NOVA
IMS
Information
Management
School

MGI

Master Degree Program in

Information Management

Applying Machine Learning Models for Decision Support in Maritime Surveillance

Predicting Illegal Activities in National Sovereignty Areas

André Francisco Taveira Seixas Nunes

Project Work

presented as partial requirement for obtaining the Master Degree in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Applying Machine Learning Models for Decision Support in Maritime Surveillance

Predicting Illegal Activities in National Sovereignty Areas

by

André Francisco Taveira Seixas Nunes

Project Work presented as a partial requirement for obtaining the Master's degree in Information Management, with a specialisation in Business Intelligence

Supervised by

Bruno Jardim, PhD, Nova Information Management School

July, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

[Lisbon, 15th July 2025]

André Francisco Taveira Seixas Nunes

DEDICATION

To my family, for their constant support. None of this would be possible without you. Thank you.

ACKNOWLEDGEMENTS

I want to express my sincere gratitude to everyone who supported me throughout this work.

I would like to especially thank the Portuguese Navy for the opportunity to carry out this project as part of my master’s degree and for the trust placed in me.

I would like to sincerely thank my supervisor, Bruno Jardim, for his invaluable guidance, support, and encouragement throughout this journey. Your insight and dedication were truly fundamental to the development of this work.

To my family, thank you for your steady support, patience, and unconditional belief in me.

Finally, I would like to thank all those who, personally or professionally, contributed to this achievement. Your support, in small and meaningful ways, has meant more than words can express.

ABSTRACT

This work explores how historical maritime inspection data can improve operational decision-making in maritime surveillance by applying machine learning models to predict the likelihood of illegal activities. In Portugal, the selection of patrol areas continues to be based on the experience of commanders and static views from previous inspections, which limits the efficiency of resource allocation in extensive maritime zones such as the Exclusive Economic Zone (EEZ). To address this challenge, a data-driven approach was adopted, using inspection data collected by the Portuguese Navy between 2014 and 2024. The dataset included over 15,000 records with spatial and temporal features such as inspection location, month, period of day, day of the week, and vessel type. Four classification models were tested—Naïve Bayes, Logistic Regression, k-Nearest Neighbour, and Classification Trees—following the CRISP-DM methodology. Model performance was evaluated using precision, recall, accuracy, F1-score, ROC-AUC, and Brier score, considering balanced and imbalanced training scenarios. Results showed that although the Classification Tree model achieved the highest precision (0.71), it had a very low recall, which indicates a conservative but limited detection capability. On the other hand, k-NN achieved a better balance with higher recall but lower precision. The results confirm that historical inspection data contains predictive value and could be applied to a decision-support system to assist in the selection of a patrol area. Probability maps generated from model outputs highlight high-risk zones and offer a practical tool for prioritising enforcement actions. Despite limitations related to data imbalance and geographic distribution of inspections, the proposed approach demonstrates the feasibility of improving maritime surveillance using structured data and machine learning. Future developments could include the integration of real-time data sources, such as AIS and weather conditions, and the use of more complex models to support wider maritime authority operations.

KEYWORDS

Maritime Surveillance; Machine Learning; Decision Support Systems; Illegal Activities Prediction; Spatiotemporal Data Analysis

Sustainable Development Goals (SDG):

14 LIFE BELOW WATER

16 PEACE, JUSTICE AND STRONG INSTITUTIONS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AIS	Automatic Identification System
ANN	Artificial Neural Networks
BI	Business Intelligence
CFAR	Constant False Alarm Rate
CFR	Community Fleet Register
CNN	Convolutional Neural Networks
CRISP-DM	Cross-Industry Standard Process for Data Mining
CT	Classification Tree
DTG	Date-Time-Group
EEZ	Exclusive Economic Zone
FISCREP	Fiscalization Report
FN	False Negative
FP	False Positive
FPR	False Positive Rate
GDH	Group-Date-Hour
GIS	Geographic Information System
GLM	Generalised Linear Model
IMO	International Maritime Organisation
IPS	Illegal Parking Score
ISAR	Inverse Synthetic Aperture Radar
IUU	Illegal, Unreported, and Unregulated
k-NN	k-Nearest Neighbours
LR	Logistic Regression
LSTM	Long Short-Term Memory
LWIR	Long-wavelength infrared

ML	Machine Learning
MONICAP	*Sistema de monitorização contínua da atividade da pesca* - Continuous Monitoring System for Fishing Activity (Portuguese acronym)
MS	Maritime Surveillance
MSA	Maritime Situational Awareness
NB	Naïve Bayes
NLP	Natural Language Processing
NN	Neural Network
NSDT	Novel Ship Detection and Tracking
PBSPCC	Patrol Boat Scheduling Problem with Complete Coverage
RCAAP	*Repositórios Científicos de Acesso Aberto de Portugal* - Open Access Scientific Repositories of Portugal (Portuguese acronym)
RF	Random Forest
ROC-AUC	Receiver Operating Characteristic - Area Under Curve
SADAP	*Sistema de Apoio à Decisão na Actividade de Patrulha* - Decision Support System for Patrol Activities (Portuguese acronym)
SAR	Synthetic Aperture Radar
SIFICAP	*Sistema de Fiscalização e Controlo das Actividades da Pesca* - Fishing Activities Inspection and Control System (Portuguese acronym)
SIPM	*Sistema de Informação da Polícia Marítima* - Maritime Police Information System (Portuguese acronym)
SOLAS	Safety of Life at Sea
ST-MM-JFKA	Skeleton theory-based multiple model joint feature knowledge association
SVM	Supported Vector Machine
SWIR	Short-wavelength infrared
TN	True Negative
TP	True Positive
TPR	True Positive Rate

TSP	Travelling Salesman Problem
UAV	Unmanned Aerial Vehicle
USV	Unmanned Surface Vehicle
VMS	Vessel Monitoring System
YOLOv3	You Only Look Once version 3

1. INTRODUCTION

1.1. BACKGROUND AND CONTEXT

Maritime security has become one of the central concerns of international security discourse (Bueger & Edmunds, 2017). Bueger and Edmunds (2017) divide maritime security into four interconnected and interdependent domains, ranging from national security to external threats, to the importance of the blue economy and human security. Given the new realities and challenges related to maritime security, states have been fighting against illegalities, safeguarding human life at sea, and protecting the marine environment (Decree Law No. 43/2002, 2002). To achieve this, Portuguese Navy ships, coordinated by the National Maritime Authority within the maritime authority system, have the mission of supervising the maritime areas under national sovereignty or jurisdiction to enforce state authority at sea (Decree Law No. 185/2014, 2014). These maritime surveillance (MS) missions are of "utmost importance to ensure the safe use of the seas and maritime border security" (Ramongassie et al., 2010, p. 3793). A key area is the Exclusive Economic Zone (EEZ), where the state has the right to explore, manage, and conserve natural resources in the waters above the seabed, and in Portugal, the smallest of the three subareas (the mainland) covers more than 287,000 km^2 (Direção-Geral de Recursos Naturais, Segurança e Serviços Marítimos, 2025). Given the importance of this task and the vast area it covers, it is crucial to optimise the allocation of resources for surveillance and inspection activities.

1.2. RESEARCH GAP AND OBJECTIVES

In recent years, numerous studies have focused on MS systems that utilise radar, air or naval assets, satellite capabilities, or maritime databases (Şengül et al., 2023). However, these systems mainly serve to increase situational awareness and detect irregularities in real time (Şengül et al., 2023), rather than supporting the proactive planning of maritime patrols. In Portugal, patrol area selection is still largely based on the experience of commanders and manual analysis of historical inspection data available in systems such as *Sistema de Fiscalização e Controlo das Actividades da Pesca* (SIFICAP) and *Sistema de Apoio à Decisão na Actividade de Patrulha* (SADAP) (da Rocha Rei, 2016). Although these platforms provide relevant vessel information, they lack an integrated analytical model to assist in selecting vessels or areas for inspection, often leading to suboptimal resource allocation, especially considering the size of the area of responsibility.

While real-time detection systems, such as Automatic Identification System (AIS), Synthetic aperture radar (SAR), and satellite imagery (Bernabé et al., 2024; Liu et al., 2024), are crucial for operational awareness, they require continuous data feeds and expensive infrastructures (Chen et al., 2022; Chircop et al., 2022). Moreover, they are reactive by nature. In contrast, historical inspection data, such as those from the Portuguese Navy's Fiscalization Report (FISCREP) records, offer structured and georeferenced information on vessel typologies, infractions, and context (da Rocha Rei, 2016; Miguel et al., 2024; Moura et al., 2024). This data can reveal spatiotemporal patterns of non-compliance that support predictive and proactive patrol planning. Previous work has demonstrated the value of this approach: Wen et al. (2012) showed that inspection records can outperform human judgment in predicting smuggling, and Jardim et al. (2022) applied a similar logic to illegal parking.

Building on these insights, this research explores how data from previous inspections can be used to predict risk zones and support data-driven patrol decisions in a maritime context.

Therefore, this project aims to answer the question:

"In what ways can historical inspection data enhance the selection of operational areas in maritime surveillance using machine learning models?"

This aligns with the objective:

"*To test machine learning models for the development of a decision-support system for maritime surveillance, aiming to predict illegal activities in areas under national sovereignty based on spatiotemporal conditions."*

1.3. METHODOLOGICAL APPROACH

In this project, the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology (Schröer et al., 2021) was used. This methodology encompasses several phases: understanding and preparing data, modelling, evaluation, and deployment. The data to be used consists of inspection reports carried out by the Portuguese Navy between 2014 and 2024. It will use information about the type of vessel inspected and spatiotemporal conditions (month, period of the day, day of the week and geographic area) to predict the probability of a possible illegal activity. We will study four ML algorithms: Naïve Bayes (NB), Logistic Regression (LR), k-Nearest Neighbour (k-NN) and Classification Trees (CT).

Through this methodology, we expect to support the use of ML models as a basis for a decision-support system that provides the conditional probability of the occurrence of illegal activities in the Portuguese mainland EEZ, subdivided into a grid.

1.4. THESIS STRUCTURE SUMMARY

This project is organised into five chapters. The first chapter introduces the research subject and context, identifies the research gap, describes the research question and objective, and provides a summary of the methodology adopted. The second chapter presents a literature review that supports the identification of the gap. The third chapter describes all methodological steps in detail. Subsequently, the fourth chapter presents and analyses the results obtained. Finally, the fifth chapter concludes the work by summarising the main results, highlighting the limitations found, and proposing directions for future research.

2. LITERATURE REVIEW

In this chapter, we present the state of the art in MS knowledge, starting with a broad overview and gradually narrowing down to its practical and specific applications within the Portuguese context.

To understand the current state of the art in the study of MS, the Scopus[1] database was consulted. This database of peer-reviewed literature (scientific journals, books and conference proceedings) offers a broad overview of worldwide scientific production in the areas of science and technology. The search in Scopus was carried out using keywords that best defined the scope of this study: “maritime surveillance”, “maritime inspection” and “maritime patrol”. The search was limited to publications from 2021 to 2024, in English, and from journals classified at least in Quartile 2 in their respective fields. This criterion was adopted to ensure that only highly reputable sources, with significant impact in their respective areas, were considered.

Subsequently, VOSviewer[2] 1.6.20 software was used to analyse the most significant keywords (document and author) that appeared at least ten times. This research identified five clusters, organised by the number of connections (see Figure 2.1). The different clusters found were used to systematise the selection of relevant articles and studies and to structure the following chapters.

The following clusters were found:

- **Cluster 1** - Automation and General Detection (red): Focused on automation, Automatic Identification System (AIS), monitoring, and maritime transportation, emphasising real-time maritime monitoring and anomaly detection using AIS.
- **Cluster 2** - Synthetic aperture radar (SAR) and Neural Network (NN) (green): Dominated by SAR, deep learning, radar imaging, and object detection, showcasing radar integration with NN for advanced maritime image analysis.
- **Cluster 3** - Radar Detection (blue): Focuses on tracking radar, clutter (information theory), target tracking, and antennas, reflecting techniques to improve radar-based detection and tracking in cluttered maritime environments.
- **Cluster 4** - Remote Sensing and maritime vehicles (yellow): Highlights remote sensing, space-based radar, satellite imagery, and maritime vehicles, focusing on satellite-based technologies for global MS.

---

[1] Available at www.scopus.com

[2] Available at www.vosviewer.com

- **Cluster 5** - Learning Algorithms (purple): Primarily represented by learning algorithms, focusing on leveraging ML to support automation and advanced maritime analyses across clusters.

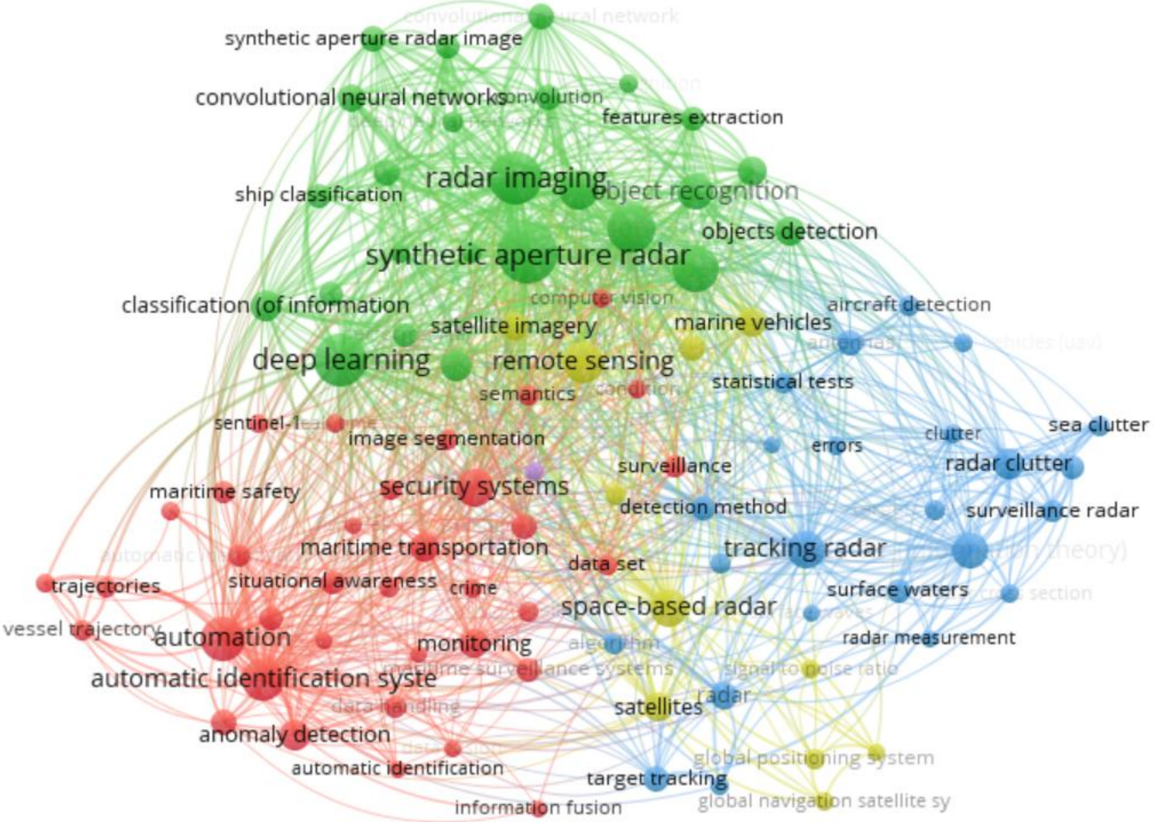


Figure 2.1 – Scopus keywords network created with VOSviewer.

After the cluster analysis, the bibliography was chosen for further reading based on the information contained in their abstracts.

Given that we intend to apply this research to the Portuguese Navy, an additional search was carried out in the Portuguese open-access scientific repository "*Repositórios Científicos de Acesso Aberto de Portugal"* (RCAAP)[3], using the Portuguese keyword "*Fiscalização Marítima*". In addition to the bibliography found through this process, others were used based on cross-reference.

2.1. MARITIME SURVEILLANCE OVERVIEW

The Maritime Situational Awareness (MSA) is the capability to achieve information superiority in the maritime environment, enabling a shared understanding of the maritime situation to

---

[3] Available at www.rcaap.pt

enhance the planning and execution of operations, and the MS is a key component of it (Soldi et al., 2021). MS plays a crucial role in ensuring the safety and security of the seas (Ramongassie et al., 2010), being crucial in areas such as fisheries monitoring, law enforcement and national security policies (Soldi et al., 2021). In the same scope, the theme of maritime security has gained renewed interest in globalisation and national strategies (Bueger & Edmunds, 2017), considering the increase of maritime security threats (Gamage et al., 2023).

With the growing need to improve MS capabilities, numerous efforts have been made in recent years to improve new technologies, data storage and processing to increase global surveillance (Chaumette, 2016). These advancements aim to enable ship monitoring, detect illegal activities, classify vessels, and optimise maritime and air patrol operations, thus maximising the use of available data.

