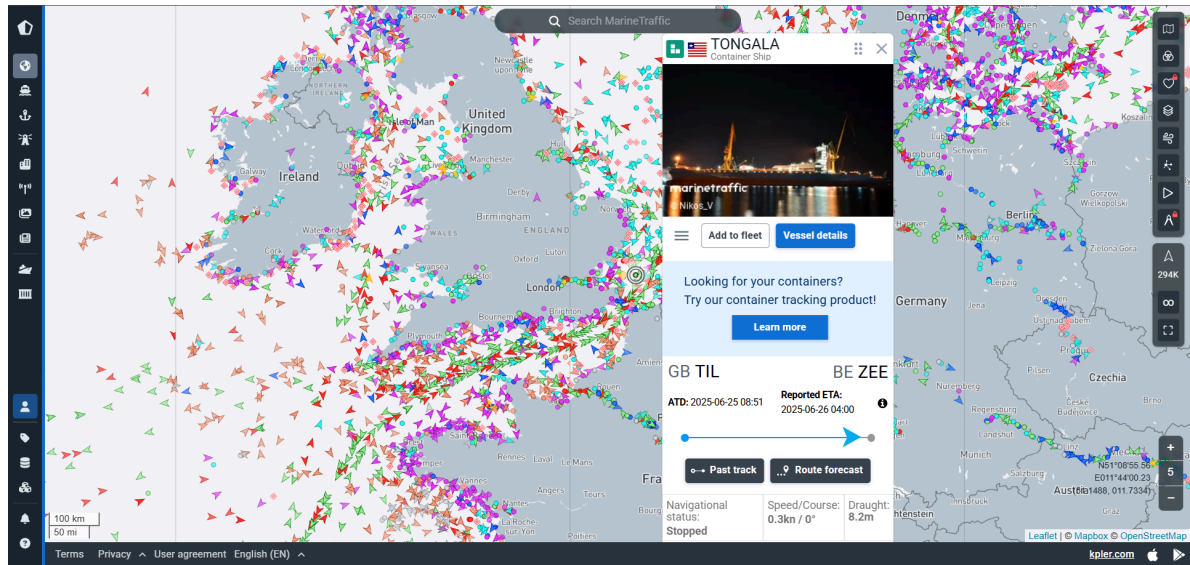
Este trabalho utiliza os dados históricos da ANTAQ como fundamento para as análises exploratórias e para o treinamento de um modelo de aprendizado de máquina, visando transformar dados brutos em insights estratégicos que contribuam para a melhoria da eficiência e competitividade do setor portuário brasileiro.

2.6 SOLUÇÕES EXISTENTES

Durante a etapa de levantamento e análise, observou-se a existência de soluções voltadas ao uso de dados e inteligência artificial no setor marítimo internacional. Essas ferramentas, ainda que com objetivos variados, demonstram o potencial da tecnologia para transformar a logística portuária e serviram como referência para o desenvolvimento deste trabalho.

Internacionalmente, destacam-se plataformas como o MarineTraffic, que fornece monitoramento em tempo real de embarcações e estimativas de chegada com base em rotas e histórico de navegação. Embora amplamente utilizado, o foco da ferramenta é voltado à visualização de tráfego marítimo, e não à análise operacional portuária, como mostra a Figura 2.

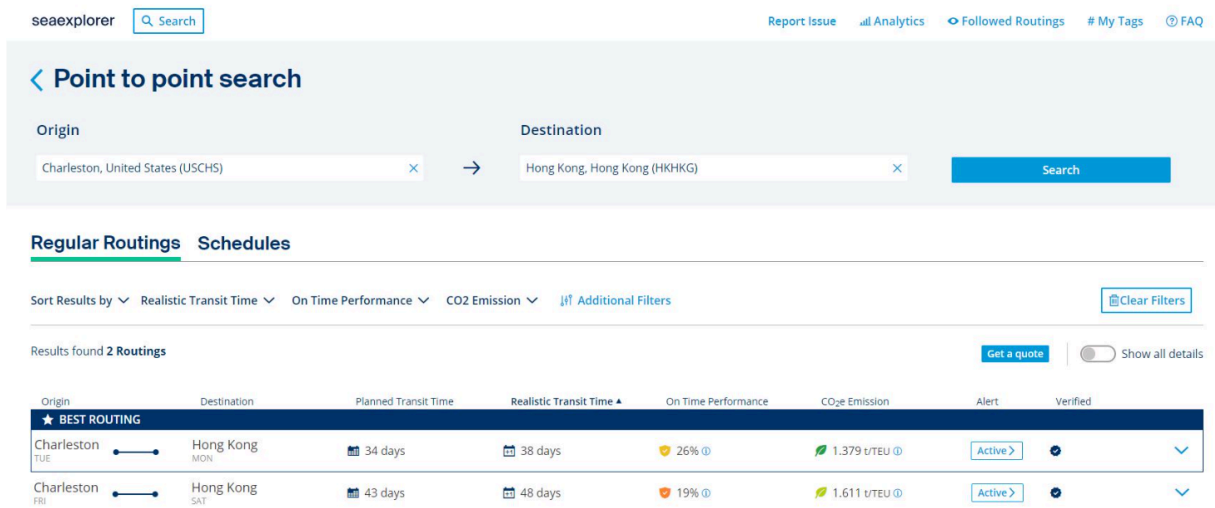
Figura 2 - Visualização da plataforma MarineTraffic.



Fonte: MarineTraffic, 2025.¹

Outra iniciativa de referência é o sistema desenvolvido por empresas de logística como a Kuehne+Nagel, que vêm adotando modelos preditivos para atrasos e planejamento de cadeia logística, baseando-se em múltiplas fontes de dados e IA, conforme ilustrado na Figura 3. Contudo, essas soluções são, em sua maioria, proprietárias e voltadas ao mercado privado global.

Figura 3 - Visualização da plataforma seaexplorer.



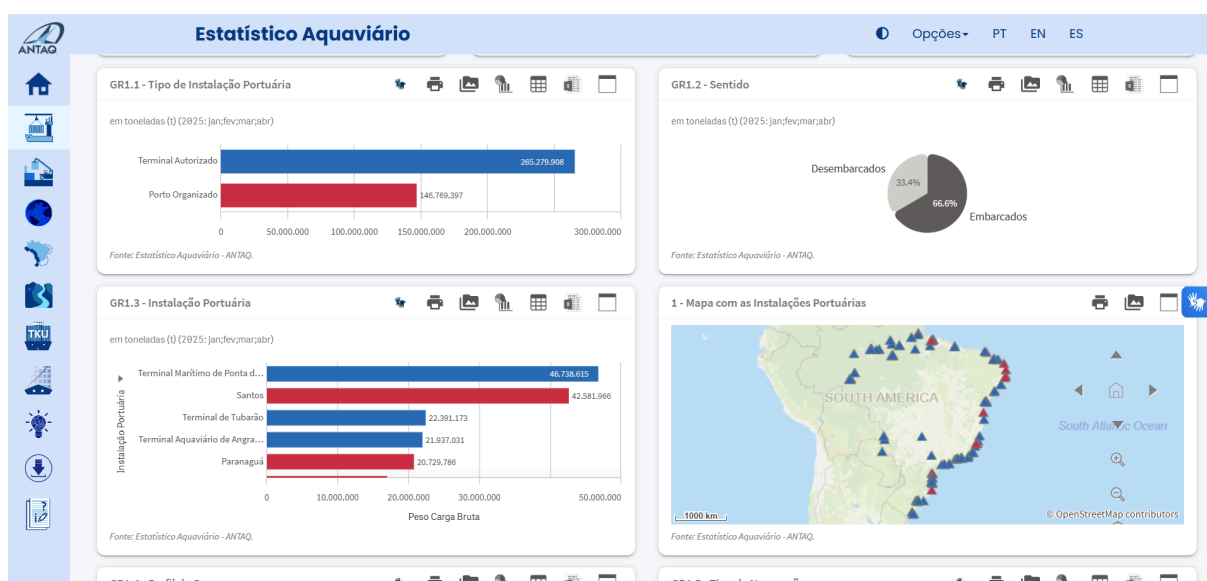
Fonte: SeaExplorer, 2025.²

¹ Disponível em: <https://www.marinetraffic.com/en/ais/home/centerx:1.9/centery:51.6/zoom:6>.

² Disponível em: <https://www.seaexplorer.com/>.

No Brasil, a ANTAQ se consolida como a principal provedora de dados públicos do setor aquaviário. Por meio do Painel Estatístico Aquaviário, a agência oferece uma plataforma interativa com indicadores sobre movimentação portuária, perfil de cargas, tipos de navegação e desempenho operacional. As visualizações permitem consultas históricas contribuindo para o acompanhamento da atividade do setor, como observado na Figura 4..

Figura 4 - Visualização do painel estatístico aquaviário.



Fonte: ANTAQ (2025).³

Embora o painel apresente um conjunto abrangente de informações e facilite uma visão geral das operações marítimas, seu foco ainda está restrito à análise estatística descritiva, disponibilizando dados em gráficos isolados e consultas segmentadas. Assim, a ferramenta, apesar de essencial para o acompanhamento do setor, não explora relações entre variáveis, nem oferece recursos preditivos que auxiliem na tomada de decisões estratégicas.

Diante desse cenário, este trabalho propõe ampliar o potencial dos dados públicos disponibilizados pela ANTAQ, desenvolvendo ferramentas que agregam valor analítico e preditivo às informações existentes. O objetivo é transformar a apresentação puramente estatística em uma plataforma de apoio à decisão, capaz de revelar padrões operacionais, antecipar gargalos e oferecer uma visão mais estratégica da gestão logística portuária nacional.

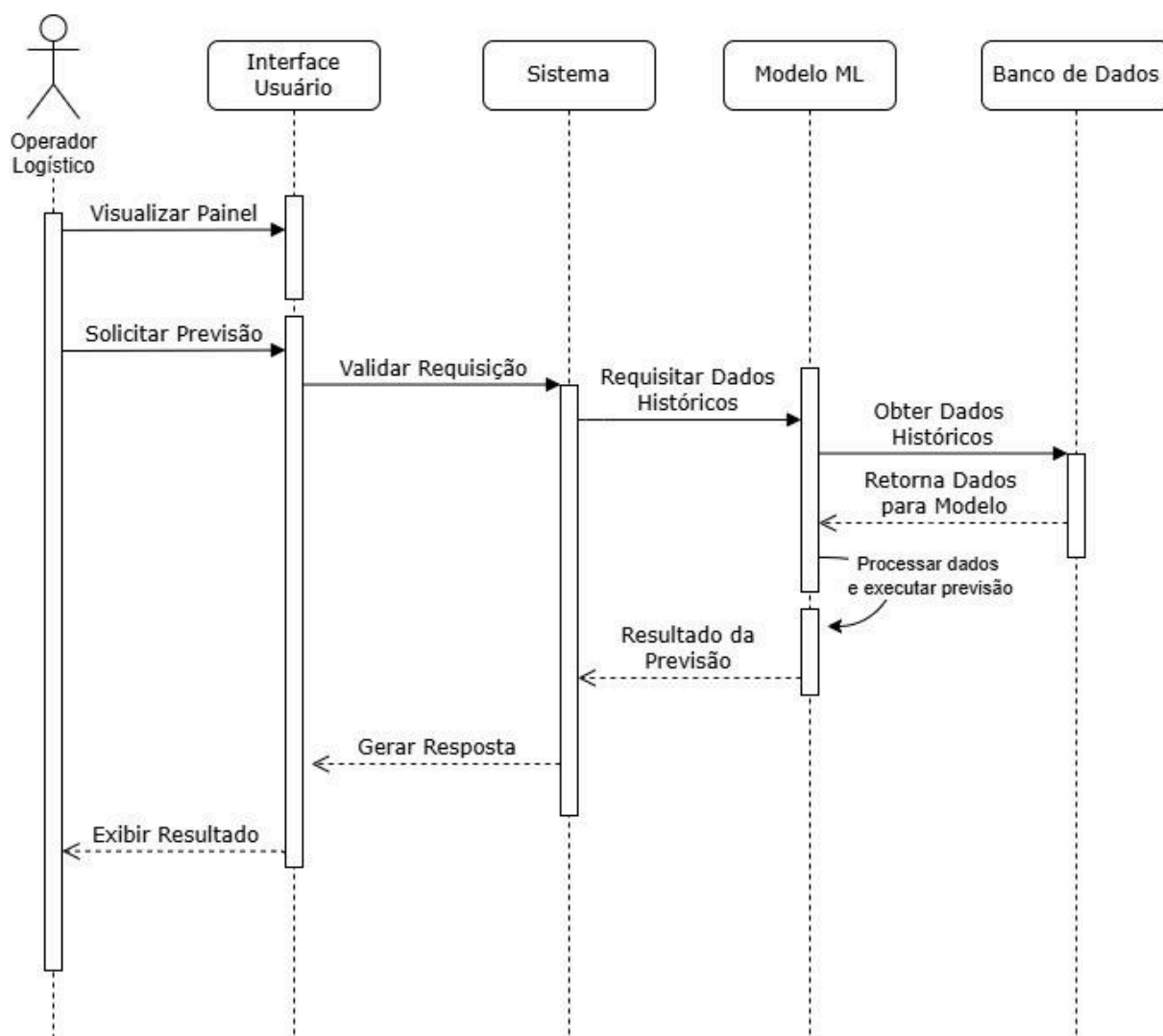
³ Disponível em: <https://web3.antaq.gov.br/ea/sense/index.html#pt>.

