

Modelando Ocorrências de Espécies Marinhas com Classificação Uma-Classe: um Estudo de Caso sobre a Caravela-Portuguesa

Henrique Margotte
Departamento de Informática
Universidade Federal do Paraná
Curitiba, PR, Brasil
hmargotte@inf.ufpr.br

Carmem Satie Hara
Departamento de Informática
Universidade Federal do Paraná
Curitiba, PR, Brasil
carmem@inf.ufpr.br

Aurora Trinidad Ramirez Pozo
Departamento de Informática
Universidade Federal do Paraná
Curitiba, PR, Brasil
aurora@inf.ufpr.br

Abstract

This study explores the application of One-Class Support Vector Machines (OC-SVM) for predicting occurrences of Portuguese Man-of-War (*Physalia physalis*) along the Brazilian coast. The model leverages daily wind data and citizen science sightings to identify environmental conditions associated with the presence of this species. Given the imbalance in the dataset — where only positive instances are available — we frame the problem as a one-class classification task. The proposed approach achieved over 70% accuracy and 64% F1 Score on two test sets, demonstrating its feasibility despite limitations in temporal and spatial feature representation. The study highlights the importance of data quality, sampling strategies, and feature extraction methods, suggesting directions for future research. These include incorporating seasonal and geospatial information and exploring alternative architectures. The findings contribute to the growing field of environmental monitoring using machine learning, especially for rare or underreported species.

Keywords

Monitoramento Ambiental, Aprendizado de Máquina, *One-Class Classification*, *One-Class SVM*, Caravela-Portuguesa

1 Introdução

Cnidário da espécie *Physalia physalis*, a caravela-portuguesa (Fig. 1) destaca-se tanto pela sua coloração vibrante quanto pelo perigo representado pelas células urticantes em seus tentáculos, sendo uma das principais causas de acidentes nos litorais brasileiros [1]. A escassez de dados sobre essa espécie, causada por sua dispersão em alto-mar e pela limitação nos meios de registro, aliada à sua suscetibilidade à força do vento — indicada por especialistas como um dos principais fatores que afetam seu deslocamento, devido ao pneumatóforo flutuante — dificultam o seu estudo.

Com o avanço tecnológico e o aumento da capacidade de processamento computacional, a Inteligência Artificial (IA) e o Aprendizado de Máquina (AM) têm se consolidado como ferramentas promissoras para resolver diversos problemas sociais e ambientais. No contexto da previsão de caravelas-portuguesas, Margotte [3] apresenta uma abordagem baseada em AM, envolvendo a análise do problema, a criação de uma base de dados e a condução de experimentos preliminares, em contraposição a estudos baseados em simulações hidrodinâmicas.

Uma das principais dificuldades para a modelagem preditiva é a escassez e o viés nos dados disponíveis. A coleta por meio de ciência cidadã [4] surge como alternativa viável, embora ainda apresente limitações. Por exemplo, os registros disponíveis referem-se exclusivamente à presença de caravelas-portuguesas, não havendo



Figura 1: Foto de uma caravela-portuguesa [2].

dados confiáveis sobre a sua ausência. Isso resulta em uma base desbalanceada, composta por apenas uma classe, o que inviabiliza a aplicação direta de modelos tradicionais de classificação. Nesses casos, técnicas de Classificação Uma-Classe (OCC) são mais apropriadas, pois permitem o treinamento do modelo apenas com exemplos positivos.

Este trabalho propõe o uso de técnicas de AM para prever o aparecimento de caravelas-portuguesas no litoral brasileiro. Para isso, utiliza-se uma Máquina de Vetores de Suporte Uma-Classe (OC-SVM), treinada a partir de padrões de vento observados nos períodos e regiões de avistamento registrados. O objetivo é identificar sequências de velocidades e direções de vento que se assemelhem aos padrões associados aos eventos de aparição.

O modelo é avaliado com base em métricas obtidas na classificação de dados reais e de dados negativos gerados via amostragem negativa¹. Os testes indicam desempenho promissor, com acurácia superior a 70%. Além disso, são discutidas propostas para trabalhos futuros, visando aprimorar o sistema preditivo e elevar a robustez dos resultados.

Os códigos desenvolvidos e bases de dados utilizadas serão disponibilizados no *GitHub* do projeto (<https://github.com/RESMA-PPGINF-UFPR-CAPES-PRINT>), para garantir a reprodutibilidade do experimento e contribuir para pesquisas relacionadas.

Este artigo está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 descreve a construção da base de dados; o modelo e o protocolo experimental são discutidos na Seção 4, com os resultados apresentados na Seção 5. As perspectivas para trabalhos futuros estão na Seção 6, e as conclusões na Seção 7.

¹Técnica em que uma amostra dos dados é sintetizada a partir do conjunto total possível.