2.1.1. Automation and General Detection

As evidenced by the research in Scopus, one of the most studied subjects is the use of AIS data as a source of information, assuming an essential role for MS (Bernabé et al., 2024). According to the International Maritime Organisation (IMO) guidelines, established by the 1980 Safety of Life at Sea (SOLAS) convention, certain types and sizes of vessels are required to be equipped with this system to increase the safety of maritime traffic (Liu et al., 2024). AIS devices provide vessel tracking through static (e.g., MMSI, ship name, type, and size) and dynamic (e.g., position, speed, and course) data (Liu et al., 2024). This spatiotemporal data enables advanced analytics in surveillance systems through supervised and unsupervised ML on the MS domain, for various purposes (Dästner et al., 2018; Newaliya & Singh, 2021). The use of this system is not new, for instance, Ramongassie et al. (2010) discuss a dual sensor constellation that integrates radar and AIS technologies, leveraging the strengths of both sensors and improving vessel detection and tracking.

The main applicability of using the AIS data is anomaly detection. That fact is evident in the work of Ribeiro et al. (2023), who, through a systematic literature review of articles from 2009 to 2020, present a taxonomy based on ML. Their study reveals that the most used data-driven approach uses ML algorithms to collect AIS data and detect patterns, as well as new AIS data to analyse and identify deviations.

In the same context, Liu et al. (2024) reference several studies suggesting that the increase in data related to this system has allowed the expansion of applications, especially for anomaly detection, trajectory clustering, trajectory prediction, maritime traffic monitoring, vessel behaviour analysis and environmental evaluation. In their research, Liu et al. (2024) present a systematic framework for maritime kinematic anomaly detection, introducing a novel approach to classify and detect anomalies in vessel movements. For instance, Do Nascimento et al. (2023) evaluated several ML models to detect fishing activity in vessel trajectories, aiming for high precision in MS. The study evaluated different data approaches, with the Long

Short-Term Memory (LSTM) model achieving 100% accuracy but being slow, while the Random Forest (RF) model showed 87% accuracy and was 1,000 times faster. The trade-off between accuracy and speed is crucial for naval inspections, where LSTM is suitable for high-risk scenarios requiring precision, and RF is more practical for real-time processing. Pohontu et al. (2023) in their study explore AI-based vessel type classification using trajectories, effectively identifying 12 main classes, including non-cooperative vessels engaged in specialised activities such as fishing, dredging, and search and rescue using handwriting signature verification methodology. Considering integration with other systems, Morando et al. (2023) study the applicability of a fusion algorithm that matches the vessels' positions provided by AIS with Radio Frequency emitters and satellite ship detections. Bernabé et al. (2024) in their work aim to detect suspicious trajectories of vessels with intentional AIS shutdowns in near-real-time using self-supervised deep learning, achieving 99.5% accuracy in predicting the receipt of AIS messages. From another perspective, Agrawal and Ramteke (2024) used AIS data to approach MS as a dynamic Travelling Salesman Problem (TSP), optimising routes to monitor 172 vessels.

2.1.2. SAR and Neural Networks

Based on the cluster analysis, the second most studied subject is the SAR. This technology uses radio waves to capture images of the Earth's surface (Wang et al., 2021). In this matter, Guida et al. (2023) presented initial results from the Nereus project, using combined AIS/SAR data to investigate illegal, unreported, and unregulated (IUU) fishing activities in the Mauritius EEZ, highlighting the need for improved detection methods and satellite solutions for more effective fisheries monitoring. The study conducted by Yan et al. (2023) proposes a SAR vessel classification method using the SMOTEBoost algorithm and AIS data, achieving 93% accuracy in classifying cargo, tanker, and fishing vessels in SAR images by transferring AIS data, with superior performance in precision, recall, and F1-Score metrics.

Looking only at SAR technology, Ghosh et al. (2021) present a three-stage methodology (raw data preprocessing, ship-like object detection using Constant False Alarm Rate (CFAR) algorithm and reduction stage) adapted to detect vessel-like objects using medium-resolution satellite images. Wang et al. (2021) propose a vessel detection method using SAR satellite imagery with a Novel Ship Detection and Tracking (NSDT) preprocessing (filtering, edge detection, and segmentation) to enhance image quality and improve ML model performance, achieving superior accuracy and metrics compared to models without preprocessing with supported vector machine (SVM) achieving 95% accuracy. Bhattacharjee et al. (2023) presents a state-of-the-art architecture and a deep learning-based algorithm for vessel localisation with SAR images to balance the accuracy, speed, and computational cost. K et al. (2023) developed an autonomous system for vessel classification using satellite images and advanced deep learning models such as You Only Look Once version 3 (YOLOv3), Convolutional neural networks (CNN), DenseNet, and ResNet. The system achieved high accuracy, precision, recall, and F1 scores (above 75%), and demonstrated real-time detection capability. By leveraging pre-trained models, the system performance has been further enhanced, enabling highly effective applications such as environmental monitoring, disaster response, and maritime surveillance.

Concerning the NN theme, Biçen and Celik (2024) explored the use of deep learning approaches to improve vessel inspections, focusing on the development of a post-inspection algorithm to identify the causes of flaws and guide future inspections. The research demonstrated that the MARCAT (the hierarchical map used to analyse the causes of observations) framework is effective in categorising inspection deficiencies, while the Natural Language Processing (NLP) based NN algorithm accurately predicts the causes of previous observations. This methodology enhances the precision and efficiency of causation analysis, helping to prevent accidents and optimise resource use in ship assessments. The study also highlights the practical application of artificial intelligence and data science in maritime safety, with potential implications for other industries reliant on inspections.

With a dual role, Ezzeddini et al. (2024) study focuses on a comparative analysis of Faster Region-based CNN and YOLOv8 models for real-time detection of fishing vessels and fish, aiming to enhance MS and fisheries management through deep learning techniques. The data source consists of a collection of images and videos recorded in various maritime environments. The research demonstrates that YOLOv8 outperforms Faster Region-based CNN in terms of precision and recall, achieving higher mAP50 scores, particularly 94% for fishing boats, highlighting its superior accuracy in marine object detection. These findings underscore the potential of YOLOv8 as a powerful tool for protecting marine resources, emphasising the need for further model improvements using diverse datasets, data augmentation, and hyperparameter optimisation.

2.1.3. Radar Detection

The third cluster focuses especially on improving radar systems for detection and classification. In this matter, Li et al. (2023) highlight the unique capabilities of deep learning models in infrared ship detection, emphasising their potential for all-day, all-weather surveillance applications. This study examines the impact of multi-band infrared images and deep learning algorithms on vessel recognition by comparing Cascade R-CNN, RetinaNet, and CenterNet. Results show that RetinaNet balances precision and speed, Cascade R-CNN excels with small vessels, CenterNet is faster, and short-wavelength infrared (SWIR) images outperform long-wavelength infrared (LWIR) in recognition. In the same scope, detection in complex backgrounds, Qian et al. (2023) propose to use the skeleton theory-based multiple model joint feature knowledge association (ST-MM-JFKA) method, which uses a hull-based dot model and a wake-based curve model to differentiate ships from sea clutter; it employs an improved curvature algorithm for candidate region extraction and a multi-scale model to handle varying object sizes, ultimately enhancing detection accuracy with specialised filtering techniques. To improve the situational awareness of field operators, Saager et al. (2021) present an approach that uses augmented reality to support during the classification phase.

2.1.4. Satellites, Maritime Vehicles, and Learning Algorithms

Clusters 4 and 5 can be jointly analysed, as they encompass studies that are thematically related to the previous clusters, though presented in a broader and less specific manner.

However, in this cluster, a new theme appears: maritime vehicles. This happens because, in addition to the monitoring systems mentioned, air or sea patrols in the area can also be employed by authorities to detect illegal activities, although their use is limited by available resources (Chen et al., 2022; Chircop et al., 2022). Although important, especially for military operations, this topic has received limited coverage in the academic literature compared to other subjects (Chen et al., 2022; Chircop et al., 2022). Assuming that Chen et al. (2022) conducted a study to maximise the efficiency of maritime patrolling through appropriate planning of a patrol route with or without the position of suspicious vessels provided by AIS or air support information. Chircop et al. (2022) presents new modelling and solution techniques for the Patrol Boat Scheduling Problem with Complete Coverage (PBSPCC) using a space-time resource network and an enhanced branch-and-price method. This study focuses on maximising the number of ships providing continuous presence, achieving 70% optimal solutions, with the most decisive factors being the length of the planning horizon and the number of patrol regions. For the use of unmanned surface vehicles (USVs), Xu et al. (2023) propose a target search algorithm (knowing the position) for complex maritime environments using a Markov prediction model to optimise search paths, enhancing efficiency and safety through improved ant colony optimisation algorithm, fuel management, and obstacle avoidance, with potential future applications in multi-target and heterogeneous searches. Similarly, Zhu et al. (2023) explore multi-agent task assignment for USV in maritime patrols, focusing on endurance constraints, which have often been overlooked in previous research. On the other hand, for air patrols using unmanned aerial vehicles (UAVs) equipped with Inverse Synthetic Aperture Radar (ISAR) and an electro-optical sensor, De Lima Filho et al. (2023) developed a methodology to search for illegal fishing vessels as quickly as possible. In this paper, the authors use AIS data and the ISAR to plan and update the route. This work reinforces the Dutch Navy perspective presented by De Wit et al. (2023) on the benefits of using UAVs in the field of MS when equipped with relatively capable sensors (e.g. SAR). In their study, Kirszenblat and Caelli (2022) demonstrate the benefits of using curvature-constrained gradient descent methods (GDM) to optimise mission path planning in imaging missions with ISAR, reducing path complexity by 80% and path length by more than 30% compared to the k-NN algorithm, with improved adaptability for complex vessel configurations and dynamic updates for real-time surveillance.

As we can see, the integration of sensors and other advanced technologies has been widely employed in MS, especially for remote detection. Additionally, the application of ML on diverse data sources, such as geospatial data, inspection records, and vessel trajectories, has proven to be essential for improving maritime monitoring activities and detecting illegalities in this field. Research into older articles and the problem of illegal fishing shows us other forms to apply learning algorithms and data sources to support decision-making. Wen et al. (2012) used data from on-board inspections of smuggled fishing vessels in the Taiwan region to apply Artificial Neural Networks (ANN) and LR models. They used vessel data to analyse and predict vessel smuggling. The ANN model achieved an average precision of 0.763, significantly higher than the 0.077 precision of human inspection, and reduced the workload by 90.47%.

Xu et al. (2017) developed a system for border security resource allocation and patrol routing plan that offers a structured framework for integrating historical records, Geographic Information System (GIS) and sensor data, and user input, enabling easy future implementation by different developers through model-driven development.

In their study, Pala et al. (2018) examined illegal fishing in the Gulf of Mexico, using a combination of data from sightings of small fishing boats and fishing gear, along with weather and moon phase information. The analysis, which employed LR and k-NN models, revealed three main peaks in fishing activity corresponding to the new, half, and full moons, suggesting that tides may influence fishing success. The study found that sightings were more frequent in the morning and showed seasonal variations, with wave height negatively impacting sighting frequency. However, other environmental factors did not show strong correlations, due to data limitations or the unpredictable nature of illegal fishing practices.

Davis and Harasti (2020) analysed illegal fishing in two no-take areas in New South Wales, Australia, using surveillance camera data. The study found that illegal fishing was higher on non-working days and during favourable weather conditions (no rain, light winds, and calm seas). A generalised linear model (GLM) effectively predicted illegal fishing at Seal Rocks but demonstrated lower accuracy at Little Broughton Island. The study suggests that weather and day type (working or non-working) influence illegal fishing patterns, offering valuable insights to target enforcement and outreach efforts to improve compliance.

In other areas of maritime knowledge, Navas de Maya et al. (2022) applied data mining techniques models (CT, Bayesian networks, decision rules and LR) to predict and rank non-conformities in oil tankers, using a 2006–2019 database to identify the most frequent issues. Models such as CT and LR successfully predicted and ranked non-conformities, aligning with key safety concerns in the tanker industry. Despite limitations in data sample size and variable complexity, the study's findings can aid decision-making and improve safety, with potential for broader application across different shipping sectors.

Table 2.1 – Studies on maritime surveillance technologies and global applications.

Author(s)	Research Focus	Methodology	Key Findings	Cluster
Wen et al. (2012)	Predicting smuggling vessels in Taiwan using data from board inspections	ANN and LR	ANN achieved 0.763 precision, reducing workload by 90.47%, significantly outperforming human inspection.	5

Author(s)	Research Focus	Methodology	Key Findings	Cluster
Pala et al. (2018)	Illegal fishing patterns in the Gulf of Mexico using sightings data	LR and k-NN	Illegal fishing peaks during specific moon phases; wave height negatively impacts sightings.	5
Davis and Harasti (2020)	Patterns of illegal fishing in no-take zones in Australia using surveillance camera images	GLM applied to surveillance camera data	Illegal fishing is higher on non-working days and during favourable weather conditions.	5
Ghosh et al. (2021)	Detecting vessel-like objects using medium-resolution satellite imagery	Three-stage methodology	Precision result of 64% and a recall of 86% for most of the scenes.	2
Wang et al. (2021)	Ship detection and tracking using SAR imagery	Novel Ship Detection and Tracking (NSDT) model	SVM achieved 95% accuracy.	2
Chen et al. (2022)	Optimising maritime patrol efficiency	Patrol route optimisation with or without position given by AIS or air patrol	The efficiency of maritime patrol increases when the positions are known.	4
Chircop et al. (2022)	New modelling and solution techniques for PBSPCC	Resource-space-time network	Achieve 70% optimal solutions.	5
Navas de Maya et al.(2022)	Predict and rank non-conformities	CT, Bayesian networks, decision rules and LR	Data mining techniques reliably identified key tanker non-conformities, mainly linked to operations, structure, and human error.	5
Bhattacharjee et al. (2023)	A network pipeline to reduce computational complexity	A single-stage detector with a two-network architecture (feature extraction and regression)	A state-of-the-art architecture and a deep learning-based algorithm for vessel localisation with SAR images.	2
Guida et al. (2023)	Investigate IUU fishing activities in the Mauritius EEZ	AIS/SAR data matching	Highlight the need for improved detection methods and satellite solutions for more effective fisheries monitoring.	2
K et al. (2023)	Vessel classification using satellite images	Advanced deep learning models like YOLOv3, CNN, DenseNet and ResNet	The used models show high accuracy, precision, recall, and F1 scores when recognising vessels.	4
Li et al. (2023)	Select the appropriate infrared detection band (radar)	Analysis of three deep learning models: Cascade R-CNN, RetinaNet, and CenterNet	RetinaNet balances precision and speed, Cascade R-CNN excels with small vessels, and CenterNet is faster.	3

Author(s)	Research Focus	Methodology	Key Findings	Cluster
De Lima Filho et al. (2023)	Develop a maritime patrol planning methodology to find illegal vessels	Use the TSP algorithm and the information given by the AIS and the ISAR vessel system	The used algorithm proves to be effective.	4
Do Nascimento et al. (2023)	Detecting fishing activities from vessel trajectories	Comparison of ML models (LSTM and RF)	LSTM achieved 100% accuracy but was slower; RF was 1,000 times faster with 87% accuracy.	1
Pohontu et al. (2023)	Classifying vessels based on their trajectories	Handwriting Signature Verification	Identify 12 main classes.	1
Qian et al. (2023)	Differentiate vessels from sea clutter	ST-MM-JFKA method	Effectively removes false wakes and successfully detects vessels.	3
Xu et al. (2023)	Optimise the search trajectory	Markov prediction model	Improve the success rate and efficiency of finding the target.	4
Yan et al. (2023)	SAR vessel classification integrating AIS data	SMOTEBoost algorithm with AIS and SAR data	Achieved 93% accuracy, excelling in cargo, tanker, and fishing ship classification.	2
Zhu et al. (2023)	Multi-agent task assignment for maritime patrols with endurance constraints	Optimisation algorithms considering endurance and multi-target searches	Enhanced USV patrols with efficient task assignment for complex maritime scenarios.	5
Agrawal and Ramteke (2024)	Optimising maritime surveillance as a dynamic TSP	AIS data for optimising monitoring routes	Efficient monitoring of 172 vessels through optimised routing.	1
Bernabé et al. (2024)	Real-time detection of suspicious AIS trajectories	Self-supervised deep learning for AIS shutdown anomaly detection	Achieved 99.5% accuracy in predicting AIS message reception, detecting suspicious vessel activities.	1
Biçen and Celik (2024)	Post-inspection algorithm to identify the causes of flaws and guide future inspections	NLP based on NN	MARCAT effectively categorised the inspection deficiencies while NLP-based NN accurately predicts the causes.	2
Ezzeddini et al. (2024)	Real-time detection of fishing vessels and fish for maritime surveillance using images and video footage	Comparative analysis of Faster R-CNN and YOLOv8 models	YOLOv8 outperformed Faster R-CNN with higher precision and recall, achieving 94% mAP50 for fishing boats.	2
Liu et al. (2024)	Systematic framework for detecting maritime kinematic anomalies	Model training with historical AIS data and real-time data analysis	Achieved an F1-score of 99.79% and reduced misclassification by 62%.	1

The literature review highlights that, although MS has significantly progressed with technologies such as AIS, SAR, radar, satellite systems, and ML techniques for anomaly detection and vessel classification, there is still a gap in decision support for selecting patrol areas based on previous inspection results such as those used by Wen et al. (2012). Most studies focus on enhancing real-time detection accuracy, automated classification and surveillance optimisation, but seldom explore leveraging historical inspection data to optimise patrol planning. An overview of the most important articles can be found in Table 2.1. Some approaches, such as those of Chen et al. (2022) and Chircop et al. (2022), emphasise the efficiency of the patrol route without integrating insights from previous inspections. Therefore, there is an opportunity for research to develop decision support systems that apply predictive analytics and historical data to identify high-risk areas.