3 METODOLOGIA E DESENVOLVIMENTO

3.1 ARQUITETURA DO SISTEMA

Esta seção apresenta a estrutura geral do sistema, descrevendo seus principais módulos, fluxos de informação e interações entre componentes.

Figura 5 - Diagrama de sequência do sistema de análise e previsão portuária.



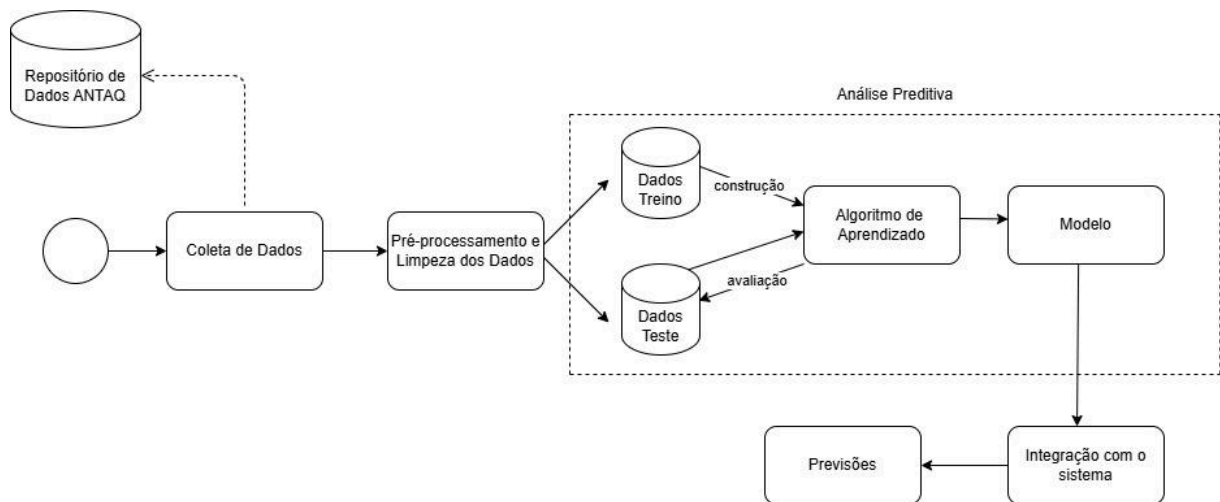
Fonte: Elaborado pela autora (2025).

O diagrama de sequência da Figura 5 ilustra o fluxo de interação entre componentes durante uma operação. O processo inicia quando o usuário acessa a interface web para visualizar indicadores no dashboard ou solicitar previsão de tempo de estadia. Após a solicitação de previsão, o sistema valida os dados de entrada (tipo de carga, porto,

características da embarcação), requisita dados históricos ao modelo ML, que processa as informações do banco de dados e retorna o resultado da previsão ao usuário por meio da interface.

O processo de análise preditiva, representado no diagrama de blocos da Figura 6, segue um pipeline estruturado em quatro etapas principais:

Figura 6 - Diagrama de bloco do aprendizado de máquina.

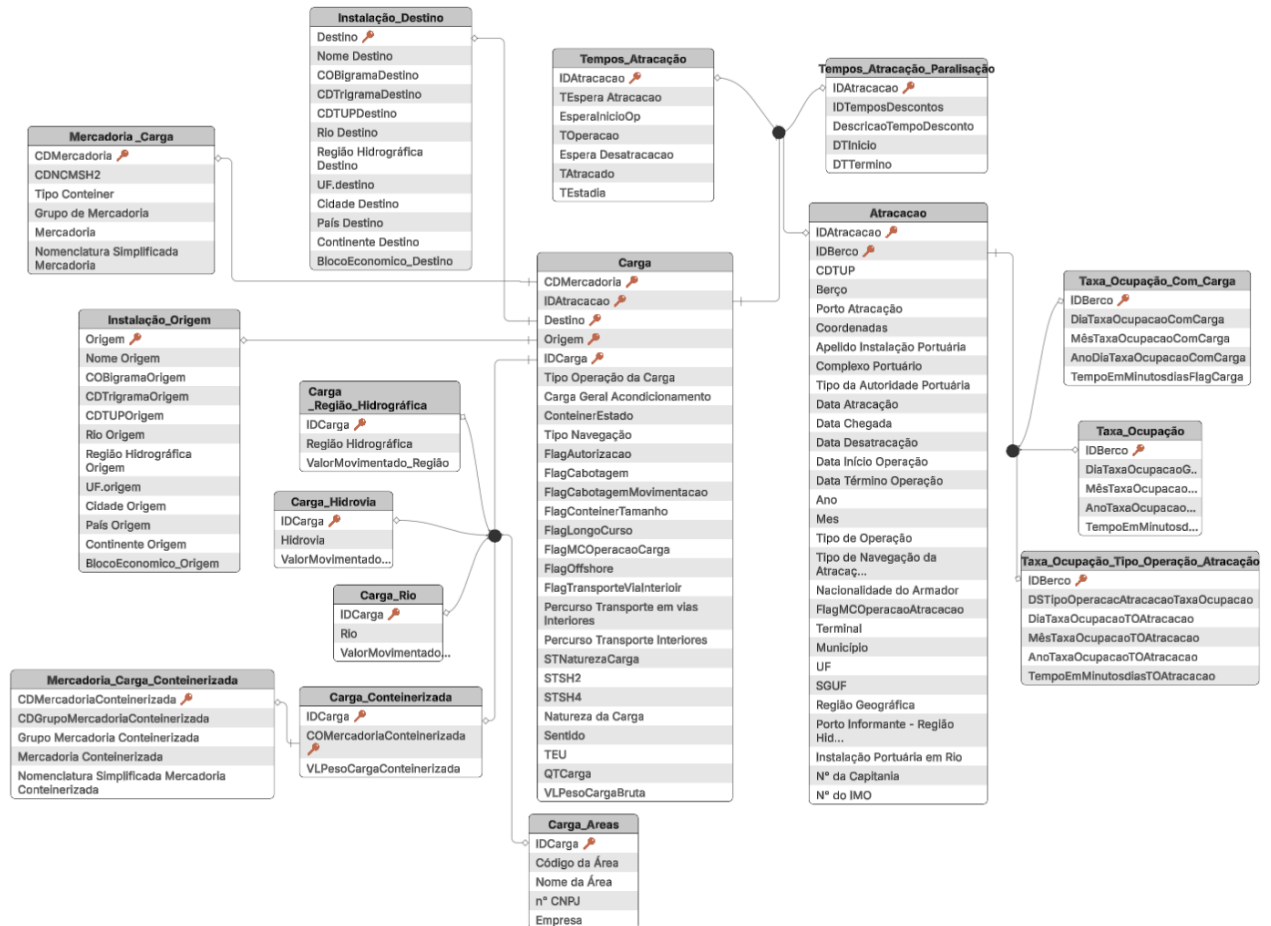


Fonte: Elaborado pela autora (2025).

1. Coleta de Dados: Os dados brutos foram obtidos do repositório de dados da ANTAQ, que contém informações históricas sobre atracções, movimentação de cargas, tempos operacionais e características das embarcações.
2. Pré-processamento e Limpeza dos Dados: Nesta etapa, foi realizado o tratamento de inconsistências, remoção de duplicatas, tratamento de valores nulos, padronização de formatos e engenharia de atributos (feature engineering), garantindo a qualidade dos dados utilizados no treinamento dos modelos.
3. Análise Preditiva: Os dados processados foram divididos em conjuntos de construção (treinamento e validação) e teste. O desempenho do modelo é avaliado utilizando-se os dados de teste, garantindo sua capacidade de generalização.
4. Integração com o Sistema: Uma vez validado, o modelo foi integrado ao sistema, permitindo que as previsões pudessem ser geradas em tempo real a partir das requisições dos usuários.

A Figura 7 apresenta o modelo de dados disponibilizado pela ANTAQ, com as tabelas e suas relações, representando a estrutura dos dados analisados no desenvolvimento do trabalho.

Figura 7 - Modelo de dados ANTAQ.



Fonte: ANTAQ (atualizado em abril/2025).

3.2 FERRAMENTAS E TECNOLOGIAS UTILIZADAS

O desenvolvimento do projeto envolveu o uso de diversas tecnologias, abrangendo desde o planejamento e a organização inicial até a construção da aplicação e o treinamento do modelo de aprendizado de máquina.

Na fase de experimentação, foi utilizado o Google Colaboratory (Colab), uma ferramenta que permite escrever e executar código Python diretamente no navegador. Seu uso facilitou o acesso a bibliotecas voltadas à ciência de dados e ao aprendizado de máquina, além de eliminar a necessidade de configurações locais. Essa praticidade favoreceu a execução rápida dos testes e a documentação dos resultados obtidos.

O sistema foi desenvolvido utilizando **Python 3.12**, escolhido por sua versatilidade e ampla utilização nas áreas de ciência de dados e aprendizado de máquina. Foram empregados frameworks e bibliotecas como Flask, Pandas e Scikit-learn, que proporcionaram uma base sólida para a criação da aplicação e para o processamento dos dados. Para o armazenamento e versionamento do código, utilizou-se o GitHub, garantindo controle sobre as versões do projeto.

3.2.1 COLETA E PRÉ-PROCESSAMENTO DOS DADOS

Os dados utilizados neste trabalho foram obtidos no portal de dados abertos da ANTAQ, disponibilizados em arquivos CSV anuais organizados por categorias (Atracação, Carga e Tempos Operacionais). Esses conjuntos contêm informações detalhadas sobre as operações portuárias brasileiras e formam a base principal do projeto.