2 Trabalhos Relacionados

O estudo da trajetória da caravela-portuguesa ainda é escasso na literatura científica, principalmente devido à dificuldade de obtenção de dados precisos sobre a posição desses organismos em alto-mar. Considerando a importância da predição antecipada de sua aparição em praias e a necessidade de compreender seu deslocamento, alguns trabalhos buscam reconstruir trajetórias com base em avistamentos costeiros por meio de técnicas de *backtracking* [5, 6]. Outros utilizam registros obtidos antes do encalhe, como medições em estações de monitoramento fora da costa [6], ou adotam pontos de partida conhecidos, quando disponíveis [7].

Todos os estudos identificados fazem uso de simuladores hidrodinâmicos para modelar o transporte de partículas no oceano, com a inclusão de um coeficiente de arrasto de vento na superfície do oceano, além das forças marítimas já presentes nos modelos. Contudo, nenhum deles considera especificamente o cenário do litoral brasileiro. Essa lacuna é abordada neste trabalho, conforme sistematizado na Tabela 1.

Problemas de OCC ocorrem quando as bases de dados apresentam majoritariamente — ou exclusivamente — elementos de uma única classe [8]. Essa situação é comum em tarefas de detecção de anomalias, nas quais os modelos são treinados para aprender o comportamento típico da classe predominante e identificar desvios como anomalias [9].

Hayashi et al. [10] apresenta uma revisão abrangente sobre técnicas de OCC e suas aplicações em diferentes domínios. Entre os métodos discutidos, destaca-se a Máquina de Vetores de Suporte (SVM) adaptada para OCC, denominada OC-SVM. Apesar da existência de variantes mais recentes, o OC-SVM ainda se mostra eficaz, especialmente por sua menor exigência de dados e poder computacional em comparação com modelos baseados em aprendizado profundo. Por essas razões, é uma escolha apropriada para estudos exploratórios e de fundamentação inicial em problemas de OCC.

3 Base de Dados

Margotte [3] apresenta uma base de dados integrada (ilustrada na Fig. 2) composta por registros de avistamentos de caravelas-portuguesas, obtidos por meio de ciência cidadã, e por dados atmosféricos de velocidade e direção do vento coletados via satélite. A base contém 425 registros de avistamentos realizados entre 2007 e 2022.

Para cada avistamento, foram associados vetores de velocidade do vento nos eixos norte-sul e leste-oeste, obtidos da coordenada mais próxima em uma grade espacial com resolução de $0,125^\circ$ (aproximadamente 12,5 km). Esses dados foram extraídos para a data do avistamento e para os quatro dias anteriores.

É importante observar que, em cenários reais de previsão — nos quais se deseja antecipar o aparecimento de caravelas em dias futuros —, é necessário estimar os vetores de vento para datas além do último dado disponível. Uma abordagem possível é o uso de modelos de AM, especialmente Redes Neurais Recorrentes (RNNs), para prever as componentes do vento a partir de séries históricas. Essa técnica pode ser integrada ao sistema de predição de caravelas, desde que a base de dados inclua uma janela temporal maior, como os dez dias anteriores ao instante de previsão, conforme sugerido

por trabalhos relacionados [11]. A Fig. 3 exemplifica a integração de um sistema de previsão de vento à previsão de caravelas.

4 Protocolo Experimental

O modelo OC-SVM é uma adaptação da SVM tradicional para tarefas de detecção de anomalias, operando de forma não supervisionada. Ambas as abordagens projetam os dados de entrada em um espaço de características, utilizando funções de *kernel* para permitir separações não lineares [12]. No entanto, enquanto a SVM convencional é treinada com exemplos rotulados de múltiplas classes, a OC-SVM assume que todos os dados de treinamento pertencem a uma única classe, considerada positiva. O modelo aprende uma função de decisão que define uma região no espaço vetorial que abrange a maior parte dos dados fornecidos, tratando como anomalias os pontos que se situam fora dessa região [13].

Um dos principais hiperparâmetros do modelo é o valor de nu , que controla simultaneamente a fração máxima de pontos de treinamento considerados como anomalias e a fração mínima de vetores de suporte. Os vetores de suporte são as observações mais próximas da fronteira de decisão aprendida, e são os principais responsáveis por definir essa região no espaço de características. A Fig. 4 ilustra a delimitação dessa fronteira utilizando um *kernel* RBF (*Radial Basis Function*).

Neste estudo, foi utilizada a implementação do modelo OC-SVM disponível na biblioteca *Scikit-Learn*, por meio da classe *OneClassSVM*, mantendo-se os hiperparâmetros em seus valores padrão. O modelo adota o *kernel* RBF, adequado para detectar padrões não lineares nos dados, com os parâmetros γ definido como *scale* e nu igual a 0,5. O objetivo da aplicação foi avaliar se os vetores de vento associados aos registros de avistamentos de caravelas-portuguesas apresentam padrões recorrentes que possam ser aprendidos por esse tipo de modelo.