2.2. MARITIME SURVEILLANCE BY THE PORTUGUESE NAVY

Maritime surveillance and inspection of ships in Portugal's coastal and oceanic waters are largely carried out by the Navy ships (Mourinha, 2012). To control, monitor and, support the planning they use the MONICAP ("*Sistema de monitorização contínua da atividade da pesca*") also designated Vessel Monitoring System (VMS) , SIFICAP ("*Sistema de Fiscalização e Controlo das Atividades da Pesca*"), SADAP ("*Sistema de Apoio à Decisão na Atividade de Patrulha*") and SIPM ("*Sistema de Informação da Polícia Marítima"*) (Beja, 2023; da Rocha Rei, 2016; Mourinha, 2012; Pinto, 2017; Rodrigues, 2021; Vala, 2023). In addition to VMS data (similar to AIS), one of the main data sources for the this systems is the Fiscalization Report (FISCREP), generated after each inspection of a vessel—whether the vessel is commercial, engaged in fishing activities, recreational, national, or foreign—that operates in waters under national sovereignty or jurisdiction (Moura et al., 2024). These reports contain information regarding the position and date of the inspection, the identification and type of ship and the type of presumed illegalities that were found.

The data resulting from FISCREP has been utilised and analysed in many ways in several academic studies. Da Rocha Rei (2016) conducted research to assess the effectiveness of enforcement actions in Portuguese maritime fishing areas, focusing on the impact of supervision on the protection of resources and compliance with regulations. Analysing FISCREP inspection data from 2006 to 2015, the study found a compliance rate of 80.7%, which was higher among vessels with VMS and those registered in Portugal. The presence of VMS was associated with fewer infractions, indicating its deterrent effect. Additionally, the research developed discriminant functions to predict noncompliance using econometric and stochastic models, achieving 79% overall accuracy and 96% accuracy for legal vessels, providing valuable tools for targeted enforcement and sustainable fisheries management.

On the other hand, Pinto (2017) analysed FISCREP data between January 2014 and December 2015, cross-referencing them with fishing statistics from the Portuguese National Statistical Institute and Global Fishing Watch. The research focused on the evolution of fisheries regulation in the Portuguese maritime ecosystem, particularly on monitored fishing methods and associated infractions. The main findings indicated that polyvalent fishing gears were the most frequently inspected, while purse-seine methods, despite fewer inspections, accounted for a considerable share of the total fish catches. The study also highlighted that illegal fishing activities were concentrated near the southeastern boundary of the EEZ, showing a decreasing trend over time, and underlined the need for enhanced monitoring, targeted enforcement, and collaborative efforts to promote sustainable fisheries management and regulatory compliance.

AIS data were used by Rodrigues (2021), who combined them with FISCREP data from 2015 to 2020, and with vessel registration databases to optimise the monitoring of fishing activity, focusing on vessels with MONICAP (around 10% of the fleet). He developed a density map and a detailed list of vessels to better allocate Navy resources, increasing the effectiveness of inspections. By using parallel processing, he reduced analysis time by 98.5%. Three new performance indicators were created, allowing for more efficient allocation of patrols in critical areas. The findings indicate that this Big Data approach significantly improves the focus and efficiency of maritime inspection operations.

Exploring the potential of Business Intelligence (BI), Beja (2023) proposed a Power BI solution to enhance the reporting and analysis capabilities of the Portuguese Navy. This approach addresses the significant challenges posed by IUU fishing, which not only affects the economy and environment but also undermines effective fisheries management and fair competition for compliant fishers. By centralising data from SIFICAP and other inspection sources, such as FISCREP, the system transforms fragmented information into actionable insights. The goal of this BI solution is to improve decision-making, streamline sea inspections, and assess the effectiveness of maritime surveillance on fisheries-related violations, strengthening the enforcement of fisheries regulations.

In a comprehensive geospatial analysis, Vala (2023) examined fishing infractions along the Portuguese coast from 2015 to 2022 using data from the FISCREP database. The study also developed metrics to assess the intensity and diversity of fishing activity, leveraging the MONICAP database with a case study focused on the Algarve region. A comparative analysis between the FISCREP and MONICAP datasets for the overlapping period (2021-2022) revealed data discrepancies due to legal changes in vessel registration formats, which were resolved through cross-referencing with external sources. Findings revealed that most of the infractions were related to vessel and crew safety rather than sustainability concerns, highlighting the safety negligence among industry operators. The study concluded that, despite the decrease in the number of inspections by the Portuguese Navy, certain regions and fishing practices still require reinforced monitoring, with regional variations underscoring the need for area-specific regulatory approaches.

Taking advantage of geostatistical techniques, Miguel et al. (2024) focus on mapping the spatial distribution of presumed fisheries infractions along the southern coast of Portugal, particularly in the Algarve. Using Bayesian hierarchical models and the INLA-SPDE methodology, the research predicts infraction probabilities based on georeferenced binary data. The resulting risk maps identify areas with high risks of infringement, mainly Sagres, Faro, and Tavira, where the probability of infringement can rise to 80% during specific times of the day. These insights aim to increase the Portuguese Navy's inspection efficiency, guiding patrol routes, thus reducing operational costs and strengthening MS efforts. This work stands out as a new application of geostatistical modelling in maritime enforcement contexts.

To support future studies and enable the implementation of data mining algorithms on FISCREP data collected from fishing vessels, Moura et al. (2024) ensure the anonymisation of data, maintaining its accuracy, and cross-referencing it with vessel information from other sources.

Table 2.2 - Studies on maritime surveillance and inspections in Portugal.

Author(s)	Research Focus	Methodology	Key Findings
Da Rocha Rei (2016)	Effectiveness of enforcement actions in Portuguese maritime fishing areas	Analysis of FISCREP data (2006-2015) using econometric and stochastic models	Found 80.7% compliance rate. MONICAP's presence is linked to fewer infractions. Developed discriminant functions predicting noncompliance with 79% overall accuracy.
Pinto (2017)	Evolution of fishery regulations and associated infractions in Portugal	Analysis of FISCREP reports (2014-2015) cross-referenced with fishery statistics	Highlighted infractions in specific fishing methods; illegal activities concentrated near the EEZ southeastern boundary.
Rodrigues (2021)	Optimising fishing activity monitoring in Portugal	Use of AIS data and inspection reports (2015-2020) with density mapping and parallel processing	Created performance indicators and improved patrol efficiency, reducing analysis time by 98.5%.
Beja (2023)	Enhancing reporting and analysis in Portuguese Navy operations using BI	Proposed Power BI system centralising data from SIFICAP and FISCREP	Improved decision-making and inspection efficiency, addressing challenges of IUU fishing.
Vala (2023)	Geospatial analysis of fishing infractions along the Portuguese coast	Comparative analysis of MONICAP and FISCREP data (2015-2022)	Uncovered regional infraction trends; most infractions linked to safety concerns. Emphasised the need for area-specific regulations.
Miguel et al. (2024)	Spatial distribution of presumed fisheries infractions along Portugal's southern coast	Bayesian hierarchical models and INLA-SPDE methodology	Identified high-risk zones (Sagres, Faro, Tavira) for infractions, guiding more efficient patrol routes.
Moura et al. (2024)	Use of FISCREP data of vessel inspections in Portugal's jurisdictional waters	Compilation of FISCREP reports for compliance and resource planning	Highlighted gaps in leveraging inspection data for strategic patrol planning.

As we can see, MS and ship inspections in Portugal are mainly conducted by the Navy, using systems such as MONICAP, SIFICAP, SADAP, and SIPM to support monitoring and decision-making. A significant source of information comes from FISCREP reports, which are generated after each vessel inspection and have been used in several studies (Table 2.2). However, despite the availability of this rich dataset, existing studies mainly focus on compliance rates and optimising monitoring efforts, without fully leveraging inspection history for strategic patrol planning. For example, although research by da Rocha Rei (2016) and Rodrigues (2021) has shown improvements in resource allocation using inspection data, there is still an opportunity to use this information more effectively to enhance patrol area selection. On the other hand, the work of Miguel et al. (2024) focuses on a specific area of Portugal. Although the types of infractions and vessel typologies extend beyond fishing activities, the studies focus exclusively on these types of vessels and infringements.

By combining international and national reviews, it is possible to identify a gap that can be addressed. Based on the work of Wen et al. (2012), inspection records can be used to predict future illegal activities. This approach would go beyond the spatial distribution presented by Vala (2023) and Miguel et al. (2024), not only expanding the geographic scope beyond the Algarve but also broadening the range of vessel typologies. Furthermore, incorporating the variables available in FISCREP (vessel type, location and date) would result in a more dynamic system, enabling the selection of areas based on historical data rather than merely the presence or absence of vessels. This approach would improve accuracy by taking into account the correlated probability of illegal activity.

Based on this concept, a data-driven approach such as the Illegal Parking Score (IPS) framework proposed by Jardim et al. (2022) could be adapted to optimise maritime patrol operations. In their study, the IPS is calculated based on the conditional probabilities of several types of parking infractions. Following a similar methodology, our goal is to use data from FISCREP reports to estimate the likelihood that any inspected vessel, regardless of type, may be in presumed violation in different zones of the Portuguese EEZ. By leveraging these probabilities, we aim to support the development of a dynamic decision-support tool that improves MS by guiding ship patrolling efforts more effectively, increasing the efficiency of maritime inspections.

3. METHODOLOGY

3.1. RESEARCH DESIGN

This work, aligned with the objective, falls within the field of data mining. For this type of initiative, Schröer et al. (2021) identify the CRISP-DM methodology as the standard process. This process consists of six iterative phases (Figure 3.1), namely: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment (Chapman, 2000).

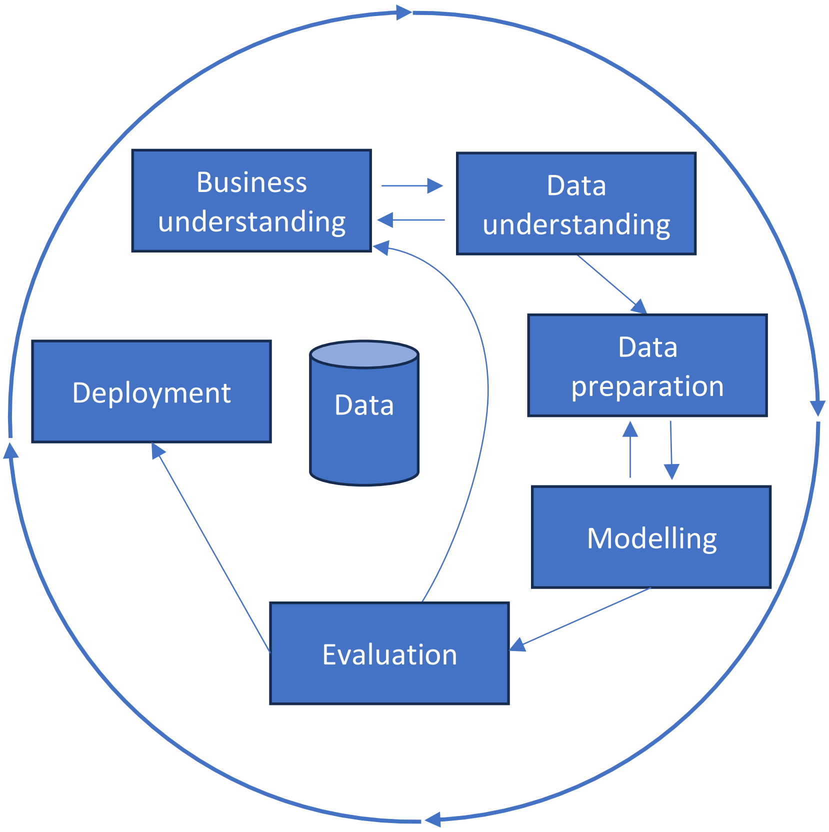


Figure 3.1 – Phases of the CRISP-DM model. Adaptation of the original image by Chapman (2000, p. 13).

In Figure 3.1, the arrows illustrate the most significant and frequent interactions or relationships between various phases of the data mining process. The outer circle emphasises the iterative nature of this process, suggesting that the cycle does not end with the deployment of a solution. Instead, it highlights that the process is continuous, involving revisiting previous phases for refinement, improvement, or adaptation based on new data or changing requirements triggered by the implemented solution (Chapman, 2000).

In the following chapters, we will explain how we applied each phase to this project.

3.2. BUSINESS UNDERSTANDING

The first phase of this type of work is to understand the objectives and requirements from a business perspective. In this context, "business" refers to surveillance and inspection activities in areas under national sovereignty and jurisdiction. More specifically, the Portuguese Navy contributes to the Maritime Authority System, as defined in Decree-Law No. 43/2002.

In simplified terms, the MS and ship inspection process (Figure 3.2) begins when the Navy commanding officer receives their instructions. These instructions may be highly specific, such as directing operations to a particular area or directing vessels using certain types of fishing gear, or more generally, instructing the commander to exercise the state's sovereignty at sea. With these instructions in hand, the commanding officer integrates them with other mission directives to plan the ship's movements (Rodrigues, 2021).

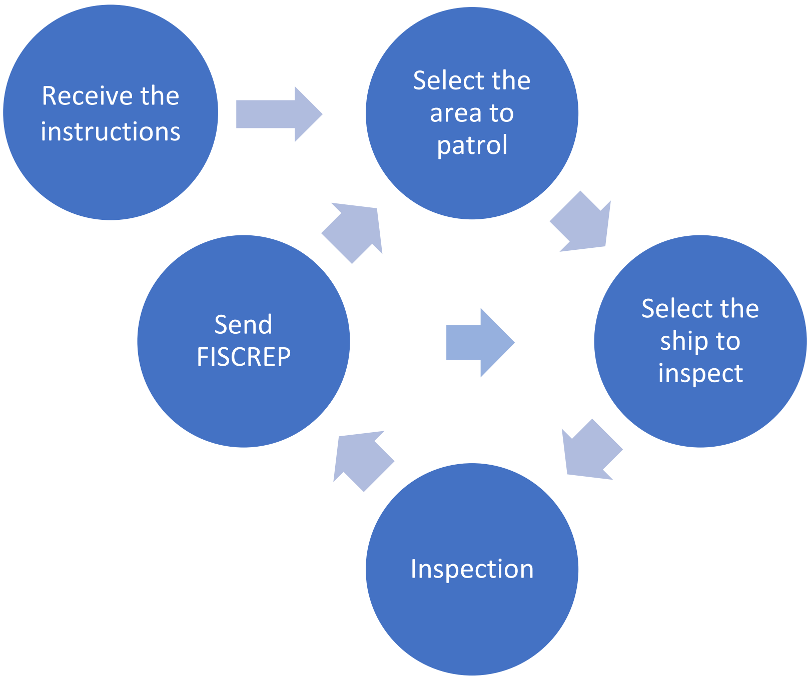
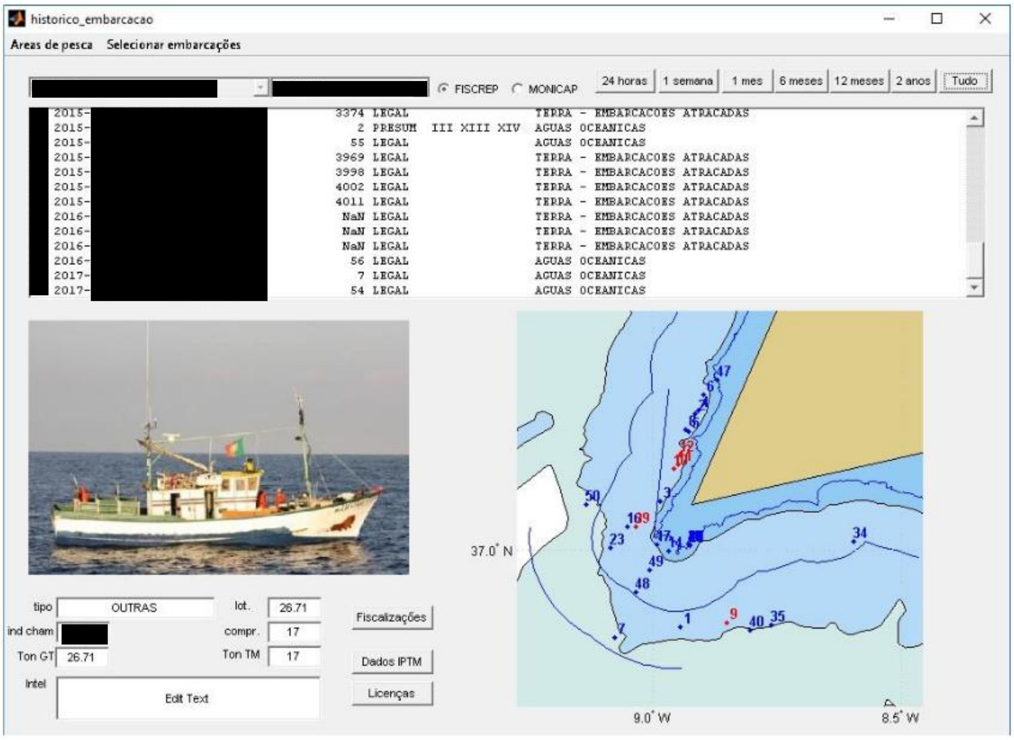


Figure 3.2 – Maritime inspection process.