O processo de consolidação dos dados foi automatizado por meio do script *filter_recent_data.py*, que realiza a unificação dos arquivos anuais em conjuntos únicos por categoria. O script foi desenvolvido para identificar automaticamente o separador utilizado (ponto e vírgula), ajustar a codificação de caracteres para UTF-8 e aplicar filtros temporais. Durante a execução, são geradas estatísticas descritivas básicas e indicadores de progresso, facilitando o acompanhamento do processamento. Os resultados são salvos no diretório

data/processed/2019_2024/, seguindo um padrão de nomenclatura (por exemplo, *atracao_2019_2024.csv*).

A etapa de pré-processamento foi conduzida pela classe *DataPreprocessor*, implementada no módulo *data_preprocessing.py*. Essa classe executa as principais rotinas de preparação dos dados:

- Conversão de tipos numéricos, corrigindo campos lidos como texto em razão do uso de vírgulas como separador decimal;
- Tratamento de colunas temporais, convertendo-as para o formato *datetime* do *Pandas*, o que permite manipulações temporais precisas;
- Otimização de memória, reduzindo o consumo de recursos com a conversão de tipos de dados para formatos mais leves.

Além disso, foram tratados valores ausentes, removidas duplicatas e aplicadas técnicas de detecção e tratamento de outliers, garantindo que o conjunto de dados estivesse limpo e padronizado para as etapas de análise exploratória e modelagem.

3.2.2 ANÁLISE EXPLORATÓRIA DOS DADOS

A Análise Exploratória de Dados (AED), ou *Exploratory Data Analysis* (EDA), constitui uma etapa fundamental em projetos de ciência de dados, abrangendo a exploração inicial e visualização de dados para identificar padrões, tendências, outliers e relações entre variáveis, permitindo a compreensão da estrutura, qualidade e características dos dados antes da aplicação de técnicas de modelagem (TUKEY, 1977).

Neste trabalho, foi realizada uma análise exploratória abrangente dos dados portuários da ANTAQ, cobrindo o período de 2010 a 2024. A exclusão dos dados de 2025 fundamentou-se na necessidade de evitar vieses decorrentes de anos incompletos, especialmente considerando padrões sazonais nas operações portuárias que poderiam distorcer análises estatísticas e temporais. O conjunto de dados abrange 15 anos completos, totalizando 1.193.226 registros de atracações, 27.595.160 registros de cargas movimentadas e 1.193.226 registros de tempos operacionais.

A análise foi desenvolvida em ambiente **Jupyter Notebook**, executado diretamente no VS Code, o que permitiu combinação de código, visualizações e documentação em documento único. A opção por utilizar o Jupyter localmente, em vez do Colab, justificou-se pela necessidade de maior estabilidade no processamento dos dados. Para garantir isolamento de dependências e reprodutibilidade, foi criado ambiente virtual Python (.venv), prática recomendada para projetos de ciência de dados.

O processo empregou um conjunto de bibliotecas Python, tanto para manipulação e processamento dos dados, quando para a visualização e análises estatísticas, sendo elas:

- **Pandas (v2.3.2):** Para manipulação e análise de dados estruturados em formato tabular (DataFrames), utilizada para leitura, limpeza, transformação e agregação dos dados portuários;
- **NumPy (v2.3.2):** Computação numérica com estruturas eficientes e operações matemáticas vetorizadas;
- **Matplotlib:** Biblioteca de visualização de baixo nível que oferece controle detalhado sobre criação de gráficos estáticos;
- **Seaborn:** Construída sobre Matplotlib, fornece interface de alto nível para gráficos estatísticos esteticamente aprimorados, facilitando a visualização de distribuições e relações entre variáveis;
- **Plotly:** Criação de gráficos interativos com zoom, filtragem e exibição de informações contextuais;
- **SciPy:** Funções estatísticas avançadas, incluindo testes de hipóteses e análise de distribuições;
- **Pathlib:** Manipulação de caminhos de arquivos de forma orientada a objetos; e
- **Datetime:** Manipulação de datas e horários para análises temporais.

Considerando o volume expressivo de dados, foi implementado script auxiliar (*consolidate_historical_data.py*) responsável pela consolidação, otimização e padronização. O script une arquivos anuais da ANTAQ em três arquivos CSV principais: *atracao_2010_2024.csv* (1.193.226 registros), *carga_2010_2024.csv* (27.595.160 registros) e *tempos_2010_2024.csv* (1.193.226 registros). O processo inclui otimização de memória através de conversão de tipos de dados para formatos eficientes e padronização mediante uniformização de codificação (UTF-8) e separadores. O sistema verifica automaticamente a

existência dos arquivos consolidados e, quando disponíveis, realiza carregamento significativamente mais rápido, reduzindo o tempo de inicialização.

Posteriormente, foi realizada uma segunda análise exploratória restrita ao período de 2019 a 2024, com o objetivo de reduzir o custo operacional e o tempo de processamento, bem como verificar se os ganhos analíticos obtidos com o conjunto completo (2010-2024) justificavam o esforço computacional adicional em relação ao uso de um subconjunto mais recente. Essa abordagem manteve aproximadamente 50% do volume total de dados e apresentou vantagens adicionais, já que diversas tabelas e campos que não existiam nos primeiros anos da série histórica passaram a estar disponíveis de forma consistente, e o período selecionado contempla contextos econômicos distintos e relevantes, incluindo cenários pré-pandemia (2019), durante a pandemia de COVID-19 (2020-2021) e recuperação pós-pandemia (2022-2024), permitindo análises comparativas de diferentes ciclos econômicos que impactaram diretamente o setor portuário brasileiro.

3.2.3 DESENVOLVIMENTO DO MODELO PREDITIVO

O modelo preditivo foi ajustado ao longo do processo, com o objetivo corrigir falhas metodológicas que pudessem comprometer sua aplicabilidade. Em um primeiro momento, foi construído um modelo baseado no algoritmo **XGBoost** (Extreme Gradient Boosting), treinado com variáveis como o tempo de operação (TOperacao) e o tempo de espera para desatracação (TEsperaDesatracacao). Embora tenha apresentado bons resultados nas métricas de desempenho, uma análise mais aprofundada revelou um problema crítico de vazamento de dados (data leakage), isso ocorria porque algumas dessas variáveis só se tornam conhecidas após o término da operação portuária, inviabilizando o uso do modelo para previsões antecipadas (*ex-ante*), que constituem o foco principal do sistema.

Para solucionar esse problema, foi desenvolvido o Modelo Operacional, que simula com precisão o cenário real de predição, utilizando apenas informações disponíveis no momento da chegada da embarcação ao porto. As principais modificações envolveram a remoção completa de variáveis pós-operacionais e a criação de novas features a partir do histórico de atracações, por meio de um processo de *feature engineering*.

Essas variáveis foram agrupadas em duas categorias principais. As features temporais, extraídas da data de atracação (ano, mês, dia da semana, trimestre e turno do dia), permitem capturar padrões sazonais e operacionais. Já as features de agregação histórica refletem o comportamento médio do sistema em períodos anteriores e são calculadas utilizando janelas expansivas que consideram apenas dados passados. Entre os principais exemplos estão o tempo médio de estadia por porto (TEstadia_Media_Porto), por terminal (TEstadia_Media_Terminal) e por tipo de operação (TEstadia_Media_TipoOperacao). A construção dessas variáveis foi feita utilizando a função *expanding().mean()* da biblioteca Pandas, o que garante que o modelo se baseie apenas em informações anteriores ao evento previsto, evitando qualquer contaminação temporal.

A escolha do **XGBoost** foi motivada por seu desempenho em dados tabulares estruturados e por sua capacidade de aprendizado sequencial, em que cada árvore de decisão é treinada para corrigir erros das anteriores. Esse mecanismo de *boosting* oferece ao modelo um poder preditivo superior em relação a outros métodos, como **Random Forest** e **LightGBM**.

Após o treinamento, o modelo foi serializado com a biblioteca Joblib, gerando o arquivo *xgb_pipeline_operacional.joblib*, que encapsula todo o pipeline de pré-processamento e predição. Esse arquivo é carregado na inicialização da aplicação web, permitindo a geração de previsões em tempo real com baixa latência. Também foi criado um arquivo de metadados (*feature_columns_operacional.json*), contendo informações sobre as colunas de entrada, seus tipos e ordem, assegurando consistência entre os ambientes de treinamento e produção.

3.2.4 DESENVOLVIMENTO DO DASHBOARD

O desenvolvimento de interfaces visuais interativas constitui elemento fundamental em sistemas de apoio à decisão, permitindo exploração intuitiva de dados complexos (Oliveira et al., 2018, p. 236). Neste trabalho, foi desenvolvido dashboard interativo para visualização e análise dos dados portuários da ANTAQ, abrangendo o período de 2019 a 2024.

O dashboard foi implementado utilizando **Streamlit**, framework open-source em Python especializado na criação rápida de aplicações web interativas voltadas para ciência de dados (STREAMLIT, 2025). A escolha se deu por sua sintaxe simples e intuitiva, e

características técnicas como integração nativa com bibliotecas de visualização Python, capacidade de reatividade automática e facilidade de implantação em ambientes web.

Para visualizações interativas, foi utilizada a biblioteca **Plotly**, que oferece recursos avançados de interatividade incluindo zoom, pan, filtros dinâmicos e tooltips informativos. A biblioteca gera gráficos baseados em JavaScript (utilizando a biblioteca D3.js internamente). Utilizou-se predominantemente a interface Plotly Graph Objects para customizações específicas alinhadas aos requisitos de análise portuária.

3.2.5 DESENVOLVIMENTO DO WEBSITE

A interface web foi desenvolvida para proporcionar um acesso intuitivo às funcionalidades de análise e predição.

O backend foi implementado utilizando Flask, micro-framework web em Python reconhecido pela simplicidade e facilidade de integração com bibliotecas de ciência de dados. A estrutura de rotas organiza funcionalidades em endpoints específicos: autenticação (/login, /register), página principal (/index), funcionalidades (/dashboard, /predicao, /sobre), gerenciamento de conta (/gerenciarconta, /editarconta, /excluirconta) e endpoint de API de predição (/predizer).

O sistema de autenticação utiliza sessões Flask para manutenção de estados de login e **PyMySQL** para comunicação com banco de dados **MySQL**, responsável pelo armazenamento de credenciais e informações de usuário.

O frontend foi desenvolvido com **HTML5**, estilização através de **Tailwind CSS** e iconografia fornecida pela biblioteca **Lucide Icons**. Tailwind CSS é um framework utilitário que permite construção rápida de interfaces customizáveis por meio de classes atômicas aplicadas diretamente aos elementos HTML (TailwindCSS, 2025). A responsividade foi implementada utilizando sistema de grid do Tailwind CSS com breakpoints responsivos, garantindo adaptação a dispositivos desktop e smartphones.

A comunicação frontend e backend para funcionalidades de predição utiliza arquitetura baseada em **REST API** com requisições assíncronas via JavaScript nativo (Fetch

API). Quando o usuário preenche o formulário de predição e submete os dados, uma requisição POST com payload JSON é enviada ao endpoint `/predizer`, que processa via pipeline de aprendizado de máquina e retorna resposta JSON com predição ou mensagens de erro.

O carregamento do modelo ocorre na inicialização da aplicação Flask, utilizando **Joblib** para deserialização do pipeline (*xgb_pipeline_operacional.joblib*) e arquivo JSON com metadados de colunas de entrada (*feature_columns_operacional.json*). Essa estratégia de carregamento único em memória reduz o tempo de resposta nas requisições posteriores.

A integração entre componentes frontend e backend utiliza template engine **Jinja2**, nativo do Flask, que permite renderização dinâmica de conteúdo HTML com variáveis Python, loops e estruturas condicionais, facilitando personalização de interfaces conforme estado do usuário e dados disponíveis.