A base de dados foi dividida em dois cenários experimentais distintos:

- **TR-21 (Corte em 2021):** a base de treino contém todos os registros até o final de 2021 (395 registros), e a base de teste é composta pelos 30 registros de 2022.
- **TR-75 (Divisão 75/25):** a base de treino compreende os 75% registros mais antigos (318 registros), enquanto os 25% mais recentes (107 registros) compõem a base de teste.

Em ambos os cenários, o OC-SVM foi treinado exclusivamente com dados positivos (i.e., avistamentos reais de caravelas), buscando aprender os padrões típicos associados à ocorrência desses eventos.

Como o modelo OC-SVM é treinado apenas com dados positivos, foi necessário construir uma base de teste com rótulos binários (1 = presença, 0 = ausência) para avaliação dos resultados. Para isso, aplicou-se uma técnica de amostragem negativa: para cada registro real da base de teste, foi criado um exemplo sintético espelhado, invertendo-se os sinais das componentes de vento (multiplicação por -1). O pressuposto adotado é que vetores de vento na direção oposta aos observados dificultariam ou inviabilizariam a chegada da caravela à costa. Assim, cada par de registros gerado para teste possui um exemplo positivo (avistamento) e um negativo (espelhado).

Para poder visualizar a distribuição dos dados gerados, foi realizada uma redução de dimensionalidade das bases de teste utilizando

Trabalho	Na Costa	Previsão	Aprendizado de Máquina	Mar Aberto
Macías et al. [7]	X	X		
Headlam et al. [6]		X		X
Ferrer and Pastor [5]	X			X
Trabalho proposto	X	X	X	X

Tabela 1: Comparação entre trabalhos sobre a trajetória de caravelas-portuguesas quanto à utilização de dados costeiros, previsão, aplicação de aprendizado de máquina e foco em mar aberto.