In addition to selecting patrol areas to maximise operational effectiveness, during transit, the ship is assumed to be fulfilling its mission. When detecting a vessel subject to inspection, the commanding officer assesses whether an inspection will be conducted (da Rocha Rei, 2016). If an inspection takes place, the results are sent to the operational command through a FISCREP (Moura et al., 2024). Subsequently, the ship may be relocated to a different area or carry out additional inspections within the same region.

The decision regarding which area to patrol or which vessel to inspect is made by the commanding officer's empirical knowledge and the evaluation of historical inspection data, using SADAP (da Rocha Rei, 2016; Mourinha, 2012). This system, which uses the FISCREP and MONICAP data, enhances decision-making by providing essential functionalities, such as visualising positions and outcomes of previous inspections (Figure 3.3), offering statistical insights into previous operations (Figure 3.4), and generating density maps that reveal activity patterns (Figure 3.5) (Rodrigues, 2021).



[redacted] [redacted] ◉ FISCREP ○ MONICAP

```
[redacted] 2015-[redacted]      3374 LEGAL                TERRA - EMBARCACOES ATRACADAS
           2015-                    2 PRESUM  III XIII XIV  AGUAS OCEANICAS
           2015-                   55 LEGAL                AGUAS OCEANICAS
           2015-                 3969 LEGAL                TERRA - EMBARCACOES ATRACADAS
           2015-                 3998 LEGAL                TERRA - EMBARCACOES ATRACADAS
           2015-                 4002 LEGAL                TERRA - EMBARCACOES ATRACADAS
           2015-                 4011 LEGAL                TERRA - EMBARCACOES ATRACADAS
           2016-                  NaN LEGAL                TERRA - EMBARCACOES ATRACADAS
           2016-                  NaN LEGAL                TERRA - EMBARCACOES ATRACADAS
           2016-                  NaN LEGAL                TERRA - EMBARCACOES ATRACADAS
           2016-                   56 LEGAL                AGUAS OCEANICAS
           2017-                    7 LEGAL                AGUAS OCEANICAS
           2017-                   54 LEGAL                AGUAS OCEANICAS
```

tipo OUTRAS
ind cham [redacted]
Ton GT 26.71
lot. 26.71
compr. 17
Ton TM 17
Intel Edit Text

Figure 3.3 – SADAP view regarding the history of inspections conducted on a specific vessel. Retrieved from Rodrigues (2021, p. 56).

Although this information is crucial, it remains static and is presented through isolated visualisations. The challenge, therefore, lies in analysing this data comprehensively. By applying data mining techniques, the objective is to extract actionable insights that enhance decision-making, ultimately optimising surveillance and inspection efforts.

[redacted]

Figure 3.4 – SADAP statistical information regarding inspections conducted in 2017. Retrieved from Rodrigues (2021, p. 54).

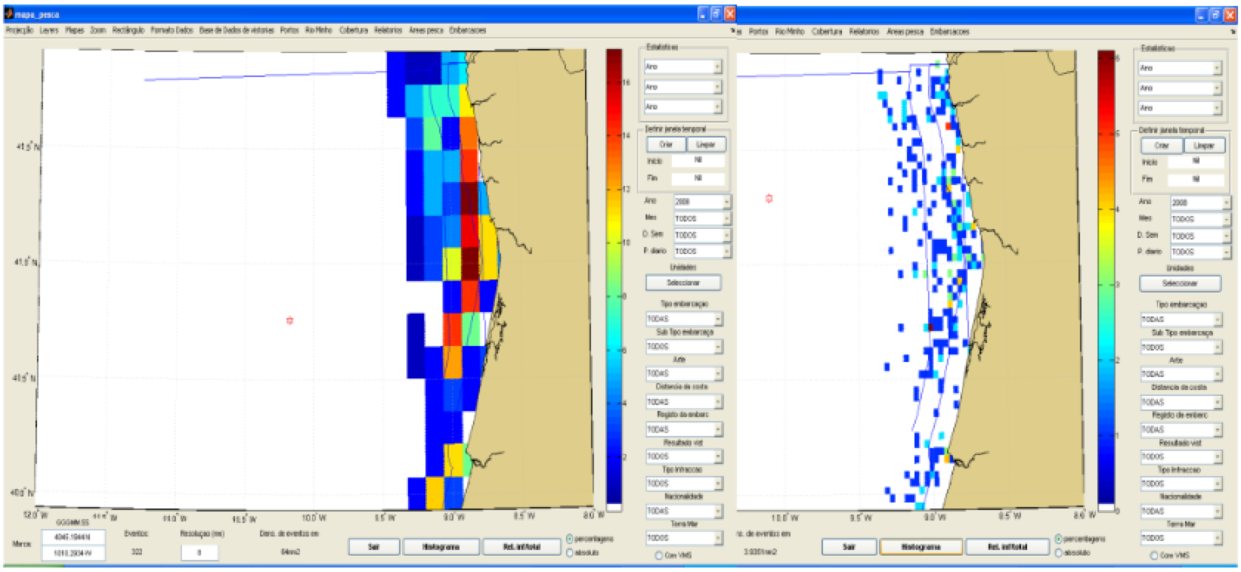


Figure 3.5 – SADAP density maps. Retrieved from Rodrigues (2021, p. 55).

In this context, our objective is to utilise FISCREP data to classify areas where encounters with suspected illegal vessels are most likely. We hypothesise that there may be a correlation between the timing of an inspection, such as the month, day of the week, or time of day, and the likelihood of detecting illegal activity (da Rocha Rei, 2016; Miguel et al., 2024). Based on these assumptions, calculations must incorporate temporal factors, including the time of day, day of the week and month in which the mission occurs. Furthermore, given the nature of the inspections, the target vessel type should also be considered to provide decision makers with a more tailored and effective system (Figure 3.6).

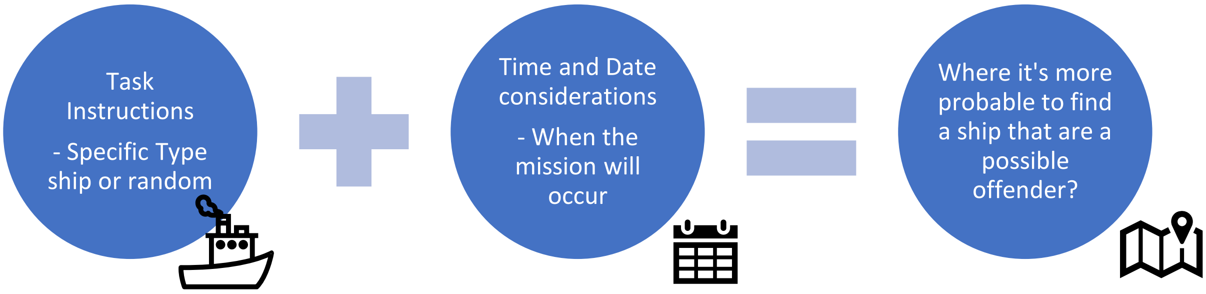


Figure 3.6 – Path to determining the areas with the highest probability of encountering suspected illegal vessels based on mission tasking and time considerations.

3.3. DATA UNDERSTANDING

To achieve our goal, it’s important to understand the data. Our main data source will be the one derived from FISCREP. This report is a formatted message; in other words, each line of sequence is identified by a subject and is filled with specific information, always in the same way and order. For example, the line regarding the identification of the inspected ship includes fields for type, subtype, and name. In most cases, each field has a predefined set of options. As a result, some fields (Table 3.1) are loaded into the SADAP database (Moura et al., 2024).

Table 3.1 – FISCREP data load on the SADAP. Adapted from Moura et al. (2024, p. 3).

Registered Information	Description
Nº	Inspection identification number assigned by the Portuguese Navy
Name	The inspected vessel name
CFR	Community Fleet Register
Nº of reg	Registration number of the vessel
Latitude	The latitude where the inspection took place
Longitude	The longitude where the inspection took place
GDH	Group-Date-Hour (GDH) or Date-Time-Group (DTG)
Unit	The Portuguese Navy ship that conducted the inspection
Vessel Type	The type of activity that the vessel is intended to perform
Sub type	The type of fishing gear allowed to be used
“Arte”	The type of fishing gear used at the time of inspection
Result	The outcome of the inspection: Legal or Presumed Offender
Infraction	Infraction codes when “Presumed Offender”

For this work, data were requested from the Portuguese Navy. Table 3.2 provides an overview of the dataset (FISCREP.xlsx). Each row represents a single inspection of a vessel at a specific location, time, and date. The inspections were conducted between January 1, 2014, and December 31, 2024.

Table 3.2 – Description of FISCREP.xlsx dataset with type and quantity of null values.

Designation	Description	Type	Quantity of nulls
Latitude	The latitude where the inspection took place	Float	5
Longitude	The longitude where the inspection took place	Float	10
GDH	Date and time when the inspection took place	Datetime	
Year	The year when the inspection took place	Integer	
Month	The month when the inspection took place	Integer	
Day	The day when the inspection took place	Integer	
Time	The time when the inspection took place	Integer	
Daily period 6 hours	One of 4th six hours day period	Integer	
Unit	The ship that conducts the inspections	String	
Vessel Type	The type of activity that the vessel is intended to perform	String	61

Designation	Description	Type	Quantity of nulls
Vessel Subtype	The specific subtype	String	510
Fishing gear	The type of fishing gear used at the time of inspection (only for fishing vessels)	String	1840
Result	The outcome of the inspection: Legal or Presumed offender	String	
Registration Nº	Registration number of the vessel	String	32
Ship Name	The vessel name	String	33
Infraction (I)…(XIV)	The quantity of each infraction, divided into 14 types	Integer	
CFR	Community Fleet Register	String	11586
FISCREP Nº	Inspection identification number assigned by the Portuguese Navy	Integer	

The dataset contains a total of 17,361 entries, after removing 14 duplicate records, and includes 31 features. Some features have missing values, as detailed in Table 3.2; how these missing values will be handled is explained in the data preparation chapter. To facilitate analysis, the features will be categorised based on the type of information they provide.

3.3.1. Position of the inspection

The position of each inspection is given by the "Latitude" and "Longitude" values. These coordinates are represented in degrees with decimal values, where negative values indicate south and west. In the dataset, the values for "Latitude" range from *-0.405667* to *37009.83*, and for "Longitude", from *-81403.20* to *757.30*.

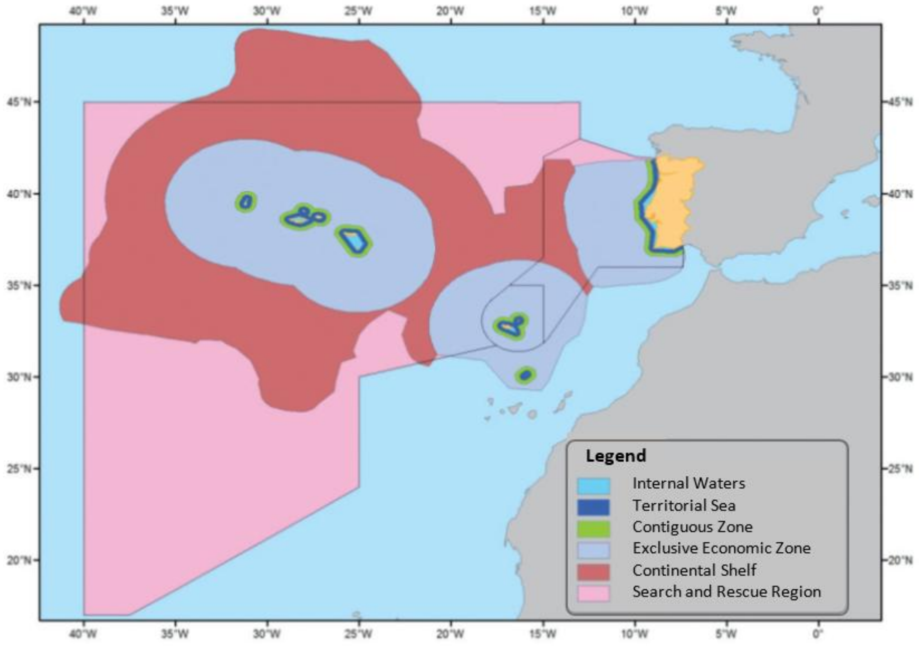


Figure 3.7 – Maritime Zones under Portuguese Sovereignty and/or Jurisdiction. Retrieved from *Direção-Geral de Recursos Naturais, Segurança e Serviços Marítimos* (2025).

Given these values, we can assume the presence of incorrect data, since the maximum latitude is 90º and the longitude is 180º, besides that the latitude in Portuguese waters varies from 25° North to 45° North, and the longitude varies from 40° West to 5° West (Figure 3.7).

3.3.2. Date and time of the inspection

The information regarding time, day and hour is represented by the features “GDH”, “Year”, “Month”, “Day”, “Time” and “Daily period 6 hours”. The main feature is “GDH”, a timestamp that contains the day, month, year, hour and minutes the inspection takes place. In this dataset, this feature is split into separate features representing the date and time components. Additionally, the “Daily period 6 hours” feature divides the day into four six-hour periods:

- **1** - represents 00:01 to 06:00,
- **2** - represents 06:01 to 12:00,
- **3** - represents 12:01 to 18:00,
- **4** - represents 18:01 to 00:00.

The “Year” feature ranges from 2014 to 2024. The year with the highest number of inspections was 2015, while the year with the lowest number of inspections was 2021 (Figure 3.8).

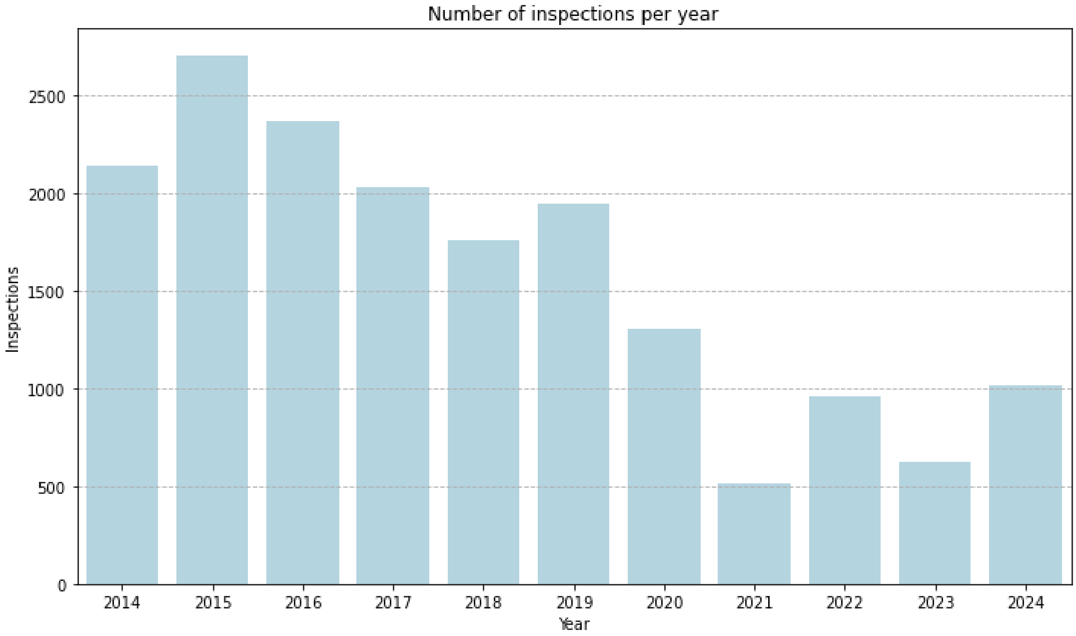


Figure 3.8 – Number of inspections per year.

The “Month” feature includes inspections from all months. Overall, the month with the most inspections was January, while the month with the fewest inspections was October (Figure 3.9).

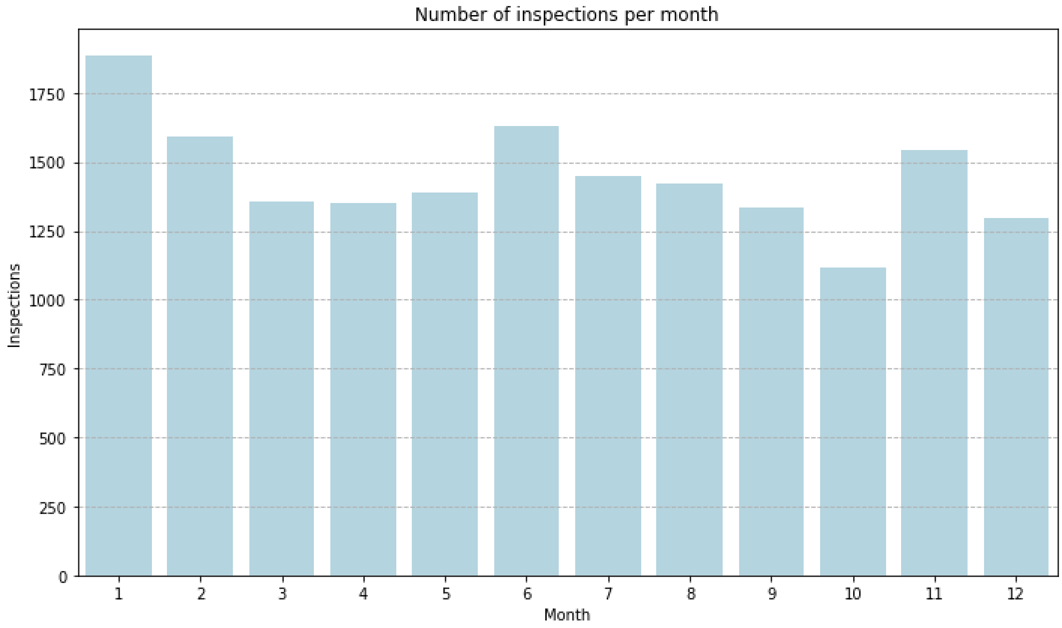


Figure 3.9 – Number of inspections per month.