3.4 PROCESSOS DE VALIDAÇÃO

Para garantir que o sistema desenvolvido fosse confiável e útil, foram realizados diferentes processos de validação, envolvendo desde a análise dos resultados do modelo até o desempenho funcional e a experiência de uso. Esses testes abrangeram os três principais componentes do projeto: o modelo preditivo, o dashboard interativo e a interface web.

A validação do modelo preditivo seguiu uma abordagem de validação temporal estrita, essencial para medir de forma realista a capacidade do modelo de generalizar para períodos futuros. Os dados foram divididos cronologicamente: o treinamento utilizou registros de 2019 a 2023, enquanto o teste foi feito apenas com dados de 2024. Esse método, conhecido como *time series split* ou *walk-forward validation*, avalia o desempenho do modelo em condições que simulam o uso real, ou seja, prever o futuro com base apenas em informações passadas (TASHMAN, 2000). Diferente da validação cruzada tradicional (*k-fold cross-validation*), que embaralha os dados de maneira aleatória, a validação temporal mantém a ordem cronológica, evitando que o modelo “aprenda” com dados futuros e produza resultados artificialmente precisos.

A avaliação quantitativa considerou métricas amplamente reconhecidas em aprendizado de máquina: o MAE (erro médio absoluto), que indica o erro médio em horas; o RMSE (raiz do erro quadrático médio), que dá mais peso a erros maiores; o R^2 (coeficiente de determinação), que mostra quanto da variabilidade dos dados o modelo consegue explicar; e o MAPE (erro percentual absoluto médio), que expressa o erro em termos percentuais.

A validação do dashboard teve foco na usabilidade e na clareza das informações apresentadas. Foram realizados testes para garantir o correto carregamento dos dados consolidados, o funcionamento dos filtros dinâmicos (como ano, região, porto e tipo de operação) e a atualização automática das visualizações conforme as interações do usuário. Também foram medidos o tempo de carregamento inicial e o tempo de resposta após a aplicação dos filtros. Para garantir consistência, os valores exibidos no dashboard foram comparados com consultas diretas aos dados brutos, assegurando a precisão dos indicadores e métricas apresentadas.

Por fim, a validação da interface web envolveu tanto testes de funcionalidade quanto de experiência do usuário. Todas as rotas implementadas foram verificadas, incluindo autenticação de usuários (com credenciais válidas e inválidas), criação de novas contas com verificação de e-mail duplicado, edição e exclusão de perfis. O endpoint da API */predizer* também foi testado com diferentes tipos de entrada: dados válidos, valores extremos e campos ausentes ou incorretos. Para assegurar boa experiência em qualquer dispositivo, a responsividade da interface foi testada utilizando a extensão *mobile first* do Chrome, que simula diferentes tamanhos de tela. Assim, confirmou-se que o layout se adapta adequadamente e que as funcionalidades permanecem acessíveis tanto em desktop quanto em dispositivos móveis.

4 RESULTADOS

4.1 FUNCIONALIDADES DO SISTEMA

A Figura 8 apresenta a interface do sistema PortoVision. As principais funcionalidades do sistema são:

- Interface web responsiva e amigável: desenvolvida com foco na usabilidade, proporcionando uma navegação intuitiva.
- Dashboard interativo de indicadores logísticos: exibe visualizações dinâmicas sobre o transporte marítimo, incluindo tipos de carga, portos mais movimentados, taxas de ocupação e produtividade portuária.
- Predição do tempo de estadia de embarcações: utiliza modelo de machine learning treinado com dados históricos para estimar o tempo de operação das embarcações em diferentes portos.

Figura 8 - Interface PortoVision.



Fonte: Sistema desenvolvido pela autora.

4.2 ANÁLISE EXPLORATÓRIA

A análise exploratória realizada neste trabalho teve como objetivos principais:

1. **Compreender a estrutura e qualidade dos dados:** Identificação de tipos de variáveis, padrões de valores ausentes, inconsistências e necessidade de tratamento;
2. **Identificar padrões operacionais:** Análise de características típicas das operações portuárias brasileiras, incluindo distribuição de tipos de carga, perfil das embarcações e produtividade média;
3. **Analisar tendências temporais e sazonalidade:** Investigação de padrões de crescimento ao longo dos anos e identificação de variações sazonais na movimentação de cargas e ocupação de berços;
4. **Mapear distribuição geográfica:** Caracterização da importância relativa de diferentes portos e regiões no sistema portuário nacional;

Esses objetivos servem como base para as análises apresentadas a seguir, nas quais são examinados os principais padrões e tendências observados nos dados. A seção 4.2.1 detalha os resultados obtidos durante a exploração dos dados portuários brasileiros no período de 2010 a 2024, destacando comportamentos recorrentes, variações temporais e características relevantes para a compreensão das operações portuárias brasileiras.

4.2.1 PADRÕES E TENDÊNCIAS IDENTIFICADOS

A partir da exploração dos dados históricos, foi possível reconhecer comportamentos recorrentes e mudanças significativas que moldam as operações portuárias brasileiras. A seguir, são discutidos os padrões e tendências identificados.

(i) Evolução temporal e ciclos econômicos

A Tabela 1 apresenta a evolução do número de atracações nos portos brasileiros ao longo do período de 2010 a 2024, acompanhada das taxas de crescimento percentual ano a ano. Essa análise histórica permite identificar ciclos econômicos distintos que impactaram diretamente a atividade portuária nacional: um período de retração entre 2012 e 2016, seguido por recuperação gradual a partir de 2017 e forte aceleração nos anos subsequentes à pandemia de COVID-19. É possível observar que, após a queda em 2016 (menor volume registrado), o setor demonstrou crescimento consistente, chegando em 2024 com taxa de expansão de 13,76%, a mais elevada da série analisada.

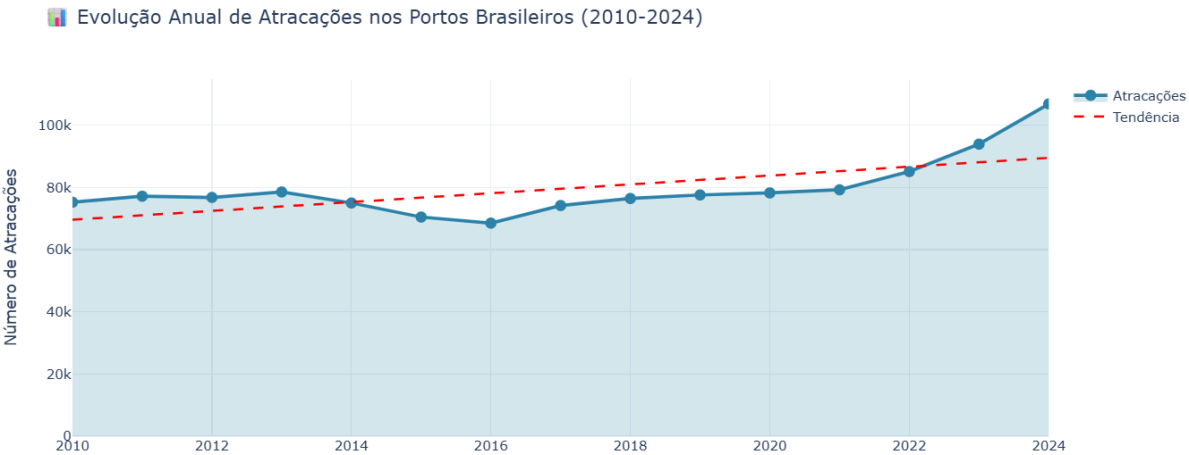
Tabela 1 - Distribuição de registros de atracções por ano.

Ano	Atracções	Crescimento (%)
2010	75246	NaN
2011	77167	2.55
2012	76777	-0.51
2013	78540	2.30
2014	74992	-4.52
2015	70469	-6.03
2016	68515	-2.77
2017	74182	8.27
2018	76440	03.04
2019	77577	1.49
2020	78232	0.84
2021	79238	1.29
2022	85092	7.39
2023	93918	10.37
2024	106841	13.76

Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

Essa análise revelou uma variação expressiva ao longo do período estudado, com o ano de 2016 registrando o menor volume (68.515 atracções) e 2024 apresentando o pico histórico (106.841 atracções), representando um crescimento de aproximadamente 56% entre esses extremos, como exibido na Figura 9.

Figura 9 - Evolução anual de atracção nos portos brasileiros.



Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

O ano de 2016 marcou o ponto mais baixo da atividade portuária brasileira, reflexo de uma crise econômica sem precedentes. O país atravessava a mais profunda recessão de sua história recente, o PIB caiu 3,8% em 2015 e mais 3,6% em 2016, acumulando retração de 7,2% no biênio (CURY; SILVEIRA, 2017). A crise atingiu de forma simultânea os setores agropecuário, industrial e de serviços, o que não acontecia desde 1996, e provocou forte queda no consumo das famílias (-4,2%), pressionado por juros altos, restrição ao crédito e deterioração do emprego e da renda. No comércio exterior, a redução das exportações de commodities foi agravada pela desaceleração da economia chinesa e pela queda nos preços internacionais do petróleo e do minério de ferro. Junto a isso tivemos também o impacto da Operação Lava Jato, que freou investimentos e atrasou operações, sobretudo em terminais ligados ao setor de petróleo (ANTAQ, 2017). O resultado foi uma expressiva redução da demanda por transporte marítimo e, consequentemente, no número de atracações registradas.

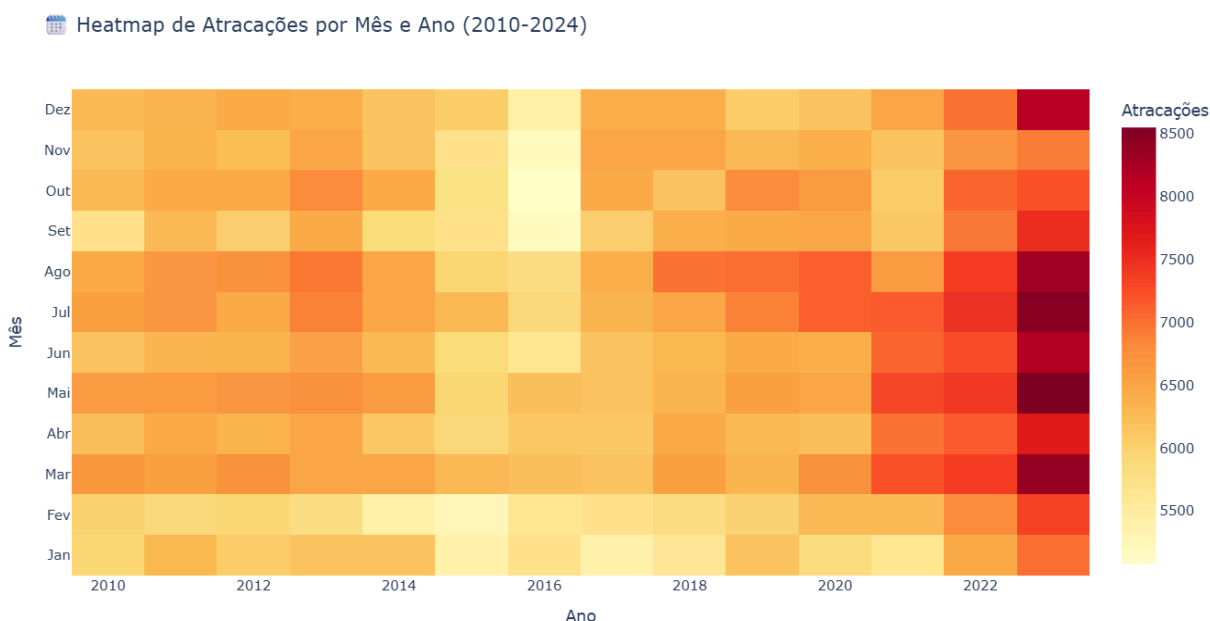
Por outro lado, 2024 foi o ano de maior movimentação portuária da série analisada, simbolizando a recuperação e o fortalecimento do setor. Segundo dados publicados pelo Governo Federal, o desempenho positivo decorreu de uma conjunção de fatores, como a retomada econômica pós-pandemia, o crescimento das exportações de commodities, a entrada em operação de novos TUPs e os investimentos em infraestrutura logística nas regiões do Arco Norte e Sudeste, voltados à redução de custos e à ampliação da capacidade operacional (BRASIL, 2024).