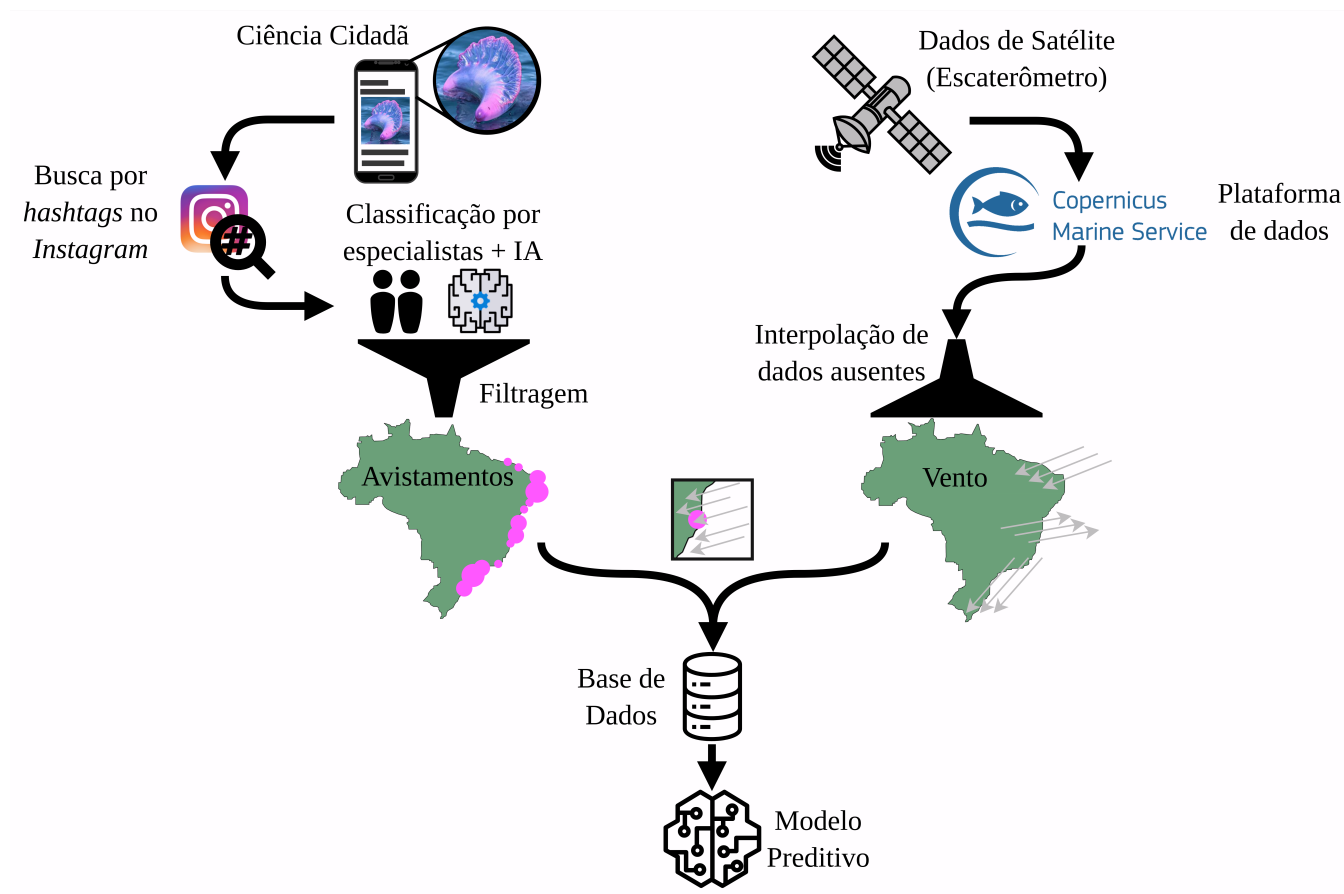


Figura 2: Diagrama representativo do processo de construção da base de dados integrada.

o algoritmo t-SNE. Os dados estão representados na Fig. 5 e em duas dimensões, onde pontos magenta representam exemplos positivos (avistamentos) e azuis representam negativos (ausência). A distribuição simétrica ilustra o comportamento dos vetores gerados por espelhamento.

Para a avaliação dos modelos, foram utilizadas quatro métricas de avaliação: acurácia, precisão, *recall* e *F1 Score*. A acurácia mede a taxa de acerto geral do modelo em ambas as classes. A precisão calcula a porcentagem de acerto do modelo dentre todos os classificados como positivos, enquanto o *recall* mede o acerto dentre todos os elementos que possuem a classe real positiva. O *F1 Score* é a média harmônica das duas últimas métricas. Como o problema é de

classificação binária, serão consideradas apenas as métricas em relação à classe positiva, ou seja, a identificação correta de avistamentos de caravelas-portuguesas.

5 Resultados

Após o treinamento do modelo, este foi utilizado para prever se as sequências de vento dos registros da base de teste indicavam ou não a presença de uma caravela, sendo os resultados comparados aos rótulos definidos. O teste de **TR-75** apresentou acurácia de 70,09%, precisão de 80,28%, *recall* de 53,27% e *F1 Score* de 64,04%, enquanto o de **TR-21** apresentou acurácia de 71,67%, precisão de 76%, *recall* de 63,33% e *F1 Score* de 69,09%, conforme apresentado na Tabela 2. Essa

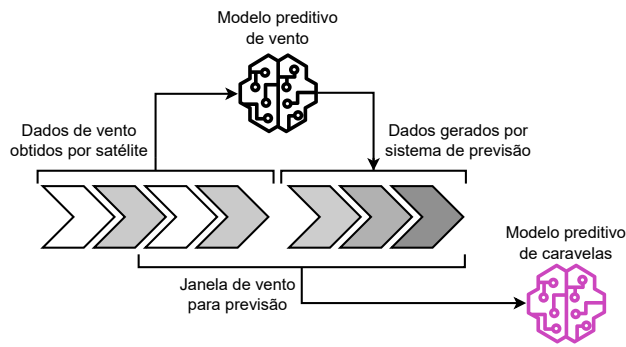


Figura 3: Esquema representando a integração de um sistema de previsão de vento para aumento da base de dados de previsão de caravelas-portuguesas.

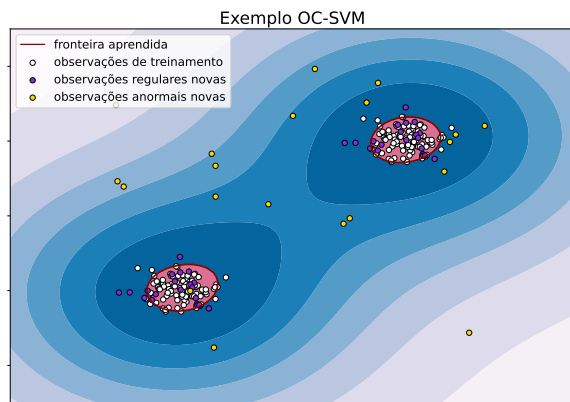


Figura 4: Representação do espaço vetorial de um OC-SVM não linear, com *kernel RBF*. Gerado por código do *Scikit-Learn*.