Regarding the time of day, most inspections occurred between 06:01 and 12:00, while the fewest inspections took place between 00:01 and 06:00 (Figure 3.10).

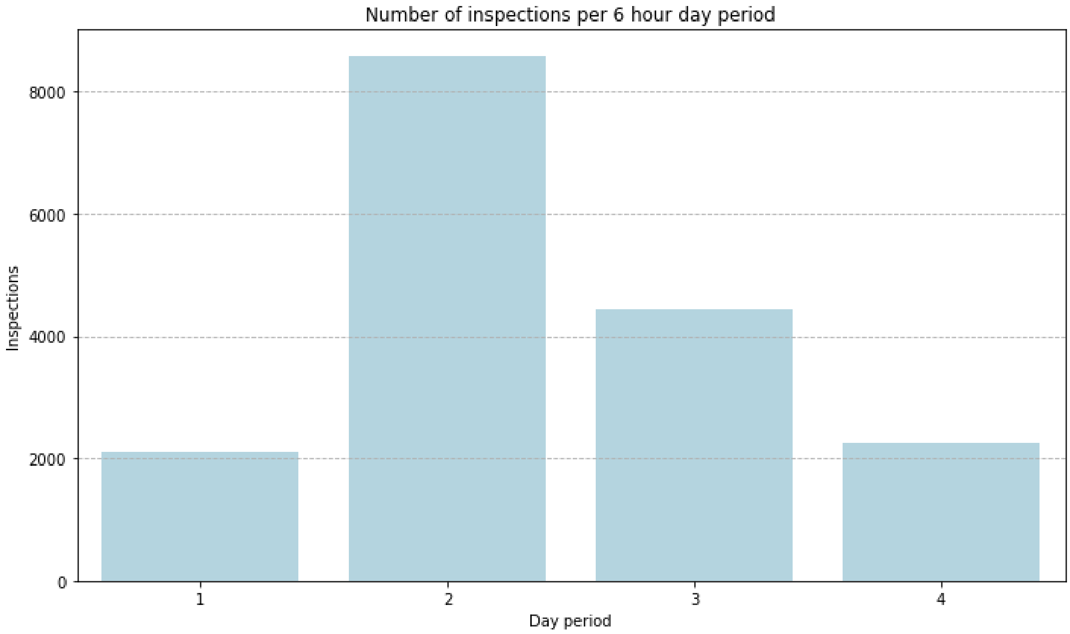


Figure 3.10 – Number of inspections per 6-hour day period.

3.3.3. Inspected vessel

To identify the inspected vessel, the name, registration number, and CFR are collected. In this case, the absence of one of these data values may indicate an error in reading the report.

In addition to information about each inspected vessel, they can be categorised by type, a corresponding subtype, and, for fishing vessels, the fishing gear used.

The "Vessel Type" feature includes all five possible categories:

- *Pesca Comercial* (Commercial Fishing),
- *Recreio* (Recreational),
- *Artes Caladas* (Set Nets),
- *Marítimo Turísticas* (Maritime Tourism),
- *Outras* (Others).

Commercial fishing vessels are the most frequently inspected (Figure 3.11). Regarding the 45 subtypes, fishing vessels classified as *Emalhar/Tresmalho* (Gillnet/Trammel Net) are the most inspected, and the fishing gear *Rede de Tresmalho* (Trammel Net) is the most inspected of the 38 registered.

For this last feature, null values are expected, as it applies only to fishing vessels.

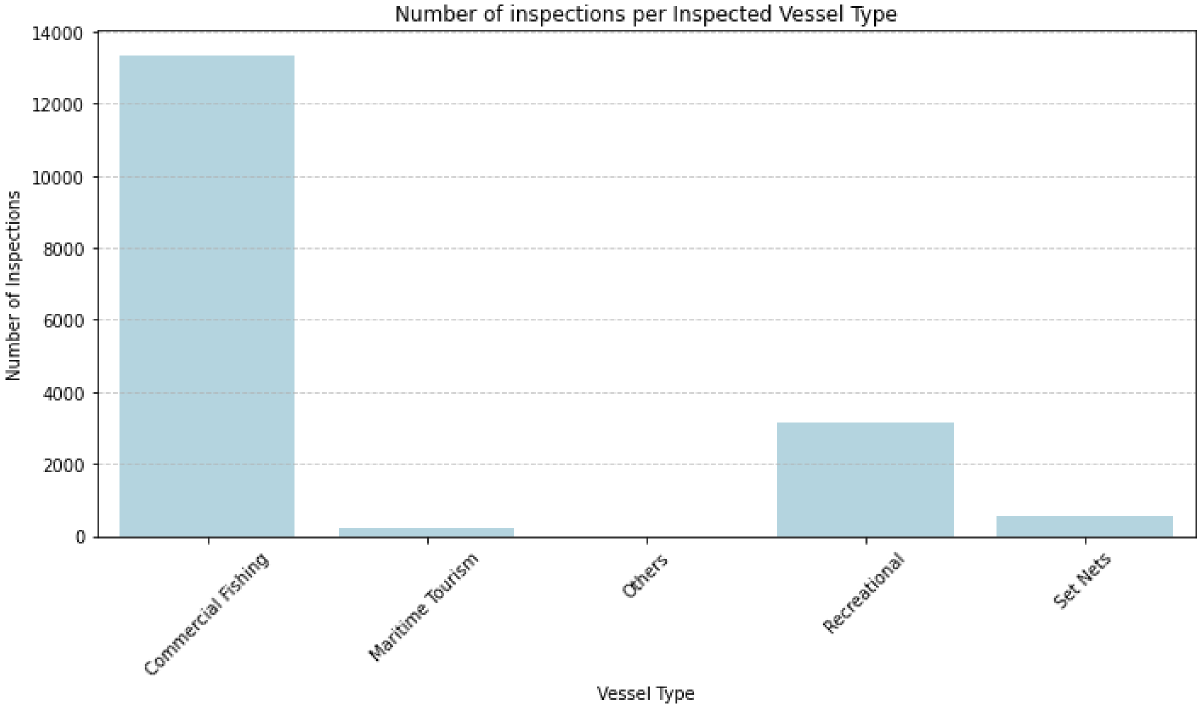


Figure 3.11 – Number of inspections per vessel type.

3.3.4. Unit Conducting the inspection

The identification of the Portuguese Navy ship carrying out the inspection is provided by the feature "Unit". Due to data classification, specific values will not be displayed in this work. However, for data understanding purposes, it is noted that the unit with the highest number of inspections conducted a total of 3,123 inspections. Each report is assigned a sequential number (“FISCREP Nº”).

3.3.5. Result

After an inspection, the result is classified as *Legal* or *Presumed Offender*. This dataset shows that 83% of the inspections were considered *Legal*. However, 1% of the inspections have their result registered as *Unknown*. Given the small number of occurrences (19), entries in the “Result” field categorised as *Unknown* were removed.

When an infraction is detected, the quantity and type are also recorded. Infractions are divided into 14 categories as follows (Moura et al., 2024, p. 11):

- **I** - Non-existent logbook,
- **II** - Incorrectly filled logbook,
- **III** - Prohibited fishing gear,
- **IV** - Fishing in a prohibited or restricted area,
- **V** - Fishing prohibited due to excessive engine power or tonnage,
- **VI** - Improper catches due to prohibited fishing,
- **VII** - Improper catches due to bycatch,
- **VIII** - Improper catches of undersized fish,
- **IX** - Activity conducted without a license or authorisation,
- **X** - Improper marking or identification of fishing gear,
- **XI** - Improper marking or identification of the vessel,
- **XII** - Miscellaneous: Invalid certificates,
- **XIII** - Miscellaneous: Non-existent/invalid maritime registration,
- **XIV** - Miscellaneous: E.g., lack of onboard documents, lack of pyrotechnics, expired life-saving equipment, expired fire extinguishers, among others.

In the dataset, the most frequent infractions found were related to non-existent/invalid maritime registrations.

3.4. DATA PREPARATION

Although the dataset contains a large amount of data, it is important to stay focused on the objective of this project. With that in mind, we select data aligned with the objectives and correct the issues identified in the previous phase to construct the dataset for the modelling phase.

3.4.1. Data sets preparation

For the first step of data preparation, we will follow the approach of Pala et al. (2018). Since the dataset is imbalanced (i.e., over 80% of the cases are *Legal*), a stratified train-test split was applied. The dataset is divided into a training set (80%) and a testing set (20%), ensuring that both subsets maintain the same proportion of *Presumed Offender* to *Legal* vessels as in the original dataset. The testing set is kept separate and is only used for the final evaluation of the model's performance after training and hyperparameter tuning.

To enable hyperparameter tuning, we will perform a second stratified split within the training set, creating training and validation subsets with the same proportion. Given the imbalanced distribution, the minority class will be over-sampled using SMOTE (Lemaître et al., 2017) on the training set after the resolution of missing values.

3.4.2. Clean data

During the data understanding phase, several inconsistencies and errors in the "Latitude" and "Longitude" values were identified in the dataset, which were subsequently addressed to ensure data integrity for further analysis. Concerning latitude values, it was found that 14 values were outside the expected range of [-90, 90], with values ranging from *-0.405667* to *37009.83*. These discrepancies were assumed to be the result of a decoding error during data entry or a typing error. To resolve this problem, it was decided to shift the decimal point one place to the right of the second digit, thereby adjusting the values within the correct geographic range (e.g. *37009.83* changes to *37.00983*).

Regarding the longitude values, 12 records were identified as outside the expected range of [-180, 0], with values spanning from *-81403.20* to *757.30*. Similarly, this has been attributed to a decoding error or typing error. To correct the values, the decimal point was moved one place to the right of the first digit, and for positive longitude values, a negative sign was applied, ensuring the data reflected the correct geographic hemisphere (e.g. *757.30* changes to-*7.5730*).

Furthermore, there were missing values in both features, five in "Latitude" and ten in "Longitude". Given that the same unit carried out the inspections in the same month and year, it was assumed that they occurred in the same geographic region. Consequently, the missing values for "Latitude" and "Longitude" were imputed with the average of the corresponding values from inspections conducted by the same unit in that month and year, ensuring consistency within the dataset. The average value was calculated from the training data and applied to the validation and test datasets.

In addition to the inconsistencies identified in the "Latitude" and "Longitude" values, 61 entries with null values were found in the "Vessel Type" field. Among these, 47 records contained information in the "Fishing Gear" field but were missing data in the "Vessel Subtype" field. To solve this issue, the missing "Vessel Type" values were imputed based on the corresponding “Fishing Gear” information. The remaining 15 entries, which did not contain “Fishing Gear” and “Vessel Subtype” data, were categorised as *Others*.

3.4.3. Select data

After the data cleaning phase, it was necessary to select the relevant data for analysis. This process was carried out in two steps: first, by selecting the inspection records relevant to the study, and second, by identifying the features to be used in the analysis.

Given the objectives of this work, the position values should be aggregated by areas corresponding to a grid system. Consequently, the first step consisted of creating a grid with a resolution of 10 by 10 nautical miles, covering the latitude range [42°N – 34.5°N] and longitude range [14°W – 7°W], based on the Portuguese EEZ (Figure 3.7). Each grid cell was given an alphanumeric code, where the letter represents the northernmost latitude, and the number denotes the westernmost longitude.

To refine the selection, two shapefiles were downloaded from the Flanders Marine Institute (2025)— one representing the continental EEZ and the other outlining Portugal’s internal waters—and were used to filter the grid (Figure 3.12). Only the grid cells corresponding to these maritime zones were retained. This refined grid was then utilised to select inspections that specifically took place within the defined areas.

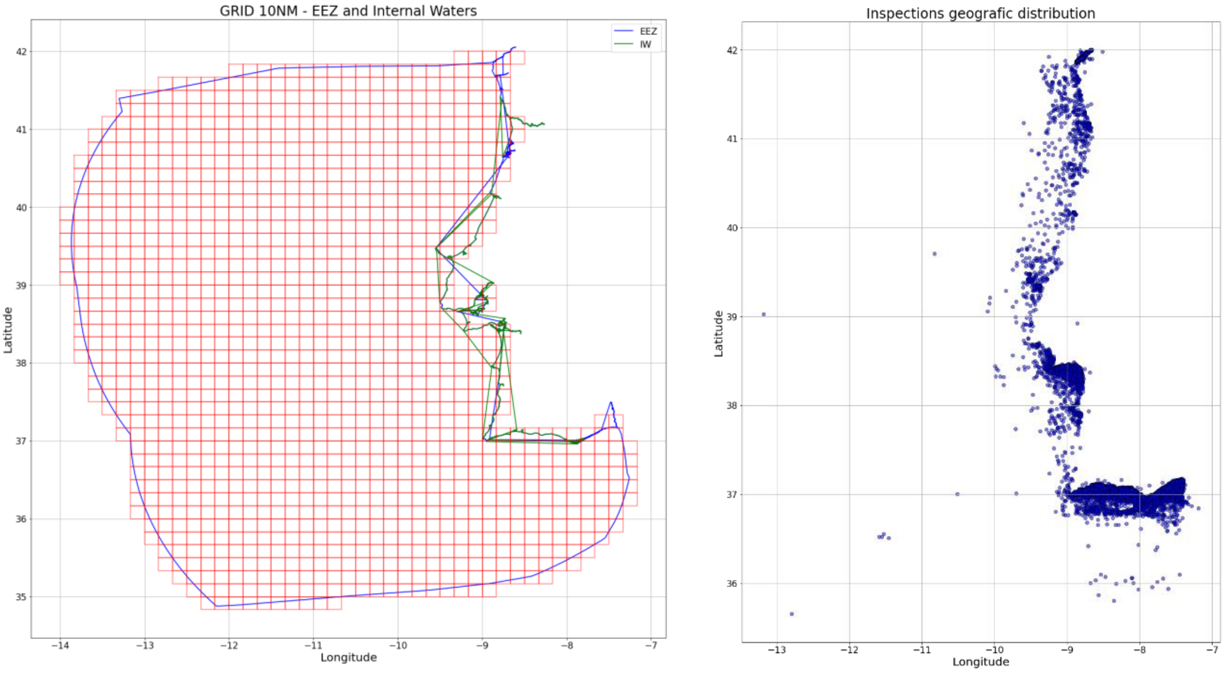


Figure 3.12 – Grid defined according to Portugal EEZ and internal waters limits, and the inspection's geographic distribution.

Regarding features, the following were discarded:

- "GDH" – The information contained in this feature is already split across other columns.
- "Vessel Subtype" and "Fishing Gear" – Given the large number of subtypes and fishing gears, the classification of the target vessel will be based solely on the broader "Type" category.
- "Registration Nº", "Ship Name", and "CFR" – The individual identification of inspected vessels is not relevant to the objectives of this project.
- "FISCREP Nº" – Apart from serving as a control mechanism for reports, this feature does not provide additional useful information.
- "Infraction (I) … (XIV)" – The specific type of infraction is outside the scope of this study.

3.4.4. Feature Engineering

In addition to the month and day period, the day of the week can also influence the occurrence of infractions (da Rocha Rei, 2016). Considering this, a new feature, "day of the week", was added to the dataset, derived from the year, month, and day values. Furthermore, to enable probability calculations for each area, the area code was included in the dataset based on the position values.

3.4.5. Dataset and Feature Selection

After the data cleaning process, the dataset consists of 15,323 historical inspection records and six features (Table 3.3), all of which are categorical.

Table 3.3 – Final dataset features description.

Designation	Description	Interval
Geographic Area	Area based on the 10x10 nm grid fit to the Portugal continental EEZ	1262 areas from A29 to AQ20
Month	The month when the inspection took place	January to December
6-hour daily period	One of 4th six hours day period when the inspection took place	]00:00 – 0600],]06:01 – 12:00],]12:00 – 18:00],]1800 – 00:00]
Day of the week	The day of the week when the inspection takes place	Monday to Sunday
Vessel Type	The type of activity that the vessel was engaged in	Commercial Fishing, Recreational, Set Nets, Maritime Tourism and Others
Result	Outcome of the inspection	Legal or Presumed Offender

The target feature, “Result”, is binary: 1 represents a *Presumed Offender* and 0 represents a *Legal* vessel. The “Geographic Area” and “Vessel Type” features are stored as string (object) types, while the remaining features (“Month”, “6-hour daily period” and “Day of the Week”) are stored as ordinal integers, reflecting the natural ordering inherent to their categories.

3.5. MODELLING

To effectively rank maritime areas based on the likelihood of detecting an illegal vessel, it’s crucial to choose the algorithm that best fits the available data and key contextual factors, such as geographic location, month, time of day, day of the week, and vessel type. To achieve this, we will employ several classifiers based on prior research:

- Naïve Bayes (Miguel et al., 2024; Navas de Maya et al., 2022),
- Logistic Regression (Davis & Harasti, 2020; Miguel et al., 2024; Navas de Maya et al., 2022; Pala et al., 2018; Wen et al., 2012),
- k-Nearest Neighbours (Pala et al., 2018),
- Classification Trees (Navas de Maya et al., 2022).