Esse recorde de 2024, sugere potencial de crescimento contínuo para os próximos anos, sustentado por perspectivas favoráveis: consolidação do Brasil como fornecedor estratégico de commodities agrícolas e energéticas no mercado global; amadurecimento de investimentos recentes em infraestrutura portuária, com ganhos progressivos de eficiência operacional; expansão da frota de navios e modernização tecnológica dos terminais; e políticas governamentais de incentivo ao setor, materializadas nos investimentos previstos pelo Novo PAC até 2030. A comparação entre 2016 e 2024 evidencia a capacidade de recuperação do sistema portuário brasileiro, reforçando a importância da análise histórica de dados para compreender ciclos econômicos e projetar tendências com base sólida.

(ii) Padrões sazonais

A análise de sazonalidade, visualizada através do heatmap apresentado na Figura 10, revelou padrões consistentes de variação mensal na atividade portuária ao longo dos 15 anos estudados. Julho foi identificado como o mês mais movimentado, com 95.216 atracções acumuladas no período, enquanto janeiro registrou o menor volume, com 83.846 atracções, resultando em variação sazonal de 12,60%.

Figura 10 - Mapa de calor de atracções por mês e ano.



Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

Dessa forma, foi revelado um padrão recorrente na movimentação portuária brasileira, com picos nos meses de julho e menor atividade em janeiro ao longo dos 15 anos analisados, que podem ser atribuídos a múltiplos fatores operacionais e comerciais. Em 2025, esse comportamento voltou a se confirmar, em julho foi registrado um novo recorde histórico de movimentação, com 124,7 milhões de toneladas transportadas, impulsionado principalmente pela cabotagem, que respondeu por cerca de 20% do total (CNN BRASIL, 2025).

Esses resultados reforçam a relevância de compreender os ciclos sazonais para aprimorar o planejamento logístico e a alocação de recursos. Identificar antecipadamente os períodos de maior movimentação permite otimizar o uso da infraestrutura, reduzir gargalos operacionais e melhorar a previsibilidade das operações. Da mesma forma, reconhecer meses

de menor atividade, como janeiro, contribui para o planejamento de manutenções, treinamentos e ajustes de capacidade.

A intensidade da coloração no heatmap também permite identificar anos atípicos. Nota-se que 2016 apresenta tons consistentemente mais claros ao longo de todos os meses, confirmando visualmente o impacto da crise econômica já discutida. Em contraste, 2023 e 2024 exibem coloração mais intensa, refletindo a expansão acelerada do período recente. Interessante observar que a sazonalidade mantém-se relativamente estável ao longo dos anos, sugerindo que os padrões de demanda por transporte marítimo estão fortemente ancorados em ciclos agrícolas e comerciais estruturais, menos sujeitos a flutuações conjunturais de curto prazo.

(iii) Distribuição geográfica e concentração portuária

A análise da distribuição geográfica das operações portuárias revela concentração significativa em determinadas regiões e portos específicos, refletindo especializações regionais e vantagens logísticas estruturais. Conforme demonstrado na Tabela 2, a região Norte lidera em volume de atracações (38,19%), seguida pelo Sudeste (29,32%) e Sul (14,78%).

Tabela 2 - Distribuição de atracações por região geográfica (2010-2024).

Região	Atracações	Porcentagem
Norte	454054	38.19
Sudeste	348667	29.32
Sul	175686	14.78
Nordeste	166995	14.05
Centro-Oeste	43596	3.67

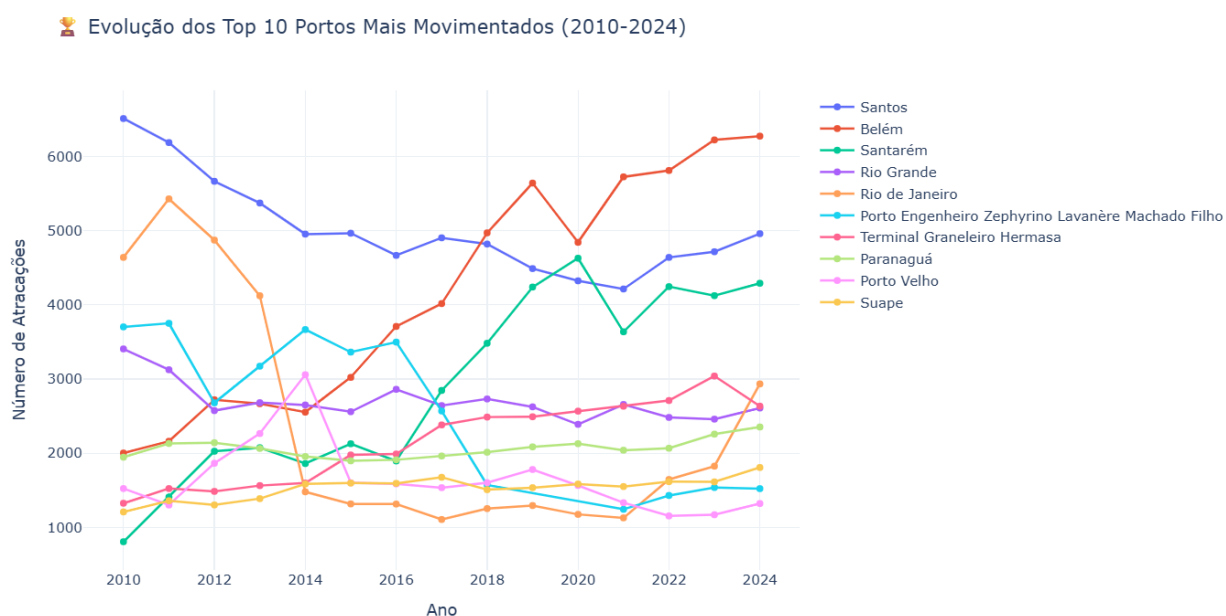
Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

A região Norte merece atenção especial, visto que historicamente o Sudeste concentrava a maior parte das operações portuárias brasileiras. Esse deslocamento gradual não ocorreu por acaso, ele é fruto de uma estratégia de desconcentração logística, impulsionada pela expansão da fronteira agrícola na região do MATOPIBA (Maranhão, Tocantins, Piauí e Bahia) e investimentos expressivos em terminais do chamado Arco Norte.

que abriram caminho para que portos como Itaquí, Barcarena e Santarém assumissem protagonismo crescente.

Apesar do Sudeste estar em segunda posição, sua importância e papel estratégico é inquestionável, pela diversidade de carga e pela proximidade de estados como São Paulo, Rio de Janeiro e Minas Gerais. O Porto de Santos, por exemplo, continua sendo a principal porta de entrada e saída do comércio exterior brasileiro, mas já não está sozinho como referência única, apesar de ser o maior porto da América Latina.

Figura 11 - Evolução dos top 10 portos mais movimentados.



Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

Outro fator analisado durante esse estudo mostra queda drástica na movimentação entre 2013 e 2014 no Porto do Rio de Janeiro. Segundo dados do Centro Nacional de Navegação Transatlântica (Centronave), citados por Fariello e Ordoñez (2017), em 2016 o porto registrou queda de 15,6% na movimentação de contêineres, três vezes superior à média nacional de 4%, seguindo tendência de queda que já havia se iniciado em 2014. A causa dessa redução decorreu de um conjunto de fatores interligados. Problemas crônicos de dragagem dificultavam a entrada de navios de grande porte, limitando a competitividade do terminal nas rotas internacionais. As restrições impostas pela Marinha, somadas ao aumento do custo com rebocadores, agravaram o cenário. No entorno, gargalos no acesso rodoviário, elevavam o tempo de espera dos caminhões, enquanto a crise do setor de petróleo e gás reduzia o volume

de cargas movimentadas. Questões de segurança pública também pesaram na decisão de armadores e operadores logísticos, que passaram a optar por portos mais eficientes e previsíveis (FARIELLO; ORDÓÑEZ, 2017).

(iv) Perfil das Operações

A análise da distribuição por tipo de operação, apresentada na Tabela 3, mostra uma clara concentração das atividades portuárias. A maior parte das atracções está ligada à movimentação de carga, que responde por 77,55% do total, seguida por operações de apoio (16,11%) e transporte de passageiros (4,08%).

Tabela 3 - Distribuição de atracções por tipo de operação (2010-2024).

Tipo	Frequência	Porcentagem
Movimentação da Carga	922506	77.55
Apoio	191583	16.11
Passageiro	48509	04.08
Abastecimento	14417	1.21
Reparo/Manutenção	7668	0.64
Marinha	2332	0.20
Misto	2273	0.19
Retirada de Resíduos	263	0.02

Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

A predominância de operações de movimentação de carga confirma a vocação comercial do sistema portuário brasileiro, fortemente orientado ao comércio exterior. Dentro desta categoria, destacam-se diferentes tipos de carga, operações de carga geral, contêineres, graneis sólidos (minério, grãos) e graneis líquidos (petróleo, combustíveis, produtos químicos). As operações de apoio ficam logo abaixo, representando cerca de 16% das atracções, ela engloba serviços auxiliares essenciais ao funcionamento portuário, como rebocadores que auxiliam manobras de atracção e desatracção; embarcações de abastecimento de plataformas de petróleo; e manuseio de âncoras.

Já o transporte de passageiros, embora menos representativo no total (4,08%), mantém importância regional. Ele se concentra principalmente em áreas onde os rios são o principal

meio de deslocamento entre comunidades, e em áreas litorâneas com travessias regulares, como as balsas que operam em São Paulo e no Rio de Janeiro. O turismo de cruzeiros também contribui para esse segmento, com destaque para portos como Santos, Rio de Janeiro e Salvador, ainda que sua participação seja relativamente pequena e limitada pela sazonalidade das viagens marítimas.

Quanto ao tipo de navegação, a Tabela 4 detalha a distribuição das atracções.

Tabela 4 - Distribuição das atracções por tipo de navegação (2010-2024).

Tipo	Atracções	Porcentagem
Interior	530573	44.60
Longo Curso	318301	26.76
Cabotagem	209450	17.61
Apoio Portuário	74799	6.29
Apoio Marítimo	56428	4.74

Fonte: Elaborado pela autora com base em dados da ANTAQ (2025).

A navegação interior, que responde por 44,6% das atracções, evidencia a relevância dos rios no território brasileiro. Em especial nas regiões Norte e Centro-Oeste, as hidrovias funcionam como verdadeiras artérias de transporte, permitindo o deslocamento de cargas, pessoas e suprimentos para localidades muitas vezes inacessíveis por rodovias. Embarcações de diversos portes navegam rios como Amazonas, Negro, Tapajós e Madeira, transportando mercadorias, passageiros e insumos essenciais para regiões mais afastadas dos grandes centros.

A navegação de longo curso (26,76%) representa o comércio marítimo internacional, com navios que conectam portos brasileiros a destinos em todos os continentes. Essas operações envolvem primariamente exportação de commodities (minério de ferro, soja, celulose, carnes, açúcar, petróleo) e importação de produtos industrializados, insumos químicos, fertilizantes e equipamentos. Já a cabotagem (17,61%) refere-se ao transporte marítimo ao longo da costa brasileira, conectando portos nacionais.