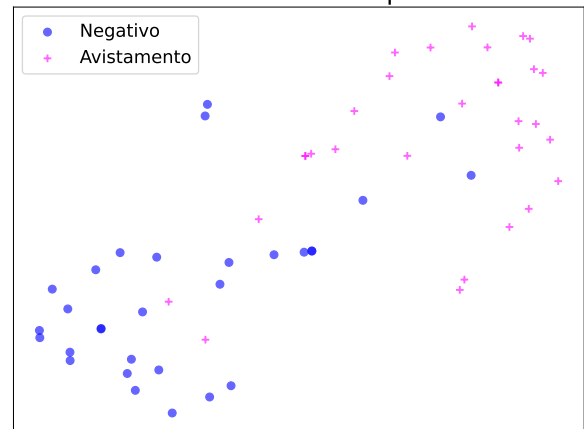
Teste	Acurácia	Precisão	Recall	F1 Score
TR-21	71,67%	76%	63,33%	69,09%
TR-75	70,09%	80,28%	53,27%	64,04%

Tabela 2: Resultado das métricas de avaliação calculadas para a classe positiva (avistamentos) nos testes em TR-21 e TR-75.

diferença pode estar relacionada à quantidade de dados disponíveis para o treinamento: a divisão **TR-21** disponibilizou um conjunto de treino maior, cobrindo um período mais extenso, o que possivelmente permitiu ao modelo generalizar melhor os padrões de vento associados à presença de caravelas, gerando acurácia, *recall* e *F1 Scores* superiores. Em contraste, a **TR-75** forneceu um volume menor de dados para o aprendizado, o que pode ter limitado a capacidade do modelo de capturar esses padrões de forma abrangente.

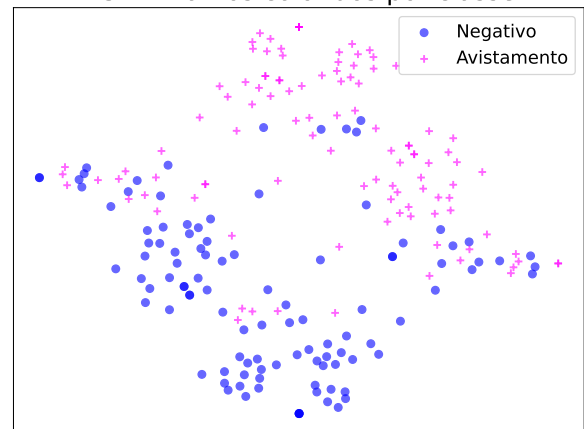
Por não ter envolvimento com a classe negativa, o *recall* se mostra uma métrica importante para avaliar a eficácia do modelo em identificar corretamente os avistamentos de caravelas. Comparando

t-SNE: Pontos coloridos por classe



(a) TR-21 (Corte em 2021).

t-SNE: Pontos coloridos por classe



(b) TR-75 (Divisão 75/25).

Figura 5: Visualização bidimensional das bases de teste via t-SNE. Os positivos magenta indicam registros de avistamento de caravelas (classe 1), e os círculos azuis representam amostras negativas obtidas por espelhamento (classe 0).

com o estimado de uma classificação totalmente aleatória, que geraria métricas próximas a 50%, o TR-75, com apenas 53,27% de *recall*, teve pouca melhora em classificar corretamente a classe positiva, enquanto o TR-21 obteve resultado superior de 63,33%, representando melhora considerável à técnica aleatória e salientando a importância de uma base de dados mais robusta para treinamento. Ainda, as demais métricas, todas superiores a 64%, indicam eficiência do modelo em evitar falsos positivos, identificando casos de ausência de caravelas com mais precisão que um sistema aleatório.

A Fig. 6 apresenta os erros e acertos do modelo nas bases de teste, a partir da mesma técnica de redução de dimensionalidade utilizada na Fig. 5, com os resultados plotados de acordo com as previsões. Observa-se nos gráficos a proximidade entre alguns registros de classes diferentes, o que pode ter contribuído para erros de classificação. Esse padrão indica que os dados de vento podem