Additionally, to establish a baseline on the test set, we implemented a randomised algorithm that generates probabilities uniformly at random, thereby simulating chance-based decisions.

To ensure consistency in model training and prediction, features were encoded appropriately based on the nature of the algorithms used. Two complementary strategies were adopted: for models requiring integer inputs (NB and k-NN), an Ordinal Encoder was used. For models based on linear separators or tree splits (LR and CT), a one-hot encoder was applied so that any unknown category simply produced an all-zero row in the one-hot representation.

To ensure a fair comparison between models, all were evaluated under the same conditions, using identical training, validation, and test sets, as well as the same features and preprocessing techniques. Each model was trained and tested using both a balanced and a non-balanced version of the training dataset, allowing the impact of class imbalance on model performance to be systematically assessed.

Although our objective is to apply a model that outputs the probability of *Presumed Offender* for each area, it’s necessary to define a threshold to evaluate the models. A threshold of 0.5 was adopted to classify whether a case is considered positive. This value was chosen for two main reasons: first, in a binary classification setting, 0.5 represents the natural midpoint and the minimum probability at which a positive class prediction becomes justifiable; second, considering that actual offences are relatively rare, the probability of a presumed offence occurring is generally expected to be less than 0.5. Therefore, using 0.5 as a cutoff provides a conservative but balanced decision point across all models.

3.5.1. Naïve Bayes

The Naïve Bayes' Theorem is a fundamental principle in probability theory and statistics, expressed as follows (Joyce, 2003; Zhang, 2004):

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}$$

where:

- $P(x|y)$ represents the posterior probability of the event (the vessel being *Presumed Offender*) given the observed evidence ("Area", "Month", "Daily period 6 hours", "Day of the week", "Vessel Type").
- $P(y|x)$ denotes the likelihood, i.e., the probability of observing given that the vessel is a *Presumed Offender*.
- $P(x)$ is the prior probability of a vessel being a *Presumed Offender*, derived from historical inspection data.
- $P(y)$ is the marginal probability of the evidence, ensuring proper normalisation.

Given that all features in the dataset are categorical, we employ the CategoricalNB from Scikit-learn (2025b), which assumes that features are conditionally independent given the target class. This specific algorithm estimates a categorical distribution for each feature i of X, conditioned on the class y (Scikit-learn, 2025a) and the probability of the category t in feature i given the class c is estimated as:

$$P\left(x_i = t \,\middle|\, y = c; \alpha\right) = \frac{N_{tic} + \alpha}{N_c + \alpha n_i}$$

where:

- $N_{tic} = \left|\left\{j \in J \middle| x_{ij} = t, y_j = c\right\}\right|$ is the number of time categories t appears in the samples x_i which belongs to the class c.
- $N_c = \left|\left\{j \in J \middle| y_j = c\right\}\right|$ is the number of samples with class c.
- α is a smoothing parameter.
- n_i is the number of available categories of features i.

The set of samples is defined as $J = \{1,\ldots,m\}$, where m is the number of samples.

In this classifier, we set the smoothing parameter α to 1 to prevent issues arising from instances with no data available. During hyperparameter tuning, we compute evaluation metrics while varying whether the prior probability (*fit_prior*) is considered or not.

3.5.2. Logistic regression

Logistic regression (LR) is a special case of linear regression, but does not assume normality due to its binary response variable. It models the probability of an event as a linear function of explanatory variables, offering a simple probabilistic classification. However, it struggles with non-linear relationships and interactive effects among variables (Yeh & Lien, 2009).

In binary logistic regression, the target variable y takes values in {0,1} for each data point x. The model predicts the probability of the positive class, $P(y = 1|x)$, using the logistic function:

$$P(y = 1|x) = \frac{1}{1 + e^{-(xw+b)}}$$

where:

- x represents the feature vector,
- w are the model's learned weights,
- b is the intercept.

The optimisation problem for binary logistic regression with a regularisation term is to minimise the following cost function:

$$J(w) = -\sum_{i=1}^{n} w_i[y_i \log(\hat{y_i}) + (1 - y_i)\log(1 - \hat{y}_i)] + \lambda R(w)$$

where:

- $J(w)$ is the total cost to minimise,
- w_i represents the weight for the sample i, which is adjusted by class weights and sample weights,
- y_i is the true class label (0 or 1),
- $\hat{y}_i$ is the predicted probability of class 1,
- $R(w)$ is the regularisation term, which penalises large weights to prevent overfitting,
- λ is the regularisation parameter.

The regularisation term $R(w)$ can take different forms, depending on the chosen penalty type:

- **L1 (Lasso)**: $\sum |wi|$
- **L2 (Ridge):** $\sum w_i^2$
- **ElasticNet**: A combination of L1 and L2, controlled by the hyperparameter α.

For this model, we will use LogisticRegression from Scikit-learn (2025f) and set the solver to "liblinear" due to the size of the dataset (Scikit-learn, 2025c). The following hyperparameters will be optimised:

- *penalty*,
- *c* – Inverse of the regularisation strength,
- *class_weight* – Weights assigned to each class to handle class imbalance. The model will be tested with both balanced ("balanced") and unbalanced (None) settings.

3.5.3. k-Nearest Neighbour

The k-Nearest Neighbours (k-NN) classifier is a non-parametric method that classifies an unknown sample based on the majority class among its k closest neighbours, determined by a specified distance metric. Its main advantage is that it does not require a predefined predictive model; however, its performance is highly influenced by the choice of distance metric and the value of k (Yeh & Lien, 2009).

For this study, we will use the KNeighborsClassifier from Scikit-learn (2025d) and optimise the following hyperparameters:

- *n_neighbors* – Number of nearest neighbours considered for classification,
- *weights* – Defines how neighbours contribute to the classification decision, with two options: "uniform" (equal weight) and "distance" (closer neighbours have higher influence),
- *leaf_size* - Affects the efficiency of the nearest neighbour search,
- *metric* - Affects the distance; it will be tested with the Euclidean and Manhattan distances.

3.5.4. Classification trees

Classification trees (CT) are hierarchical models in which internal nodes test attributes, branches represent outcomes, and leaf nodes indicate classes. They classify observations based on explanatory variables, minimising impurity to create simple classification rules while handling non-linear interactions. However, their sequential nature makes them highly dependent on the training data, meaning that small changes can significantly alter the tree structure, limiting generalisation across different contexts (Yeh & Lien, 2009).

For this classifier, we will use the DecisionTreeClassifier from Scikit-learn (2025e), optimising the following hyperparameters:

- *criterion* – defines the measure of the quality of the split,
- *max_depth* – defines the depth of the tree,
- *min_samples_split* – defines the minimum number of samples required to split,
- *min_samples_leaf* – defines the minimum number of samples required to be at a leaf node,
- *class_weight* – defines the class weight. It will be tested for no weight and "balanced".

3.6. EVALUATION

To evaluate the predictive capability of the model, the following standard classification metrics were calculated (Vujović, 2021):

- **Accuracy**: Measures the proportion of correctly classified instances.
- **Precision**: The proportion of true positive (TP) predictions among all predicted positives.
- **Recall (Sensitivity)**: The proportion of TP correctly identified among the actual positives.
- **F1-score**: The harmonic mean of precision and recall, balancing both metrics.
- **ROC-AUC (Receiver Operating Characteristic - Area Under Curve)**: Evaluates the model's ability to discriminate between classes by plotting the true positive rate (TPR) against the false positive rate (FPR) at various thresholds.

Although all evaluation metrics were considered during the evaluation, the selection of the hyperparameter combinations process was primarily guided by maximising precision, assuming the same threshold. This choice reflects a deliberate effort to minimise false positives (FP)—cases in which the model incorrectly identifies a *Presumed Offender*.

Since the final objective is the probability itself, regardless of the threshold, another metric used to compare models was the Brier score. It evaluates the accuracy of probabilistic predictions in binary classification by measuring how close the predicted probabilities are to the actual outcomes. Lower values indicate better performance, with 0 representing perfect predictions and 1 the worst (Chicco et al., 2021).

3.7. DEPLOYMENT

To simulate the implementation of the models in the real world, a post-training phase was conducted after selecting the best-performing model configurations, including hyperparameters and the choice between balanced or unbalanced data. In this phase, the trained models were applied to the training dataset to generate predicted probability scores for each observation. This simulation aimed to assess how the models behave when classifying new instances, specifically analysing the distribution of predicted probabilities associated with the positive class (i.e., potential illegal activity).

Although controlled fallback mechanisms were used during the hyperparameter tuning and validation phases, the final model was trained using the same encoders and fallback strategies and subsequently tested on the designated test set. Both training and testing were conducted with consistent preprocessing, ensuring that the evaluation accurately reflected real-world application conditions while maintaining robustness to previously unseen categories.

3.8. ASSUMPTIONS AND LIMITATIONS

A fundamental assumption to consider when using NB is derived from the mathematical basis of the Naïve Bayes theorem: all features are assumed to be independent of each other concerning the target variable (Zhang, 2004).

Limitations become evident when analysing the dataset. As previously mentioned, the primary dataset exhibits a significant imbalance in the target feature, with more than 80% of all inspections resulting in a *Legal* outcome. However, to mitigate how it affects them, we use oversampling procedures. Additionally, its annual distribution, particularly over the past four years, shows a decline compared to previous years. While the “Year” feature itself is used to train the model, this trend may still impact performance, particularly concerning the feature representing the day of the week. Another unbalanced feature is the “Type”, given that the most inspected vessels were *Commercial Fishing*.

A possible additional limitation, given the objective of the study, is the geographic distribution of inspections. As shown in Figure 3.12, certain areas have a high concentration of inspections, while others have only a single inspection. Furthermore, when comparing the northern, central, and southern regions, the southern region has a significantly higher number of inspections. Besides that, we only have the probability for the areas where inspections occurred.

A further important limitation is that the dataset only includes vessels that have been inspected. This means it excludes potentially illegal activities or vessels that have never been detected or targeted. As a result, there is a risk of underestimating the true extent of non-compliance. This surveillance bias may cause the model to overfit to known patterns and overlook areas or behaviours where illegal activities remain undetected.

4. RESULTS AND ANALYSIS

In this chapter, we present and discuss the results concerning the research question:

"In what ways can historical inspection data enhance the selection of operational areas in maritime surveillance using machine learning models?"

To address this question, four classification models were trained, tested, and evaluated: NB, LR, k-NN, and CT. These models were developed using historical maritime inspection data and assessed on both the balanced version, using SMOTE, and the unbalanced version of the training dataset. The selection of hyperparameters was guided by maximising precision on the validation set.

The results presented refer to the models trained with the best-performing hyperparameters and subsequently evaluated on the test set. Both the validation and test datasets were created using stratified sampling to maintain the original distribution of *Legal* and *Presumed Offender* cases. As a benchmark, a random probability prediction algorithm was also applied to the test set to simulate chance-level performance.

The main objective - *"To test machine learning models for the development of a decision-support system for maritime surveillance, aiming to predict illegal activities in areas under national sovereignty based on spatiotemporal conditions."* - was to evaluate whether such data, when properly structured and modelled, can reliably predict the likelihood of illegal activity in maritime zones under Portuguese jurisdiction, thereby supporting more effective operational decision-making.

The results in this chapter are presented across four main dimensions:

1. Model performance based on standard metrics, prioritising precision.
2. The impact of using a balanced dataset on model behaviour.
3. Features importance on each model.
4. Model performance in simulation, through the analysis of predicted probability distributions.

4.1. MODEL EVALUATION: PRIORITISING PREDICTION CONFIDENCE

It is assumed that, given the operational context, it is crucial to minimise the risk of deploying resources to areas where there are no infractions. Therefore, the selection of the model's hyperparameter emphasised precision — the metric that reflects the proportion of TP among all areas flagged as *Presumed Offender*. This metric was prioritised in the context of limited maritime resources, where FP can lead to wasted patrols, and false negatives (FN) allow illegal activities to go undetected. Therefore, acting only on highly reliable alerts—those with a high likelihood of being correct—is operationally preferable, even if that means accepting that some infractions may be missed. Due to the unbalanced nature of the target variable, with only 20% of instances labelled as *Presumed Offender,* the four classification algorithms were trained using both balanced and unbalanced datasets.

Table 4.1 summarises the performance of the best configuration (by precision) for each model, based on the test dataset.

Table 4.1 – Model Performance on test dataset (Precision-Based ranking).

Model	Balanced	Precision	Recall	Accuracy	F1-Score	ROC_AUC	Brier Score
CT	No	0.71	0.02	0.85	0.04	0.63	0.12
LR	No	0.66	0.04	0.85	0.08	0.69	0.12
k-NN	No	0.60	0.08	0.86	0.14	0.68	0.12
NB	No	0.47	0.07	0.85	0.13	0.68	0.12
k-NN	Yes	0.26	0.37	0.75	0.31	0.63	0.20
LR	Yes	0.22	0.51	0.67	0.31	0.65	0.22
CT	Yes	0.21	0.57	0.61	0.30	0.63	0.22
NB	Yes	0.20	0.55	0.61	0.30	0.65	0.22
Random		0.14	0.47	0.49	0.21	0.47	0.34

The CT trained on the unbalanced dataset achieves the highest precision at 0.71, meaning that when it identifies as a *Presumed Offender*, it is correct in 71% of cases. In operational terms, this indicates a high degree of trust in each alert: most interventions based on the model's predictions are likely to target actual offender areas. This level of reliability is crucial when resources are limited and need to be allocated with confidence. However, this model captures only 2% of all true infractions (recall = 0.02), meaning that nearly all actual infractions remain undetected. This makes it highly conservative: it goes wrong on the side of caution, issuing warnings only when it is almost certain, but at the cost of large detection. According to the confusion matrix in Table 4.2, this model misses 441 true infractions, false negatives (FN), while making only 4 FP errors, evidencing its highly conservative behaviour. While it rarely raises incorrect alerts, it also fails to detect nearly all illegal activity, which can severely undermine the overall effectiveness of enforcement efforts.

The LR model exhibits similar tendencies, with a precision of 0.66 and slightly higher recall (0.04), identifying 19 TP out of 451 actual offences. Its balance between caution and detection is only marginally better than CT, as reflected in a low F1-score of 0.08.

The k-NN model demonstrates a more favourable trade-off between precision (0.60) and recall (0.08), achieving the highest recall among all unbalanced models, although low. It identifies 37 true infractions—nearly four times more than CT—at the cost of 25 FP. This makes it more suitable in operational settings where expanding detection is a priority, even if it leads to an increased rate of false alerts. With an F1-score of 0.14, k-NN offers the best overall balance among the unbalanced models.

Table 4.2 – Confusion matrices.

Model	Balanced	TN	FP	FN	TP	Precision	Recall
CT	No	2588	4	441	10	0.71	0.02
LR	No	2582	10	432	19	0.66	0.04
k-NN	No	2567	25	414	37	0.60	0.08
NB	No	2555	37	418	33	0.47	0.07
k-NN	Yes	2127	465	284	167	0.26	0.37
LR	Yes	1802	790	223	228	0.22	0.51
CT	Yes	1593	999	193	258	0.21	0.57
NB	Yes	1608	984	202	249	0.20	0.55
Random		1267	1325	241	210	0.14	0.47

The NB model ranks lowest in terms of precision (0.47), though it maintains a recall like k-NN (0.07), identifying 33 TP. However, it also produces the highest number of FP (37), leading to a modest F1-score of 0.13. Although it offers broader detection than CT and LR, its reliability per alert is significantly lower.

All four models exhibit high accuracy (between 0.85 and 0.86), a misleadingly positive result that reflects the dominance of the negative class in the dataset rather than true predictive skill. More meaningful in this context is the Brier score, which measures the calibration of probabilistic predictions. Each unbalanced model achieves a low Brier score (0.12), suggesting good alignment between predicted probabilities and actual outcomes, even if classification performance varies.

When compared to the random baseline, which achieves a precision of 0.14, a recall of 0.47, and a Brier Score of 0.34, the unbalanced models show clear advantages. Even the weakest among them (NB) more than triples the precision of random guessing, and the CT model improves it by a factor of over five. Moreover, its probabilistic predictions are significantly more reliable.

In summary:

- CT offers excellent alert reliability but detects virtually none of the true infractions.
- k-NN provides the best compromise, improving recall while maintaining acceptable precision.

  - LR is moderately effective, offering slightly improved coverage than CT, but is still limited.
  - NB achieves relatively high detection at the cost of low precision.

These findings highlight the operational trade-offs inherent in each model: CT prioritises precision and caution; k-NN aims for broader detection with acceptable risk; while LR and NB occupy intermediate positions, balancing precision and recall to varying extents.

4.2. EFFECT OF DATA BALANCING

Applying balanced training data leads to a significant change in model behaviour, moving from highly conservative, precision-driven predictions to broader detection through significantly enhanced recall. Under these conditions, models such as NB and CT achieve recall rates exceeding 0.55, indicating their capacity to identify more than half of all actual presumed offenders. This trend, observed across multiple algorithms, suggests that when class imbalance is corrected, the models are more capable of learning and generalising from infraction patterns.

However, this improvement in recall comes at a considerable cost to precision**.** Most models trained on balanced data exhibit precision values between 0.20 and 0.22, meaning that approximately four out of every five alerts are FP. The CT model exemplifies this shift clearly: while its unbalanced counterpart issued only four FP with ten TP, the balanced version produces 999 FP along with 258 TP. This dramatic increase in false alarms reflects a transition from focused, pattern-specific classification to broader, more inclusive decision boundaries. In operational terms, this means that enforcement resources would be misallocated in most cases, resulting in lower efficiency.