Os padrões e tendências identificados na análise exploratória revelam que o sistema portuário brasileiro está em constante transformação. As análises subsequentes, dashboard interativo e modelo de aprendizado de máquina, foram desenvolvidas considerando esses

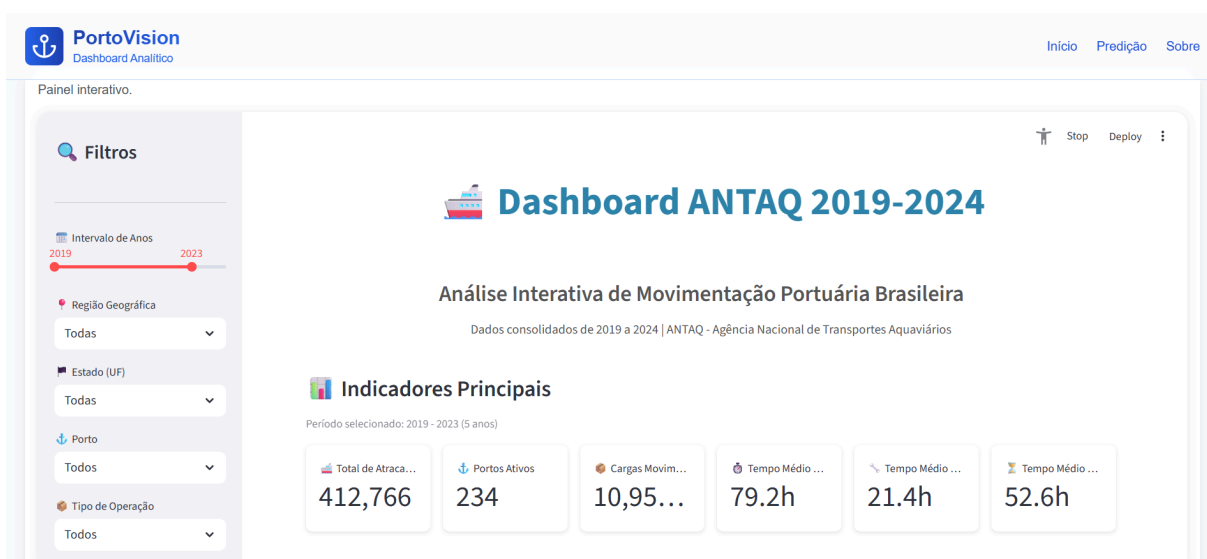
achados, buscando transformar dados históricos em ferramentas práticas de apoio à gestão portuária estratégica.

4.3 DASHBOARD DESENVOLVIDO

A tela inicial do dashboard apresenta um cabeçalho informativo que contextualiza o período dos dados analisados (2019-2024) e o propósito da ferramenta, seguido por um painel de indicadores-chave (KPIs) que apresenta métricas operacionais do sistema portuário brasileiro. Esses indicadores incluem o total de atracações registradas no período, o número de portos ativos, o tempo médio de estadia das embarcações e o tempo médio de operação.

Na lateral esquerda, a barra de filtros permite que o usuário segmente os dados de acordo com diferentes critérios: intervalo de anos, região geográfica, estado, porto específico, tipo de operação e mês. Esses filtros foram projetados para funcionar de forma interdependente e cascadeada, por exemplo, ao selecionar uma região geográfica, apenas os estados pertencentes àquela região aparecem como opções; ao escolher um estado, apenas os portos localizados nele ficam disponíveis. Todos os gráficos e métricas são atualizados conforme os filtros são aplicados, essa abordagem de *linked views*, onde diferentes visualizações reagem simultaneamente às interações do usuário, é reconhecida como técnica eficaz para exploração visual de dados multidimensionais (SHNEIDERMAN, 1996).

Figura 12 - Tela dashboard PortoVision.



Fonte: Sistema desenvolvido pela autora.

A área central do dashboard organiza as visualizações em diferentes seções que abrangem desde indicadores gerais até análises operacionais avançadas. Inicialmente, são apresentados os indicadores gerais, que incluem os vinte portos mais movimentados, a distribuição por tipo de operação, a evolução mensal das atracações, a distribuição por região e o ranking dos dez estados com maior movimentação. Em seguida, são exibidas as análises de tempos operacionais, que apresentam estatísticas detalhadas sobre o tempo de operação em horas, e as análises de carga, que exploram a distribuição por sentido e destacam as dez principais naturezas de carga. Por fim, as análises operacionais avançadas abordam aspectos de produtividade por tipo de carga, instalação e turno, bem como métricas de ineficiência operacional (gaps de tempo) e a utilização inadequada de berços. Todos os gráficos foram desenvolvidos com recursos interativos oferecidos pela biblioteca Plotly, permitindo ao usuário explorar os dados de forma dinâmica, ao posicionar o cursor sobre os elementos visuais, são exibidas as informações com valores, também sendo possível aplicar zoom, alternar modos de visualização e exportar as visualizações em formato PNG.

Ao fim da tela, há uma área com os dados detalhados das visualizações, onde é possível explorar os registros individuais de atracações, cargas movimentadas e tempos operacionais. Nela é possível exportar os registros filtrados em formato CSV, facilitando análises adicionais em outras ferramentas.

4.4 MODELO PREDITIVO

A interface de predição foi desenvolvida para tornar a utilização do modelo de aprendizado de máquina acessível a usuários sem conhecimento técnico. A tela apresenta um formulário onde o usuário preenche informações sobre a operação portuária que deseja prever: porto de atracação, tipo de operação, natureza da carga e outras variáveis relevantes. Os campos do formulário incluem validações que orientam o preenchimento correto, exibindo mensagens de erro caso informações obrigatórias estejam ausentes ou em formato inadequado.

Após o preenchimento e submissão do formulário, o sistema processa os dados em tempo real, aplicando o pipeline de pré-processamento e acionando o modelo XGBoost operacional treinado. O resultado da predição é exibido de forma destacada na tela, informando o tempo de estadia estimado em horas.

Figura 13 - Tela previsão PortoVision.

PortoVision
Predição Operacional

Previsão de Tempo Total de Estadia

Dados Operacionais

Tempo de Espera para Atracação (h)	Porto de Atracação
6.21	Porto de Santos
Terminal	Tipo de Operação
Tecon Santos	Movimentação da Carga
Tipo de Navegação da Atracação	UF
Cabotagem	SP
Natureza da Carga	Turno
Granel Sólido	Dia
Ano	Mês

Tempo previsto de estadia: 51.17 horas

Fonte: Sistema desenvolvido pela autora.

4.4.1 TREINAMENTO E VALIDAÇÃO

Conforme descrito na seção de metodologia, o conjunto de dados foi dividido cronologicamente, utilizando registros de 2019 a 2023 para treinamento e validação, e reservando exclusivamente os dados de 2024 para teste final. Essa estratégia de validação temporal garante que o modelo seja avaliado em sua capacidade real de prever o futuro, sem acesso a informações que ainda não estariam disponíveis no momento da predição.

Durante o treinamento, o modelo XGBoost foi ajustado com 300 estimadores, profundidade máxima de árvores igual a 7 e taxa de aprendizado de 0,05, parâmetros selecionados para equilibrar capacidade preditiva e prevenção de overfitting. O algoritmo de gradient boosting constrói sequencialmente árvores de decisão, onde cada nova árvore é treinada para corrigir os erros das anteriores, resultando em um modelo progressivamente mais preciso.

4.4.2 MÉTRICAS DE DESEMPENHO

As métricas quantitativas obtidas no conjunto de teste (dados de 2024) demonstram desempenho satisfatório do modelo operacional para aplicação prática. O MAE (Mean Absolute Error) de 8,35 horas indica que, em média, as predições do modelo diferem em

aproximadamente 8 horas e 20 minutos do tempo real de estadia observado. Considerando que o tempo médio de estadia no sistema portuário brasileiro varia entre 30 e 60 horas dependendo do tipo de operação, esse erro representa desvio percentual aceitável para planejamento logístico.

O RMSE (Root Mean Squared Error) de 29,53 horas, valor superior ao MAE, indica que o modelo ocasionalmente produz erros maiores, mas ainda dentro de faixa administrável. A diferença entre RMSE e MAE sugere presença de outliers ou casos atípicos onde as predições se distanciam mais significativamente dos valores reais, situação esperada dado que operações portuárias podem ser afetadas por eventos imprevisíveis como condições meteorológicas adversas, greves ou problemas técnicos não capturados nas features disponíveis.

O R^2 (Coeficiente de Determinação) de 0,78 demonstra que o modelo consegue explicar 78% da variância no tempo de estadia, resultado expressivo para problema de regressão em domínio tão complexo quanto operações portuárias, esse valor indica que a maior parte da variabilidade nos tempos de estadia pode ser atribuída às variáveis incluídas no modelo.

O MAPE (Mean Absolute Percentage Error) de 28,33% indica o erro médio do modelo em relação aos valores reais observados. À primeira vista, esse percentual pode parecer alto, mas se analisarmos em operações com tempo de estadia curto, por exemplo, cerca de 10 horas, um erro de apenas 3 horas já representa 30% em termos percentuais, embora, na prática, seja uma diferença pequena e aceitável para o planejamento portuário. Já em estadias mais longas, esse erro percentual tende a ser menor.

É válido ressaltar que esses resultados representam desempenho realista do modelo operacional, que utiliza apenas informações disponíveis no momento da chegada do navio. Conforme discutido na metodologia, o modelo inicial que incluía variáveis com vazamento de dados apresentava métricas artificialmente superiores (MAE próximo a 5 horas e R^2 de 0,87), mas era inaplicável na prática.

4.5 WEBSITE

A interface web desenvolvida serve como canal de acesso às funcionalidades do sistema, oferecendo experiência visual e intuitiva. Ao acessar o sistema após autenticação, o usuário é apresentado à página inicial que organiza os três módulos principais.

O **módulo Dashboard** permite que gestores logísticos, analistas portuários e outros usuários explorem métricas operacionais através de visualizações interativas. Esse módulo foi criado para auxiliar na tomada de decisão estratégica. A possibilidade de aplicar múltiplos filtros simultaneamente torna a ferramenta versátil para diferentes contextos de análise, desde visões macro do sistema portuário nacional até investigações mais detalhadas.

O **módulo Predição** disponibiliza o modelo de aprendizado de máquina para previsão de tempo de estadia. Essa funcionalidade contribui significativamente para planejamento mais assertivo das operações portuárias, permitindo que gestores antecipem a duração total da permanência de embarcações, informação crucial para alocação eficiente de berços, dimensionamento de equipes operacionais e coordenação com transportadores terrestres.

O **módulo Metodologia** oferece informações sobre o desenvolvimento do sistema, tecnologias empregadas e fundamentos metodológicos do trabalho. Essa seção inclui links para acesso aos dados oficiais da ANTAQ, promovendo transparência e permitindo que usuários interessados explorem as fontes primárias de informação. A disponibilização de documentação técnica e contexto metodológico contribui para a credibilidade do sistema e facilita a compreensão dos fundamentos das análises e predições oferecidas.

A funcionalidade de gerenciamento de conta, acessível através do menu de usuário na tela inicial, permite que usuários mantenham controle sobre seus dados. A interface de edição de perfil possibilita atualização de nome, email e senha, com validação de senha opcional, caso o campo seja deixado em branco, a senha anterior é mantida. A operação de exclusão de conta foi implementada com mecanismo de confirmação para prevenir exclusões acidentais, e realiza limpeza completa da sessão após remoção do registro do banco de dados, seguindo princípios de privacidade e controle de dados pessoais estabelecidos por regulamentações como a LGPD.