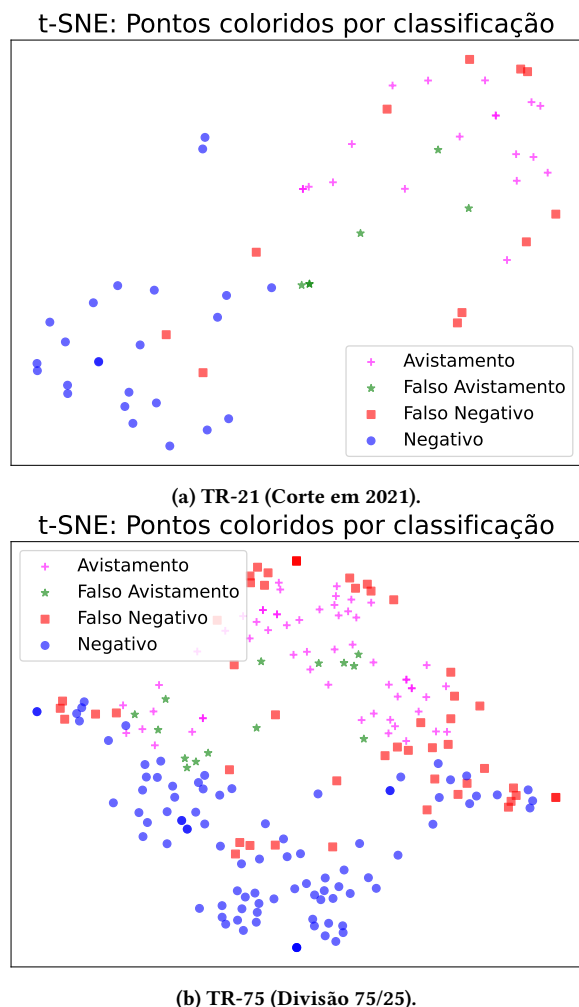


Figura 6: Gráficos de representação bidimensional das classes atribuídas pelo modelo OC-SVM nas bases de teste de TR-21 (6a) e TR-75 (6b). Positivos magenta indicam classificações corretas de presença de caravela; círculos azuis, classificações corretas de ausência; quadrados vermelhos indicam presença de caravela classificada erroneamente como ausência; estrelas verdes representam ausência classificada como presença.

ser insuficientes para prever a presença de caravelas com alta precisão, ou que o método de amostragem negativa empregado não seja o mais adequado para validação. Algumas alternativas a essas limitações são discutidas na Seção 6.

A Fig. 7 apresenta as matrizes de confusão para as duas divisões, permitindo a análise do desempenho do modelo em cada classe. Em ambas as divisões, observa-se um viés do modelo em prever a ausência de caravelas, resultando em maior acurácia nessa classe. Por outro lado, os falsos negativos (avistamentos classificados como ausências) ocorrem com maior frequência, sugerindo baixa sensibilidade para eventos positivos. A pequena diferença entre classificações corretas e incorretas nos registros de avistamento

reforça a limitação do modelo em capturar padrões positivos consistentes. Isso é esperado, dado que a amostragem negativa afeta apenas a etapa de validação e não contribui diretamente para o aprendizado do OC-SVM.

Devido a ausência de trabalhos abordando o mesmo problema através de AM e no cenário brasileiro, conforme apresentado na Seção 2, não foi possível realizar comparações com outros trabalhos, destacando também a característica de inovação do presente artigo. Um estudo comparativo ainda pode ser executado ao simular os mesmos dados em sistemas hidrodinâmicos, porém a instalação, adaptação e operação desses sistemas se mostrou além do proposto por este trabalho.

6 Trabalhos Futuros

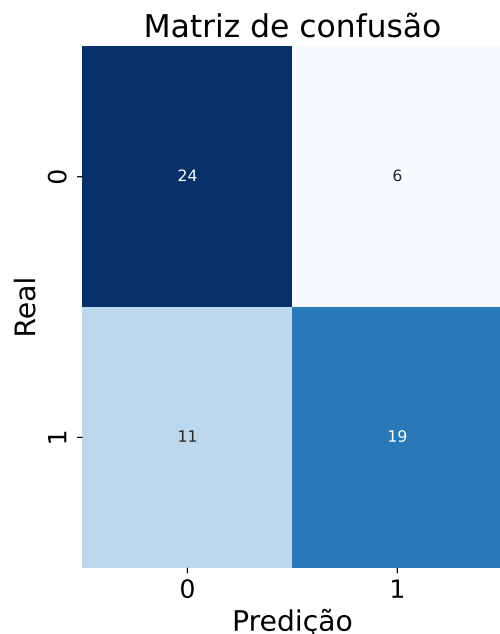
Em Margotte [3] são discutidas abordagens que podem aprimorar os resultados obtidos neste trabalho, destacando-se a extração de características temporais e espaciais para inclusão no treinamento. A dimensão temporal, como observado em modelos de previsão de vento baseados em séries temporais [11], é importante para que o modelo aprenda relações sequenciais nas entradas — o que não ocorre na abordagem atual com OC-SVM, cuja estrutura não contempla dinâmicas temporais. A dimensão espacial, por sua vez, refere-se ao uso de múltiplas coordenadas ao redor do ponto de avistamento e à exploração da correlação espacial entre elas, o que pode influenciar a modelagem ao representar trajetórias ou regiões de origem das caravelas-portuguesas.

Hayashi et al. [10] discute diferentes técnicas de extração de características para problemas de OCC, incluindo o uso de Autocodificadores (AEs), como também proposto por Margotte [3]. Embora AEs sejam eficazes na extração de representações compactas, sua principal função é a reconstrução de dados, o que os torna menos apropriados para tarefas de classificação. Há modelos de aprendizado profundo especificamente desenvolvidos para extração de características voltadas à classificação, que podem ser mais eficazes nesse contexto.