Despite this decline in precision, balanced models offer significant improvements in coverage**.** In scenarios where the primary objective is to avoid missing infractions—even at the cost of more false alerts—this trade-off may be acceptable. Among these models, k-NN demonstrates the most favourable compromise: it achieves the highest precision (0.26) and a respectable recall (0.37)**,** resulting in the best F1-score (0.31) among balanced configurations. In absolute terms, it identifies 167 TP while generating 465 FP, providing a detection rate over five times higher than the unbalanced CT, although with greater operational responsibility.

The LR model follows a similar pattern, with a recall of 0.51 and a precision of 0.22, closely aligning with NB and CT in terms of classification behaviour. Although LR is often favoured for its interpretability and robustness, it does not offer distinctive performance benefits under balanced conditions. Their confusion matrix reveals 228 TP but also 790 FP, reinforcing concerns about practical deployment where the costs of misclassification are high.

An additional effect of data balancing is observed in the Brier score**,** which increases in all models (ranging from 0.20 to 0.22). This indicates that probability estimates become less accurate and less well-calibrated under balanced training conditions. While the models become better at identifying infractions, they do so with reduced confidence in their predictions—an important consideration for real-world scenarios in which decision-makers rely not just on binary classifications but also on the trustworthiness of probability estimates.

In conclusion, data balancing significantly enhances recall and expands infraction detection, but it comes at the expense of alert quality and calibration. For missions that prioritise wide coverage and tolerance to FP, such as exploratory or preventive operations, models such as balanced k-NN may represent a viable solution. However, in contexts where efficiency and precision are paramount, the use of balanced data can undermine operational goals due to increased uncertainty and inadequate allocation of efforts.

4.3. FEATURES IMPORTANCE

To understand the role played by individual predictors across distinct modelling approaches, a detailed analysis of feature importance for each classifier was conducted. Due to the different internal mechanisms of the models, the method of quantifying the influence of the features was adapted accordingly. This section presents a comparative analysis based on these model-specific approaches, all implemented using the Scikit-learn library (Pedregosa et al., 2011).

4.3.1. Classification Tree

Feature importance was extracted using the built-in *feature_importances_* attribute, which quantifies the impurity reduction attributed to each feature (Scikit-learn, 2025i). Since categorical features were one-hot encoded, importances were aggregated across original variable names. The results (Figure 4.1) revealed “Area” as the dominant feature (0.646), followed by “Type” (0.197) and “Month” (0.089). The remaining variables, “Period” and “Day of the week”, contributed minimally to the decision-making process.

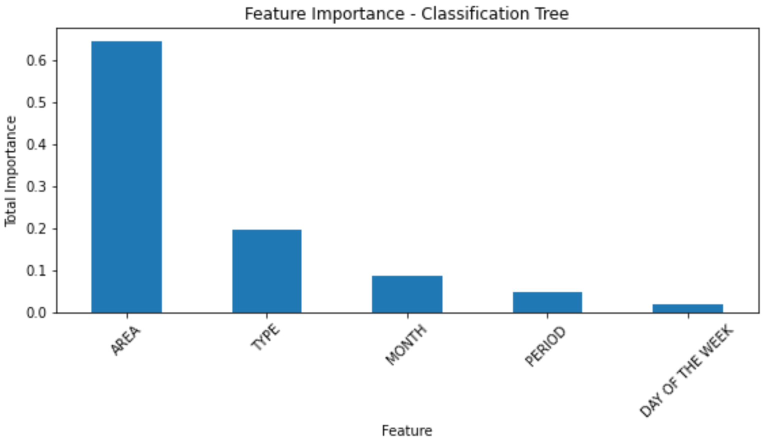


Figure 4.1 – Feature importance – Classification Tree.

4.3.2. Logistic Regression

For LR, the influence of the features was assessed through the average of the absolute coefficient magnitudes per original variable (Scikit-learn, 2025h). Due to the one-hot encoding of categorical features and the presence of numerical variables, standardisation was applied. The most influential feature was “Type” (0.377), followed by “Area” (0.296) and “Period” (0.153), whereas “Month” (0.038) and “Day of the week” (0.026) had a smaller impact (Figure 4.2).

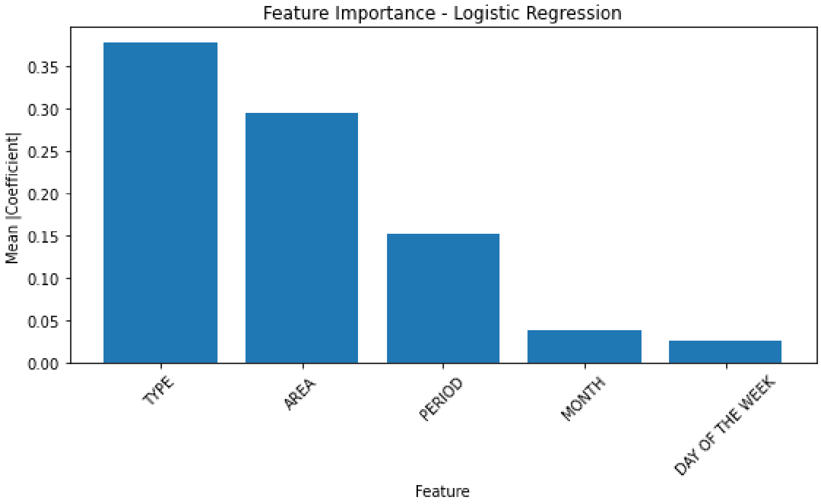


Figure 4.2 – Feature Importance – Logistic Regression.

4.3.3. Naïve Bayes

Since NB calculates class-conditional probabilities for each categorical feature, the importance of a variable was derived from the mean absolute difference in log-probabilities between classes (Scikit-learn, 2025b). Here too, “Area” (1.019) and “Type” (0.813) emerged as the most discriminative, with “Period”, “Month”, and “Day of the week” showing substantially lower class-separating capacity (Figure 4.3).

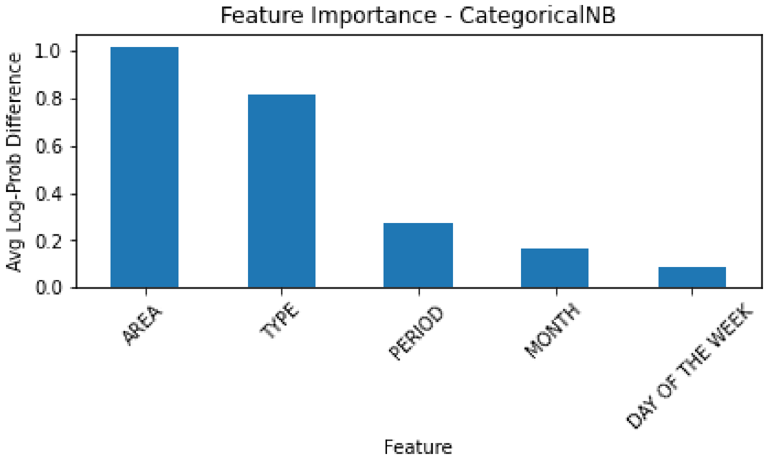


Figure 4.3 – Feature Importance – Naïve Bayes.

4.3.4. K-Nearest Neighbours

As k-NN is a non-parametric, instance-based algorithm, it does not generate internal model parameters, such as coefficients or splitting criteria, that can be directly interpreted to assess feature importance. Consequently, permutation importance (Scikit-learn, 2025g) was employed as an alternative measure of feature influence. This method evaluates the decrease in model performance, in this case, precision, when the values of a given feature are randomly permuted. The results revealed that “Area” was by far the most influential feature (0.366), followed by “Type” (0.095) and “Month” (0.069), which demonstrated a moderate impact. In contrast, “Period” showed negligible importance (0.001), indicating limited relevance in the context of distance-based classification (Figure 4.4).

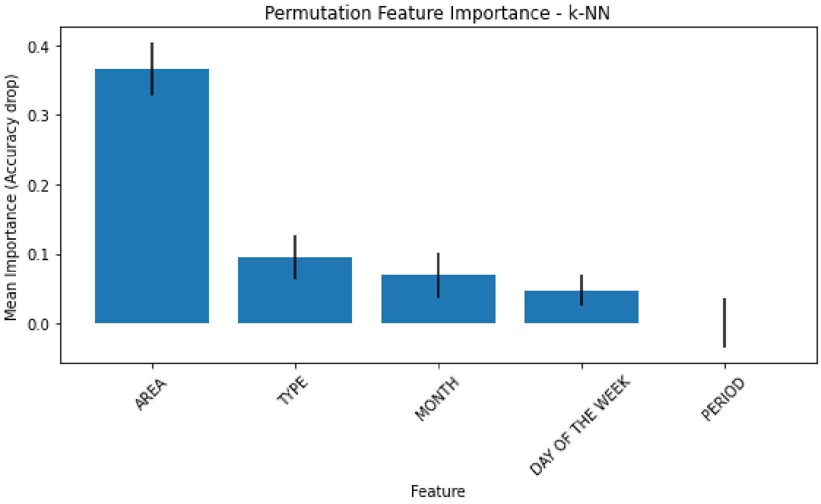


Figure 4.4 – Feature Importance – k-NN.

Across all models, “Area” and “Type” were consistently ranked as the most influential features. Notably, “Area” demonstrated the highest discriminative power in the CT, NB, and k-NN models, while “Type” proved most impactful in LR. In contrast, “Period”, “Month”, and “Day of the week” generally exhibited a weaker influence, especially in non-linear or distance-based models such as k-NN.

These findings underscore the importance of spatial and categorical contextual features in the classification task, with temporal features contributing to a reduced degree. Moreover, the consistency in feature rankings across models supports the robustness of the main identified predictors, thereby reinforcing their relevance regardless of model architecture.

4.4. SIMULATION BEHAVIOUR: DISTRIBUTION OF PREDICTED PROBABILITIES

As this project aims to explore the use of temporal historical data to predict the probability of *Presumed Offender* by area—thus allowing the ship’s captain to prioritise one area over another—it is essential to analyse how the probability forecasts behave when applied in practice and how each model spatially differentiates zones in terms of presumed offender risk. To this end, the four best-performing models were used to generate probability estimates for each area, considering the same values for month, period, day of the week, and type of vessel.

Figure 4.5 presents the spatial distribution of the predicted probabilities for each area on the map, offering a visual representation of how risk is spread across the EEZ.

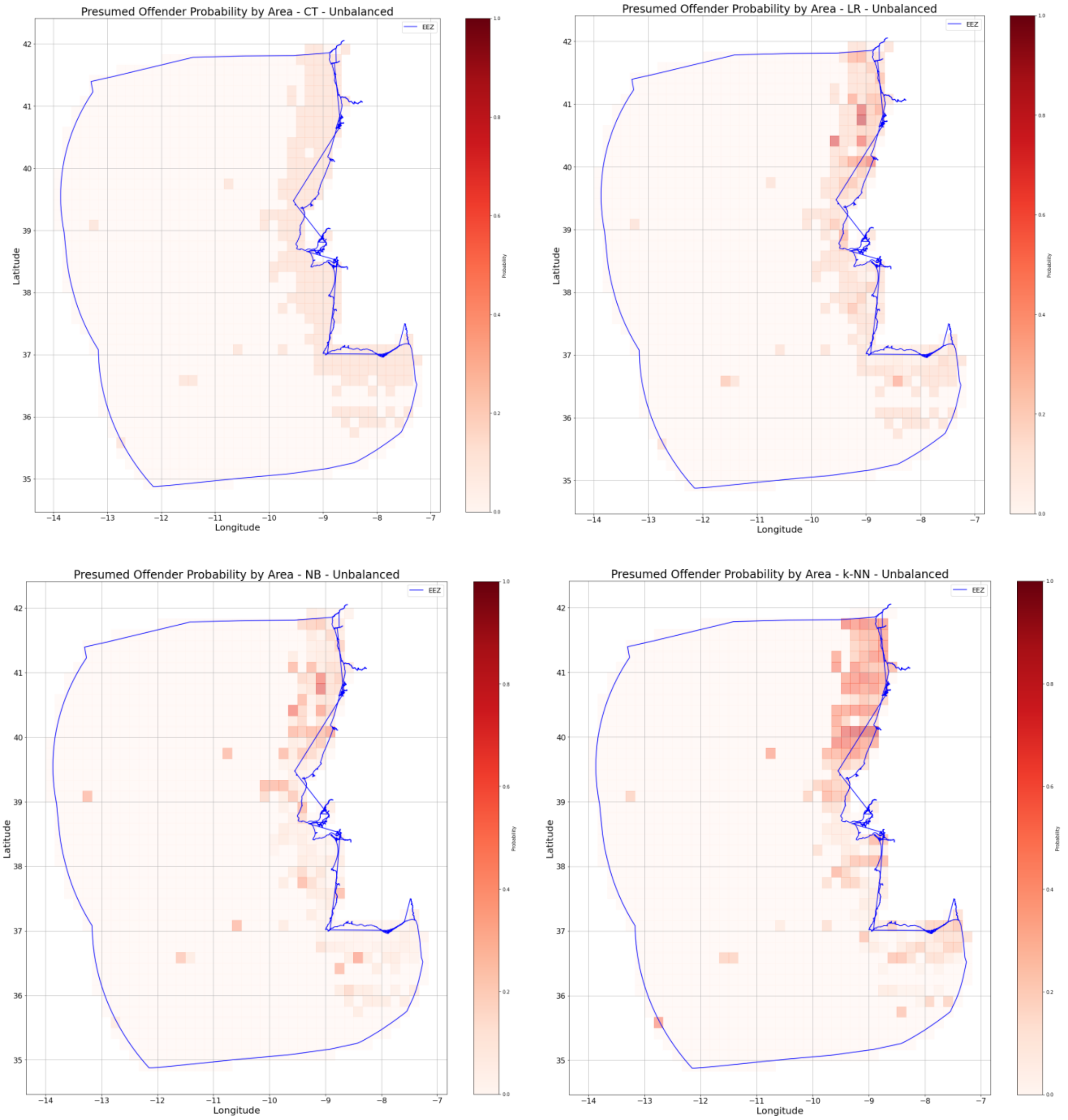


Figure 4.5 – *Presumed Offender* probability result by Area.

Complementarily, Figure 4.6 displays a histogram of these probabilities, providing insight into their overall distribution and concentration.

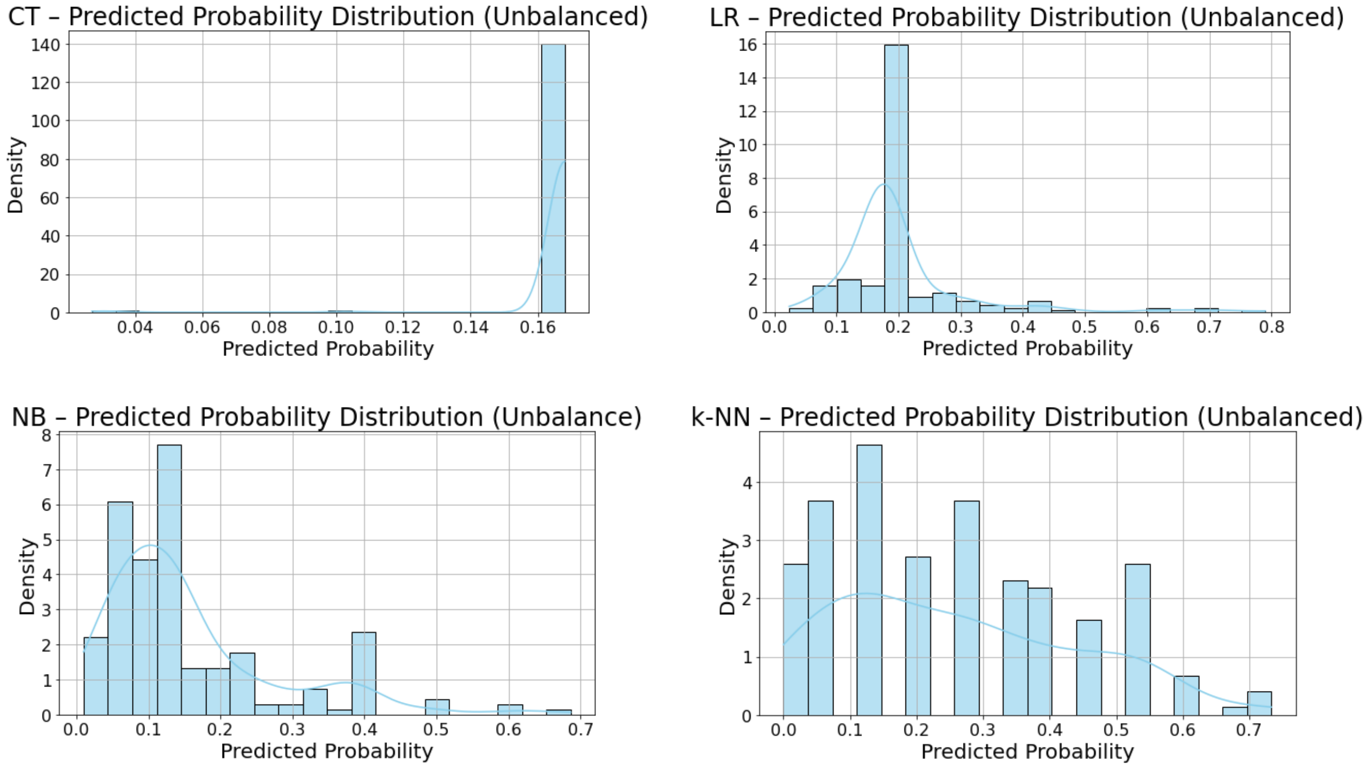


Figure 4.6 – *Presumed Offender* probabilities value distribution.