Figura 14 - Tela inicial do PortoVision.



Fonte: Sistema desenvolvido pela autora.

Durante o desenvolvimento, foi dada atenção à responsividade da interface, garantindo que o sistema funcione bem e mantenha boa aparência tanto em computadores quanto em smartphones. O layout se ajusta automaticamente ao tamanho da tela, reorganizando elementos e adaptando menus e visualizações.

Figura 15 - Tela inicial do PortoVision Mobile.



Fonte: Sistema desenvolvido pela autora.

5 CONCLUSÃO

O presente trabalho apresentou o desenvolvimento do **PortoVision**, um sistema de apoio à decisão voltado à análise e otimização do transporte marítimo brasileiro, construído a partir de dados públicos disponibilizados pela ANTAQ. Além da análise exploratória de dados, o sistema reúne visualizações interativas e um modelo preditivo baseado em aprendizado de máquina, evidenciando o potencial da ciência de dados como ferramenta estratégica para modernizar a gestão portuária no país.

Os objetivos definidos no início do projeto foram atingidos, apesar dos prazos de desenvolvimento fugirem um pouco do escopo planejado. A análise exploratória permitiu compreender com profundidade o comportamento das operações portuárias ao longo de 15 anos, revelando padrões, tendências e sazonalidades importantes. O dashboard oferece visualizações claras que apoiam a tomada de decisões estratégicas. Já o modelo preditivo fornece estimativas de tempo de estadia que contribuem para um planejamento mais assertivo das operações. A integração desses elementos em um sistema web acessível e intuitivo mostra que é possível transformar dados brutos em inteligência operacional. Este trabalho contribui para o avanço do uso de técnicas de ciência de dados e aprendizado de máquina nesse setor, um área que, apesar de sua relevância para a economia nacional, ainda apresenta amplas oportunidades de modernização.

As limitações encontradas durante o desenvolvimento não comprometem a validade dos resultados, mas apontam caminhos claros para aperfeiçoamentos futuros. A dependência de dados estruturados da ANTAQ, embora assegure transparência e reprodutibilidade, limita a inclusão de variáveis potencialmente relevantes que não constam nos registros oficiais, como condições meteorológicas específicas ou problemas técnicos específicos em terminais. A validação do sistema foi feita principalmente com métricas quantitativas e testes funcionais, sem a etapa de validação extensiva com usuários reais, o que poderia oferecer feedback importante sobre usabilidade e adequação às necessidades práticas de gestores portuários. Além disso, o desenvolvimento enfrentou restrições de hardware, especialmente quanto à capacidade de processamento e memória. Esses pontos indicam caminhos importantes para trabalhos futuros, que poderão incorporar novas fontes de dados, melhorar o desempenho computacional e validar o sistema em ambiente operacional real.

O sistema desenvolvido não substitui o julgamento humano especializado, mas o amplifica, fornecendo informações fundamentadas, análises e projeções baseadas em padrões históricos que o auxiliam a agir com maior assertividade e confiança. Os resultados obtidos evidenciam que investimentos em inteligência operacional baseada em dados representam um caminho promissor para o fortalecimento da competitividade do setor portuário nacional no comércio internacional.

5.1 TRABALHOS FUTUROS

O desenvolvimento do PortoVision abre espaço para diversas possibilidades de aprimoramento e expansão. Em trabalhos futuros, o sistema pode evoluir tanto em termos técnicos quanto funcionais, ampliando sua utilidade e impacto no setor portuário.

Uma das principais frentes de avanço é a integração de novas fontes de dados, como informações meteorológicas, registros de manutenção, feriados, greves e características detalhadas das embarcações. Esses dados adicionais poderiam tornar as previsões mais precisas e permitir análises mais completas sobre fatores que influenciam o tempo de estadia.

Também há potencial para expandir o escopo das predições, desenvolvendo modelos específicos para estimar tempos de espera, operação e desatracação separadamente. Isso tornaria o sistema ainda mais útil para diferentes públicos, como autoridades portuárias e operadores de terminais. Do ponto de vista técnico, o uso de modelos mais avançados de aprendizado de máquina, como ensembles, redes neurais profundas e aprendizado por transferência, pode aumentar a precisão e a robustez das predições. Em paralelo, melhorias na interface e experiência do usuário, como relatórios automáticos e alertas inteligentes, tornariam o sistema mais dinâmico e adaptado às necessidades reais.

Em resumo, o PortoVision estabelece uma base na qual diversas direções de pesquisa e desenvolvimento podem ser exploradas. A evolução contínua do sistema, incorporando avanços tecnológicos, feedback de usuários e conhecimento especializado do domínio portuário, tem potencial de transformá-lo em uma ferramenta estratégica para gestão do transporte marítimo brasileiro, contribuindo para o fortalecimento da competitividade logística nacional.

REFERÊNCIAS

AGÊNCIA BRASIL. **Balança: queda de importações e commodities contribuiu para saldo recorde.** 1 ago. 2016. Disponível em: <<https://agenciabrasil.ebc.com.br/economia/noticia/2016-08/saldo-recorde-da-balanca-deve-s-e-queda-de-importacoes-e-commodities>>. Acesso em: 3 nov. 2025.

ALMEIDA MARTINS, Ariele de; LACERDA, Fabiana Praxedes; SANTOS, João Almeida. Problemas da logística portuária brasileira. São Paulo, 24–25 nov. 2021. Disponível em: <https://www.fateczl.edu.br/engetec/engetec_2021/4_EnGeTec_paper_55.pdf>. Acesso em: 21 jun. 2025.

ALVES, F. **Caracterização da oferta de contêineres no transporte marítimo brasileiro de longo curso.** Digital Library of Theses and Dissertations (Universidade de São Paulo), 5 jun. 2020. Disponível em: <<https://doi.org/10.11606/d.3.2020.tde-20012021-162540>>. Acesso em: 4 nov. 2025.

ANDRADE, Bernardo; BIAZON, Tássia. **O transporte marítimo e o uso sustentável do oceano.** Jornal da USP, São Paulo, 06 out. 2021. Disponível em: <<https://jornal.usp.br/artigos/o-transporte-maritimo-e-o-uso-sustentavel-do-oceano/>>. Acesso em: 7 abr. 2025.

ANTAQ. **Competências institucionais da Agência Nacional de Transportes Aquaviários,** 2020. Disponível em: <https://www.gov.br/antag/pt-br/acesso-a-informacao/institucional/copy_of_competencias>. Acesso em: 10 nov. 2025.

ANTAQ. **Dados Abertos.** Disponível em: <<https://www.gov.br/antag/pt-br/acesso-a-informacao/dados-abertos>>. Acesso em: 8 abr. 2025.

ANTAQ. **Estatístico Aquaviário 2023.** Disponível em: <<https://web3.antag.gov.br/ea/sense/movport.html#pt>>. Acesso em: 8 abr. 2025.

ANTAQ. **Formulário para Proposição de Ato Normativo**. Brasília: ANTAQ, 2017. Disponível em:

<<https://www.gov.br/antag/pt-br/acao-a-informacao/acoes-e-programas/governanca-regulatoria/02FormulrioparaProposiodeAtoNormativoGRM.pdf>>. Acesso em: 5 nov. 2025.

ANTAQ. **Granéis sólidos e líquidos atingem maior movimentação para o mês de abril**. Disponível em:

<<https://www.gov.br/antag/pt-br/noticias/2025/graneis-solidos-e-liquidos-atingem-maior-movimentacao-para-o-mes-de-abril>>. Acesso em: 20 mai. 2025.

ANTAQ. **Relatório anual de 2016 – Ouvidoria**. [S.l.: s.n.], 2016. Disponível em: <https://www.gov.br/antag/pt-br/canais_atendimento/relatrioanualdeouvidoria2016.pdf>. Acesso em: 3 nov. 2025.

ANTAQ. **Relatório detalha impactos da pandemia da Covid-19 no transporte marítimo**. 28 jul. 2022. Disponível em: <<https://www.gov.br/antag/pt-br/noticias/2022/relatorio-detalha-impactos-da-pandemia-da-covid-19-no-transporte-maritimo>>. Acesso em: 3 nov. 2025.

ANUÁRIO Estatístico Aquaviário - 2016: **Apresentação**. [S.l.: s.n.], 2016. Disponível em: <http://sophia.antag.gov.br/index.asp?codigo_sophia=27572>. Acesso em: 3 nov. 2025.

BANCO CENTRAL DO BRASIL. **Relatório de inflação – dezembro 2016**. Brasília, v. 18, n. 4, p. 1-55, dez. 2016. Disponível em: <<https://www.bcb.gov.br/content/ri/relatorioinflacao/201612/RELINF201612-ri201612P.pdf>>. Acesso em: 31 out. 2025.

BRASIL. Casa Civil. **Com R\$ 54,9 bilhões em investimentos, Novo PAC acelera modernização dos portos e apresenta resultados na Câmara dos Deputados**. 20 ago. 2025. Disponível em: <<https://www.gov.br/casacivil/pt-br/assuntos/noticias/2025/agosto/com-r-54-9-bilhoes-em-investimentos-novo-pac-acelera-modernizacao-dos-portos-e-apresenta-resultados-na-camara-dos-deputados>>. Acesso em: 5 nov. 2025.

BRASIL. Ministério da Agricultura e Pecuária. **Portos do arco norte representam 31,6% das exportações de milho e soja, em março.** 2024. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/noticias/portos-do-arco-norte-representam-31-6-das-exportacoes-de-milho-e-soja-em-marco>>. Acesso em: 12 nov. 2025.

BRASIL. Ministério de Portos e Aeroportos. **Criada no Brasil, associação internacional vai fortalecer a logística e a gestão entre portos mundiais.** Disponível em: <<https://www.gov.br/portos-e-aeroportos/pt-br/assuntos/noticias/2024/10/criada-no-brasil-associacao-internacional-vai-fortalecer-a-logistica-e-a-gestao-entre-portos-mundiais>>. Acesso em: 25 mai. 2025.

BRASIL. Ministério dos Portos e Aeroportos. **Portos brasileiros registram maior movimentação da história com 1,32 bilhão de toneladas em 2024.** Brasília, 18 fev. 2025. Disponível em: <<https://www.gov.br/portos-e-aeroportos/pt-br/assuntos/noticias/2025/02/portos-brasileiros-registram-maior-movimentacao-da-historia-com-132-bilhao-de-toneladas-em-2024>>. Acesso em: 5 nov. 2025.

BRASIL. Ministério dos Transportes. **BR do Mar: sancionada lei que institui programa de incentivo à cabotagem no Brasil.** Publicado em 10 jan. 2022. Disponível em: <<https://www.gov.br/transportes/pt-br/assuntos/noticias/2022/01/br-do-mar-sancionada-lei-que-institui-programa-de-incentivo-a-cabotagem-no-brasil>>. Acesso em: 9 abr. 2025.

BRASIL. Secretaria de Comunicação Social. **Brasil registra recorde na movimentação portuária em 2024, com 1,32 bilhão de toneladas.** Publicado em: 18 fev. 2025. Disponível em: <<https://www.gov.br/secom/pt-br/assuntos/noticias/2025/02/brasil-registra-recorde-na-movimentacao-portuaria-em-2024-com-1-32-bilhao-de-toneladas>>. Acesso em: 10 nov. 2025.

BRASIL. Secretaria de Comunicação Social. **Investimentos federais impulsionam escoamento da safra agrícola.** 2023. Disponível em: <<https://www.gov.br/secom/pt-br/assuntos/noticias/2023/07/investimentos-federais-impulsionam-escoamento-da-safra-agricola>>. Acesso em: 12 nov. 2025.