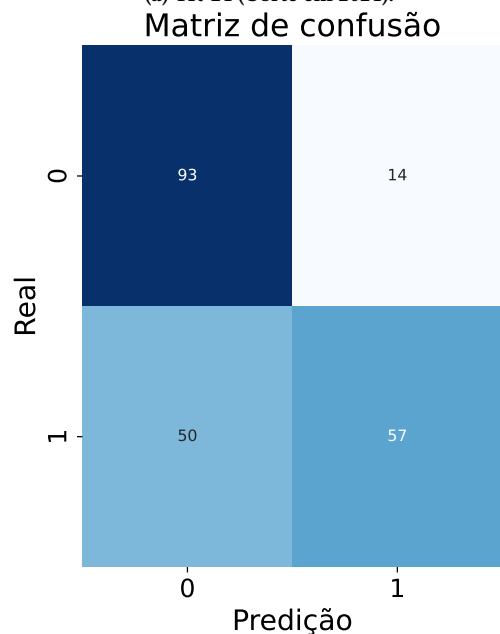
Além da dimensão temporal e espacial dos dados, outras variáveis podem ser relevantes para a predição da ocorrência de caravelas-portuguesas. Do ponto de vista temporal, a própria data pode exercer influência, uma vez que se observa uma maior concentração de avistamentos em determinados períodos do ano. Sob o aspecto espacial, as coordenadas geográficas podem ser úteis para distinguir comportamentos distintos entre regiões, como o Sul e o Nordeste do Brasil. Além disso, a distância entre o ponto de avistamento e a coordenada de vento mais próxima pode ser um fator adicional relevante na caracterização espacial da amostra.

Outra possibilidade a ser explorada consiste em tratar os próprios avistamentos de caravelas como uma série temporal, na qual a ocorrência de eventos passados influencia a probabilidade de eventos futuros. Essa abordagem permitiria investigar, por exemplo, se a presença de uma caravela em uma determinada localização aumenta a chance de novos avistamentos na mesma região em dias subsequentes, sugerindo um comportamento de propagação ou persistência no tempo.

Hayashi et al. [10] também aborda alternativas para problemas de OCC baseadas na criação de classes artificiais para possibilitar a aplicação de técnicas tradicionais de classificação supervisionada.



(a) TR-21 (Corte em 2021).



(b) TR-75 (Divisão 75/25).

Figura 7: Matrizes de confusão comparando a quantidade de acertos e erros do modelo OC-SVM para cada classe nas bases de teste de TR-21 (7a) e TR-75 (7b). Linhas indicam os rótulos reais; colunas, as classes atribuídas. Cores mais escuras representam maior concentração de registros.

Nesse contexto, os dados gerados por amostragem negativa — utilizados neste trabalho para fins de validação — poderiam ser empregados como uma classe adicional, transformando o problema em uma tarefa de classificação binária. No entanto, a eficácia dessa estratégia depende de uma fundamentação teórica sólida para garantir que os exemplos negativos representem adequadamente a ausência da classe de interesse durante o treinamento.

Diversas outras arquiteturas, técnicas de modelagem e configurações de hiperparâmetros poderiam ser exploradas na abordagem deste problema. Adicionalmente, variáveis ambientais como temperatura da superfície do mar, correntes marítimas, concentração de plâncton e composição química da água podem fornecer informações complementares para a predição de caravelas-portuguesas. Contudo, a adoção de modelos mais sofisticados e a incorporação de múltiplas variáveis aumentam significativamente a complexidade do sistema, exigindo maior volume de dados e métodos de coleta consistentes — um desafio considerável no uso de dados reais de campo.

7 Conclusão

Este trabalho investigou o uso de técnicas de AM para a predição de ocorrências de caravelas-portuguesas a partir de dados de vento e registros de ciência cidadã. Considerando a configuração desbalanceada da base de dados disponível, na qual apenas a presença da espécie é registrada, foram abordadas técnicas de OCC, com foco na utilização do modelo OC-SVM como solução inicial para o problema.

Os resultados indicaram acurácia superior a 70% e *F1 Score* superior a 64% nos testes realizados, com um destes atingindo *recall* de 63,33%, evidenciando o potencial da abordagem baseada em OC-SVM. No entanto, observou-se limitações importantes, como a dependência da consistência e completude dos dados de entrada, bem como a ausência de características temporais e espaciais mais ricas no modelo atual. Além disso, destacou-se a relevância de uma amostragem negativa representativa e bem fundamentada para a validação do modelo, com impacto direto nas métricas obtidas, e de uma base de dados de avistamento robusta para o treinamento do modelo de AM.

Este estudo contribuiu com a aplicação pioneira do OC-SVM em dados ambientais para a modelagem de ocorrências esparsas de caravelas-portuguesas, ressaltando os desafios e potenciais dessa abordagem. Foram discutidas direções promissoras para o aprimoramento do modelo, incluindo o uso de técnicas de extração de características que incorporem trajetórias espaciais e padrões sazonais, a integração de outras variáveis ambientais relevantes e o desenvolvimento de estratégias mais robustas para geração de amostras negativas.