CARVALHO, Maria Paula. **Jogos Rio-2016: no meio do caminho, uma crise política e financeira.** UOL Notícias, RFI, 17 ago. 2023. Disponível em: <<https://noticias.uol.com.br/ultimas-noticias/rfi/2023/08/17/jogos-rio-2016-no-meio-do-caminho-uma-crise-politica-e-financeira.htm>>. Acesso em: 31 out. 2025.

CASTRO, Jose Roberto. **As commodities e seu impacto na economia do Brasil.** Nexo Jornal, 31 mar. 2016. Atualizado em 28 dez. 2023. Disponível em: <<https://www.nexojornal.com.br/explicado/2016/03/31/As-commodities-e-seu-impacto-na-economia-do-Brasil>>. Acesso em: 3 nov. 2025.

CNN BRASIL. **Transporte entre portos brasileiros bate recorde de movimentação em julho.** 15 set. 2025. Disponível em: <https://www.cnnbrasil.com.br/economia/investimentos/transporte-entre-portos-brasileiros-bate-recorde-de-movimentacao-em-julho/>. Acesso em: 12 nov. 2025.

CURY, Anay; SILVEIRA, Daniel. **PIB recua 3,6% em 2016, e Brasil tem pior recessão da história.** G1, 7 mar. 2017. Disponível em: <<https://g1.globo.com/economia/noticia/pib-brasileiro-recua-36-em-2016-e-tem-pior-recessao-da-historia.ghtml>>. Acesso em: 31 out. 2025.

Druzdzal, M. J.; Flynn, R. R. (2002). **Decision support systems.** In A. Kent (Ed.), Encyclopedia of Library and Information Science (pp. 1-15). New York: Marcel Dekker.

EQUIPE TOTVS. **BR do Mar: entenda os detalhes da lei que visa agilizar a logística nacional.** Disponível em: <<https://www.totvs.com/blog/gestao-logistica/br-do-mar/>>. Acesso em: 09 abr. 2025.

EQUIPE TOTVS. **Logística portuária: o que é, importância, principais desafios e carreira no setor.** 22 fev. 2024. Disponível em: <<http://totvs.com/blog/gestao-logistica/logistica-portuaria/>>. Acesso em: 5 nov. 2025.

EQUIPE TOTVS. **Logística portuária: o que é, importância, principais desafios e carreira no setor.** Disponível em: <<https://www.totvs.com/blog/gestao-logistica/logistica-portuaria/>>. Acesso em: 12 jun. 2025.

FARIELLO, D.; ORDOÑEZ, R. **Movimento no Porto do Rio cai três vezes mais que a média nacional.** O Globo, 26 mar. 2017. Disponível em: <<https://oglobo.globo.com/economia/movimento-no-porto-do-rio-cai-tres-vezes-mais-que-media-nacional-21115030>>. Acesso em: 3 nov. 2025.

GAMMA, E. et al. **Design Patterns: Elements of Reusable Object-Oriented Software.** Reading: Addison-Wesley, 1995. Disponível em: <<https://www.javier8a.com/itc/bd1/articulo.pdf>>. Acesso em: 7 nov. 2025.

GUIA, C. **O que é Tailwind CSS.** Disponível em: <<https://tailwindcss.com.br/guia-tailwind/o-que-e-tailwind-css>>. Acesso em: 5 dez. 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Em 2016, PIB chega a R\$ 6,3 trilhões e cai 3,3% em volume.** Agência de Notícias, Estatísticas Econômicas, 09 nov. 2018. Disponível em: <<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/22936-em-2016-pib-chega-a-r-6-3-trilhoes-e-cai-3-3-em-volume>>. Acesso em: 31 out. 2025.

JORNAL BE NEWS. **Arco Norte x Porto de Santos.** Sindaport, 28 jul. 2023. Disponível em: <<https://sindaport.com.br/conteudo-pesquisa.php?id=27721>>. Acesso em: 3 nov. 2025.

KUEHNE+NAGEL. **Transporte marítimo: nossas soluções – Brazil.** Disponível em: <<https://br.kuehne-nagel.com/pt/-/servicos/transporte-maritimo>>. Acesso em: 8 abr. 2025.

LIU, Yuxi (Hayden). **Machine learning com Python na prática: desbloqueie as melhores práticas de aprendizado de máquina com casos de uso do mundo real.** 1. ed. São Paulo: Blucher, 2025. E-book. Disponível em: <<https://plataforma.bvirtual.com.br>>. Acesso em: 10 nov 2025.

MITCHELL, T. M. **Machine Learning.** New York: McGraw-Hill, 1997. Disponível em: <<https://www.cs.cmu.edu/~tom/files/MachineLearningTomMitchell.pdf>>. Acesso em: 24 out. 2025.

MOLITZAS, Ricardo. **Desafios logísticos dos portos brasileiros.** IBL – Instituto Brasil Logística, [s.l.], 26 mar. 2024. Disponível em: <<https://ibl.org.br/artigo-desafios-logisticos-dos-portos-brasileiros/>>. Acesso em: 5 nov. 2025.

NOTTEBOOM, T.; RODRIGUE, J. P.; Espaço Aberto; PPGG -UFRJ, V. **Regionalização Portuária: Rumo a uma Nova Fase no Desenvolvimento Portuário**. Port Regionalization: Towards a New Phase in Port Development. n. 2, p. 123–146, 2014. Disponível em: <<https://dialnet.unirioja.es/descarga/articulo/5522528.pdf>>. Acesso em: 7 nov. 2025.

OLIVEIRA, Rodrigo; WILDNER, Maria; PRETTO, Fabrício. **Técnicas de visualização de informações como apoio à gestão estratégica**. Revista Destaques Acadêmicos, v. 10, n. 1, 2018. Disponível em: <<https://www.univates.br/revistas/index.php/destaques/article/download/1778/1319/4500>>. Acesso em: 7 nov. 2025.

PORTO DE SANTOS AUTORIDADE PORTUÁRIA. **Infraestrutura portuária**. Disponível em: <<https://www.portodesantos.com.br/conheca-o-porto/infraestrutura-portuaria/>>. Acesso em: 3 nov. 2025.

PORTO DE SANTOS AUTORIDADE PORTUÁRIA. **Porto de Santos mantém crescimento e registra recordes novamente em setembro**. 28 out. 2025. Disponível em: <<https://www.portodesantos.com.br/2025/10/28/porto-de-santos-mantem-crescimento-e-registra-recordes-novamente-em-setembro/>>. Acesso em: 3 nov. 2025.

PORTO, P. C. DE S.; SOUZA FILHO, V. J. DE. Porto e desenvolvimento regional. **DRd - Desenvolvimento Regional em debate**, v. 14, p. 133-148, 22 maio 2024. Doi: <<https://doi.org/10.24302/drd.v14.3656>>. Acesso em: 3 nov. 2025.

POWER, D. J. **Decision Support Systems: Concepts and Resources for Managers**. Westport: Quorum Books, 2002. Disponível em: <<https://scholarworks.uni.edu/cgi/viewcontent.cgi?article=1066&context=facbook>>. Acesso em: 7 nov. 2025.

QUINTELLA, Marcus. **Desafios e oportunidades nos portos brasileiros**. Revista Mundo Logística, 23 jan. 2024. Disponível em: <<https://mundologistica.com.br/artigos/desafios-e-oportunidades-nos-portos-brasileiros>>. Acesso em: 7 nov. 2025.

RAZZOLINI FILHO, Edelvino. **Transporte e modais: com suporte de ti e si**. 1. ed. Curitiba: Intersaberes, 2012. E-book. Disponível em: <<https://plataforma.bvirtual.com.br>>. Acesso em: 31 out 2025.

REDAÇÃO PORTOGENTE. **Kuehne+Nagel usa tecnologia e IA para otimizar a produtividade da logística**. Disponível em: <<https://portogente.com.br/noticias/transporte-logistica/116648-kuehne-nagel-usa-tecnologia-e-ia-para-otimizar-a-produtividade-da-logistica>> . Acesso em: 09 abr. 2025.

REUTERS. **Maersk diz que interrupção de transporte no Mar Vermelho gera efeitos globais**. CNN Brasil, 17 jul. 2024. Disponível em: <<https://www.cnnbrasil.com.br/economia/negocios/maersk-diz-que-interruptao-de-transporte-no-mar-vermelho-gera-efeitos-globais/>>. Acesso em: 8 abr. 2025.

SANTOS, Nadyne. **Indicadores de desempenho de portos públicos e privados: estudo de caso de terminais de contêineres brasileiros**. Dissertação (Mestrado em Engenharia Civil) - Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2019. Disponível em: <<http://repositorio.ufpa.br/jspui/handle/2011/12045>>. Acesso em: 3 nov. 2025.

SHNEIDERMAN, B. **The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations**. In: IEEE SYMPOSIUM ON VISUAL LANGUAGES, 1996, Boulder. Anais [...]. Boulder: IEEE, 1996. p. 336-343. Disponível em: <<https://www.cs.umd.edu/~ben/papers/Shneiderman1996eyes.pdf>>. Acesso em: 12 nov. 2025.

STOPFORD, Martin. **Economia marítima**. 1. ed. São Paulo: Blucher, 2017. E-book. Disponível em: <<https://plataforma.bvirtual.com.br>>. Acesso em: 08 nov 2025.

STREAMLIT INC. **Streamlit documentation**. [S.l.: s.n.], 2025. Disponível em: <<https://docs.streamlit.io/>>. Acesso em: 7 nov. 2025.

TASHMAN, L. J. **Out-of-sample tests of forecasting accuracy: an analysis and review**. International Journal of Forecasting, v. 16, n. 4, p. 437-450, 2000. Disponível em: <https://www.researchgate.net/publication/223319987_Out-of-sample_tests_of_forecasting_accuracy_An_analysis_and_review>. Acesso em:

TECHNOMAR. **A tecnologia a favor do desenvolvimento do setor aquaviário e portuário no Brasil.** 18 set. 2024. Disponível em:

<<https://www.technomar.com.br/institucional/a-tecnologia-a-favor-do-desenvolvimento-do-setor-aquaviario-e-portuario-no-brasil/>>. Acesso em: 7 jun. 2025.

TUKEY, J. W. **Exploratory Data Analysis.** Reading: Addison-Wesley, 1977. Disponível em:

<<https://www.consoleflare.com/blog/wp-content/uploads/2022/09/Exploratory-Data-Analysis-1977-John-Tukey.pdf>>. Acesso em: 5 nov. 2025.

UNCTAD. **Performance-indicators for port and maritime transportation.** [S.l.]:

UNCTAD, 2023. Disponível em:

<https://unctad.org/system/files/official-document/dtltlb2023d2_en.pdf>. Acesso em: 10 nov. 2025.

UNCTAD. **Review of Maritime Transport 2023.** Disponível em:

<<https://unctad.org/publication/review-maritime-transport-2023>>. Acesso em: 9 abr. 2025.

VIEIRA, Guilherme Bergmann Borges. **Transporte internacional de cargas.** 2. ed. São Paulo: Aduaneiras, 2011. 150 p.

WICKHAM, H. **Tidy Data.** *Journal of Statistical Software*, v. 59, n. 10, p. 1-23, 2014.

Disponível em: <<https://www.jstatsoft.org/article/view/v059i10>>. Acesso em: 17 out. 2025.

WINDWARD. **Maritime Logistics.** Disponível em:

<<https://windward.ai/glossary/what-is-maritime-logistics/>>. Acesso em: 6 nov. 2025.

WS_ADMIN. [INFOGRÁFICO] **Principais desafios da logística portuária brasileira.** 12

fev. 2019. Disponível em: <<https://wilsonsons.com.br/pt-br/blog/logistica-portuaria/>>. Acesso em: 20 jun. 2025.