Acknowledgments

Este trabalho contou com suporte parcial da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Programa de Excelência Acadêmica (PROEX). Os autores também reconhecem o uso de ferramentas baseadas em inteligência artificial, como o modelo *ChatGPT*, para apoio na redação e revisão textual do manuscrito.

Referências

- [1] Mayana Mendes E Silva Cavalcante, Zulimar Marita Ribeiro Rodrigues, Rachel Ann Hauser-Davis, Salvatore Siciliano, Haddad Junior, Jorge Luiz Silva Nunes, et al. Health-risk assessment of portuguese man-of-war (*physalia physalis*) envenomations on urban beaches in sao luis city, in the state of maranhao, brazil. *Revista Da Sociedade Brasileira De Medicina Tropical*, 2020.
- [2] A Carneiro, LS Nascimento, MA Noernberg, CS Hara, and ATR Pozo. Social media image classification for jellyfish monitoring. *Aquatic Ecology*, 58(1):3–15, 2024.
- [3] Henrique Margotte. Aprendizado de máquina para previsão do aparecimento de caravelas-portuguesas no litoral brasileiro através do vento. Master's thesis, Universidade Federal do Paraná, 2024.
- [4] Lorena Silva do Nascimento, Carmem Satie Hara, Miodeli Nogueira Júnior, and M Noernberg. Instagram como fonte de dados alternativa no monitoramento da# caravelaportuguesa (*physalia physalis*, cnidaria). In *Livro de Memórias do IV SUSTENTARE e VII WIPIS: Workshop Internacional de Sustentabilidade, Indicadores e Gestão de Recursos Hídricos. Anais Piracicaba (SP) Online*. https://www.even3.com.br/anais/sustentare_wipis_2022/584935, 2022.
- [5] Luis Ferrer and Ane Pastor. The portuguese man-of-war: Gone with the wind. *Regional Studies in Marine Science*, 14:53–62, 2017. ISSN 2352-4855. doi: <https://doi.org/10.1016/j.rsma.2017.05.004>. URL <https://www.sciencedirect.com/science/article/pii/S2352485517300075>.
- [6] Jasmine Headlam, Kieran Lyons, Jon Kenny, Eamonn Lenihan, Declan Quigley, William Helps, Michel Dugon, and Thomas Doyle. Insights on the origin and drift trajectories of portuguese man of war (*physalia physalis*) over the celtic sea shelf area. *Estuarine Coastal and Shelf Science*, 246, 10 2020. doi: 10.1016/j.ecss.2020.107033.
- [7] Diego Macías, Laura Prieto, and Elisa García-Gorrioz. A model-based management tool to predict the spread of *physalia physalis* in the mediterranean sea, minimizing risks for coastal activities. *Ocean Coastal Management*, 212:105810, 2021. ISSN 0964-5691. doi: <https://doi.org/10.1016/j.ocecoaman.2021.105810>. URL <https://www.sciencedirect.com/science/article/pii/S0964569121002933>.
- [8] Shehroz S. Khan and Michael G. Madden. One-class classification: taxonomy of study and review of techniques. *The Knowledge Engineering Review*, 29(3): 345–374, 2014. doi: 10.1017/S026988891300043X.
- [9] Yanmiao Li, Yingying Xu, Yankun Cao, Jiangang Hou, Chun Wang, Wei Guo, Xin Li, Yang Xin, Zhi Liu, and Lizhen Cui. One-class lstm network for anomalous network traffic detection. *Applied Sciences*, 12(10), 2022. ISSN 2076-3417. doi: 10.3390/app12105051. URL <https://www.mdpi.com/2076-3417/12/10/5051>.
- [10] Toshitaka Hayashi, Dalibor Cimr, Hamido Fujita, and Richard Cimler. Critical review for one-class classification: recent advances and the reality behind them, 2024. URL <https://arxiv.org/abs/2404.17931>.
- [11] Henrique Margotte, Aurora Pozo, and Carmem Hara. Predicting Oceanic Wind Speed and Direction Using LSTM. In *Proceedings XXV GEOINFO*, São José dos Campos, SP, Brazil, 2025.
- [12] Mokhtar Mohammadi, Tarik A. Rashid, Sarkhel H.Taher Karim, Adil Husain Mohammed Aldalwie, Quan Thanh Tho, Moazam Bidaki, Amir Masoud Rahmani, and Mehdi Hosseinzadeh. A comprehensive survey and taxonomy of the svm-based intrusion detection systems. *Journal of Network and Computer Applications*, 178:102983, 2021. ISSN 1084-8045. doi: <https://doi.org/10.1016/j.jnca.2021.102983>. URL <https://www.sciencedirect.com/science/article/pii/S1084804521000102>.
- [13] Y. Wang, J. Wong, and A. Miner. Anomaly intrusion detection using one class svm. In *Proceedings from the Fifth Annual IEEE SMC Information Assurance Workshop, 2004.*, pages 358–364, 2004. doi: 10.1109/IAW.2004.1437839.