The CT model produces a substantially uniform spatial output, with almost all zones receiving the same probability (~0.17). The narrow histogram spike in Figure 4.6 confirms this. The lack of variation suggests the CT relies on a limited number of dominant decision divisions, failing to adapt spatially under simulation. Despite strong training results, it proves to be too rigid to offer meaningful spatial guidance.

The LR model displays a marked spatial differentiation, with certain zones reaching probabilities above 0.8. However, its histogram reveals a sharp peak around 0.2, indicating that while some areas are classified as high risk, the model tends to cluster most predictions into that intermediate level. This may reflect a cautious baseline probability, overlaid by location-specific adjustments that push a few predictions higher. The distribution suggests a combination of moderate baseline confidence with sporadic strong assertions, potentially arising from sensitivity to spatial patterns in the training data or to the influence of class imbalance.

The NB model presents a more balanced behaviour. Spatially, it identifies some areas of higher risk, keeping most others low and stable. Its histogram shows a concentrated distribution between 0.1 and 0.3. This conservative profile likely stems from the model's assumption of feature independence, which restricts complex interactions but enhances general stability.

The k-NN model reveals the greatest local variation. The map shows severe contrasts between neighbouring zones, and the histogram displays a multimodal distribution over a wide range of probabilities (up to 0.7). This suggests a strong sensitivity to training data density, enabling sharp local adaptation but risking inconsistency in under-sampled regions.

These simulation results reinforce the behavioural differences observed between the models. LR and k-NN offer the strongest spatial contrast, with LR showing a dominant peak around 0.2 and occasional high-probability zones, and k-NN exhibiting wide local variation due to its reliance on proximity-based learning. In contrast, NB adopts a more conservative stance, with stable outputs concentrated in a lower probability range. The CT, despite performing well during training, fails to differentiate meaningfully in space, displaying nearly uniform outputs. These behavioural profiles are critical for deployment scenarios, where reliable and interpretable probability estimates must guide operational prioritisation in dynamic maritime environments.

4.5. INTERPRETATION OF RESULTS

The results confirm that historical inspection data can provide significant support for more informed and targeted maritime surveillance. Through a comprehensive evaluation across model performance, dataset balancing, feature importance, and probability simulation, several key insights emerge regarding the practical utility and limitations of this data-driven approach.

Firstly, the use of precision-oriented model selection is appropriate in an operational context, where minimising FP is critical. Among the evaluated models, the CT trained on the unbalanced dataset achieved the highest precision (0.71), significantly outperforming the random baseline (0.14). This indicates that historical data can indeed capture patterns strongly associated with actual infractions, enabling highly reliable alerts—an asset in settings where patrol resources are limited.

However, this high precision came at a severe cost: the CT's recall was only 0.02, meaning that it failed to detect most infractions. This trade-off exemplifies the inherent tension between certainty and coverage. While CT avoids false alarms, it also overlooks most illegal activity. The same trend, though less extreme, was observed in LR and NB, both of which offered moderate precision (0.66 and 0.47, respectively) but with equally low recall. The k-NN model, in its unbalanced configuration, achieved the highest recall (0.08) among these models, detecting significantly more infractions while maintaining reasonable precision (0.60). This made it the most balanced performer under precision-prioritised conditions.

When training with balanced data, the model's behaviour shifted significantly. Recall increased dramatically, exceeding 0.50 in several cases, reflecting broader detection of presumed offenders. However, this came with a notable drop in precision (to around 0.20–0.22), leading to a high FPR. This trade-off suggests that while balanced training improves detection, it may compromise operational efficiency. Among the balanced models, k-NN again achieved the best F1-score (0.31), representing the strongest trade-off between coverage and reliability. Still, even this configuration generated more false alarms than accurate alerts.

Additionally, Brier score increased across all balanced models (rising to 0.20–0.22), indicating decreased calibration of predicted probabilities. This decline is operationally significant: poorly calibrated models can mislead decision-makers by overestimating low-risk zones or underestimating high-risk ones, thereby distorting patrol planning and resource allocation.

These performance dynamics were further examined through spatial simulations. Probability maps and histograms (Figures 4.5 and 4.6) illustrated the behaviour of each model under controlled input conditions, with location as the only variable. Both LR and k-NN models demonstrated strong spatial resolution and probability variation, making them valuable for zoning and prioritisation. LR tended to produce sharply defined risk gradients, with a dominant concentration around 0.2 and a few high-probability peaks. k-NN showed multimodal probability distributions and sharp contrasts between neighbouring zones, reflecting its sensitivity to local data density. This makes it responsive, but potentially unstable in undersampled areas.

In contrast, the NB model exhibited more cautious and consistent behaviour. Its probabilities were lower and more evenly spread, making it suitable for broad scanning operations where stability and robustness are preferable to fine-grained variation. Meanwhile, the CT model, despite its high training precision, produced near-uniform outputs (~0.17) across the EEZ in simulation, revealing poor adaptability to generalised spatial inputs and undermining its practical value for dynamic area prioritisation.

Taken together, these results show that historical inspection data hold substantial predictive value for MS. However, the ability to operationalise that value depends heavily on managing trade-offs between precision, recall, and calibration. No model performed best in all aspects, but each showed specific strengths aligned with operational needs, whether that be precision-led enforcement or wide-area detection. Table 4.3 summarises the operational profile of each model based on the hyperparameter selection, focusing on precision, offering a practical reference for decision-makers.

Table 4.3 – Models operational profile.

Model	Operational Fit	Strength	Limitation
CT	Targeted interventions	High precision	Misses most infractions
k-NN	Exploratory patrols	Balanced F1	Sensitive to data sparsity
NB	Broad scanning	Robustness	Moderate performance
LR	High-contrast zoning	Spatial variation	Risk of overfitting

Ultimately, these results demonstrate that ML algorithms, when properly trained on historical data, can enhance maritime decision support. However, its effective deployment demands alignment between model behaviour (precision, recall, and calibration) and strategic enforcement goals. Choosing the appropriate model is not a purely technical decision but a tactical one, guided by the mission priorities and constraints.

4.6. COMPARISON WITH LITERATURE

The literature review shows that most international studies on MS focus on real-time ship or anomaly detection, typically using AIS or SAR data (Agrawal & Ramteke, 2024; Bernabé et al., 2024; Do Nascimento et al., 2023; Ezzeddini et al., 2024; Ghosh et al., 2021; Guida et al., 2023; K et al., 2023; Liu et al., 2024; Newaliya & Singh, 2021; Pohontu et al., 2023; Wang et al., 2021; Yan et al., 2023). However, other studies, such as Wen et al (2012) and Navas de Maya et al. (2022) looked at the prediction of illegal activity based on former inspections. In the specific case of Portugal, maritime inspection-related works such as Vala (2023) and Miguel et al. (2024) highlight the importance of analysing historical data.

Observing the results, we can see that the selected features, temporal (month, day of the week and day period), space (area) and type, can be correlated and used to predict illegal activities. The same conclusions can be found in studies such as Davis and Harasti (2020) or Da Rocha Rei (2016), who observed a correlation between illegal activities and working days, or Miguel et al. (2024), who identified high-risk zones in the southern Portugal EEZ.

In terms of predictive metrics, the models in this study achieved precision values below those found in Ghosh et al (2021) or K et al. (2023), which exceeded 60%, or Wen et al (2012), which reached 76% for smuggling prediction. However, these detection-focused studies used data such as SAR imagery, well-labelled AIS streams or vessel characteristic data. Although it achieves a lower precision value, the best model, CT, isn't too far from the results of other studies.

In the specific case of Portuguese maritime zones, this study addresses a clear gap. Previous works, such as that of Miguel et al. (2024), focused only on the southern region and on the spatial occurrence of infractions, without incorporating temporal variables such as month or time of day. This work increases the scope geographically and methodologically by integrating temporal dimensions (month, day of the week and period of the day) and applying predictive modelling to support prioritisation decisions. It thus constitutes a novel contribution in the national maritime enforcement context. Besides that, prior works only look for the fishing vessels and, as he has seen, the feature "Type" shows some importance in these models.

Finally, by comparing the proposed approach with existing decision-support tools, this study reinforces previous concerns (Rodrigues, 2021; Vala, 2023) about the underutilisation of structured inspection data. It highlights the need to integrate insights derived from FISCREP into broader maritime surveillance systems, complementing tools such as SADAP.

5. CONCLUSIONS AND FUTURE WORKS

5.1. SYNTHESIS OF FINDINGS

This study evaluated the potential of historical maritime inspection data to improve operational area selection in MS by applying ML models. Specifically, it aimed to evaluate whether such data could support the development of a decision-support system capable of predicting illegal activities in maritime zones under national sovereignty, based on spatiotemporal conditions.

The results confirm that, when properly structured and modelled, historical inspection data offer meaningful predictive value, enabling more informed and strategically targeted decision-making than traditional approaches based solely on standard protocols or the ship commander's perception. This demonstrates that patrol planning informed by empirical data can outperform judgment-based decisions, particularly in extensive and complex maritime environments.

Three key findings emerged:

1. **Predictive value of historical data**: All models trained on unbalanced datasets outperformed the random baseline in terms of precision and recall, confirming that patterns embedded in past inspections are informative and can be used to prioritise areas with a higher likelihood of infractions. This highlights the potential of historical data to support proactive surveillance planning.
2. **Precision–recall trade-off**: Models trained on unbalanced data prioritised precision, producing fewer but more reliable alerts, minimising operational noise and unnecessary deployments. Conversely, balanced training substantially increased recall, capturing a greater number of potential infractions but at the expense of a higher FPR. This trade-off underscores the importance of aligning model configuration with operational goals: precision-focused strategies are preferable when enforcement capacity is limited, while recall-oriented approaches are better suited for deterrence missions or early detection frameworks.
3. **Simulation insights**: Spatial probability maps revealed distinct behavioural profiles across models. LR and k-NN showed strong spatial discrimination, making them suitable for highlighting high-risk zones. NB exhibited a more conservative and stable pattern, with probabilities mostly concentrated between 0.1 and 0.3, offering moderate but consistent risk assessments. CT, despite its high precision during testing, failed to meaningfully differentiate zones in simulation, limiting its utility for spatial prioritisation.

In summary, the findings demonstrate that historical inspection data can substantially contribute to more effective MS. However, implementation in the real world requires not only accurate models but also thoughtful consideration of how prediction outputs translate into actionable insights. As such, model selection is not a purely technical decision—it must reflect the practical constraints and enforcement philosophy of maritime operations.

5.2. IMPLICATIONS OF FINDINGS

The findings of this study carry implications at both strategic and operational levels. At the operational level, the models developed can serve as a foundation for a decision-support system that assists patrol planning based on historical patterns of infraction, rather than relying solely on the presence of vessels or static visualisations used by commanders. The decision regarding which model to adopt should be guided by the specific objectives of the Portuguese Navy's mission. Nevertheless, by guiding resources toward areas with higher predicted risk, this approach may increase deterrence and optimise inspection outcomes.

A feasible implementation could involve deploying the models as an interactive dashboard, integrated into platforms such as SADAP or linked to the BI solution proposed by Beja (2023) to provide commanders with dynamic patrol recommendations based on risk probabilities.

Ultimately, these findings reinforce the position of historical inspection data as a valuable operational asset. Success will also depend on the model's transparent behaviour, interpretability, and ability to gain user trust—factors that are critical in high-stakes, real-time operational contexts.

5.3. LIMITATIONS AND CHALLENGES

Despite the promising results, the study faces several noteworthy limitations:

- **Class imbalance**: Most records reflect legal situations (>80%), making it difficult to learn patterns associated with infractions. Nevertheless, and hopefully, vessels engaging in illegal activities are presumably fewer, so the records are consistent with reality. However, this imbalance poses challenges for training ML algorithms, as most are inherently biased towards the majority class unless compensatory techniques are applied. It is also worth noting that the number of inspections carried out on commercial fishing vessels exceeds those conducted on other types of vessels.
- **Uneven geographical distribution**: Some areas had very few records, reducing the reliability of estimations.
- **Incomplete or inconsistent data**: Certain variables, such as the type of vessel or fishing gear used, were missing from some records, requiring imputation and simplification.

- **Limited generalisability**: Models only classify areas with inspection history. Application in regions without records would require extrapolation or additional training. Besides that, areas with few inspections can be affected by model generalisation.
- **Surveillance bias**: The dataset only includes information from inspections performed. This means it does not account for illegal vessels or activities that have never been detected or inspected, which may underrepresent the true scope of infractions.

Despite these limitations, the study confirms the feasibility of employing historical data and ML algorithms to support MS area selection.

Given the operational context and the systems currently in use, a phased deployment strategy is likely to be the most practical approach. Implementation should begin with offline testing during pre-mission planning, allowing gradual integration into live operations after validation of models through performance analysis and crew feedback.

However, a key limitation lies in the inherent difficulty of validating model predictions in real time. Since presumed infractions can only be confirmed through onboard inspections, it becomes challenging to measure the models' true effectiveness in isolation. A rigorous operational evaluation would require simultaneous patrols by separate navy ships in different areas under identical conditions—something that is logistically complex and operationally costly.

Despite this constraint, and as discussed previously, the models consistently outperformed random selection baselines, demonstrating that even limited integration into planning workflows can yield more targeted and effective patrol routes than intuition alone.

5.4. FUTURE RESEARCH DIRECTIONS

Future research should focus on increasing model robustness, diversifying data sources, and extending applicability to a wider range of maritime contexts and vessel typologies. Given the relevance of historical inspection data in identifying risk patterns and supporting the definition of operational priorities, and recognising the added value of integrating complementary data sources, future research could adopt a more comprehensive and multidimensional approach. These investigations should also build on previous work carried out in the field of MS, particularly concerning maritime zones under Portuguese sovereignty or jurisdiction, where effective oversight and resource allocation are of growing strategic importance.

Within this context, the following possibilities for future research are proposed:

- **Exploring more sophisticated predictive models:** This could include the assessment of various deep learning architectures, such as feedforward ANN used by Wen et al. (2012), and recurrent models like LSTM, employed by Do Nascimento et al. (2023) for time-series forecasting.

- **Operational interface development:** Intuitive and interpretable dashboard design, tailored to the cognitive and workflow needs of ship commanders, ensuring usability in dynamic conditions.
- **Integration of geo-referenced meteorological data:** Incorporation of weather data specific to the areas under analysis, to more accurately reflect environmental factors that influence risk patterns and maritime activity.
- **Use of navigation monitoring data:** Integration of data from systems such as AIS and SAR imagery, to enrich situational awareness and improve the detection of anomalous behaviour.
- **Application of predictive models for targeted inspection selection:** Development of approaches that enable the prioritisation of individual vessels for inspection based on risk indicators, rather than relying solely on selection based on geographic area.
- **Route optimisation for autonomous platforms:** Design of predictive-navigation modules to support UAV and USV deployment, particularly in scenarios requiring flexible and high-precision coverage of smaller or high-risk maritime zones.

Ultimately, while integrating predictive models into military or enforcement contexts requires caution and validation, the Portuguese Navy already possesses the foundational systems (e.g. SADAP) necessary for gradual adoption. This facilitates realistic expectations for implementation, particularly if such tools are introduced in collaboration with end-users and through iterative prototyping.

BIBLIOGRAPHICAL REFERENCES

Agrawal, A., & Ramteke, M. (2024). Traveling of multiple salesmen to dynamically changing locations for satisfying multiple goals. *Digital Chemical Engineering*, *11*, 100149. https://doi.org/10.1016/j.dche.2024.100149

Beja, J. A. O. (2023). *BI Solution, in Support of the Portuguese Navy´ s Fishery Surveillance* [Universidade Nova de Lisboa]. http://hdl.handle.net/10362/160447

Bernabé, P., Gotlieb, A., Legeard, B., Marijan, D., Sem-Jacobsen, F. O., & Spieker, H. (2024). Detecting Intentional AIS Shutdown in Open Sea Maritime Surveillance Using Self-Supervised Deep Learning. *IEEE Transactions on Intelligent Transportation Systems*, *25*(2), 1166–1177. IEEE Transactions on Intelligent Transportation Systems. https://doi.org/10.1109/TITS.2023.3322690

Bhattacharjee, S., Shanmugam, P., & Das, S. (2023). A Deep-Learning-Based Lightweight Model for Ship Localizations in SAR Images. *IEEE Access*, *11*, 94415–94427. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3310539

Biçen, S., & Celik, M. (2024). A deep learning approach to analyse ship inspection reports via natural language processing integrated with artificial neural network. *Journal of Marine Engineering & Technology*, *23*, 1–14. https://doi.org/10.1080/20464177.2024.2353407

Bueger, C., & Edmunds, T. (2017). Beyond seablindness: A new agenda for maritime security studies. *International Affairs*, *93*(6), 1293–1311. https://doi.org/10.1093/ia/iix174

Chapman, P. (2000). *CRISP-DM 1.0: Step-by-step data mining guide*. SPSS. https://api.semanticscholar.org/CorpusID:59777418

Chaumette, P. (2016). *Maritime areas: Control and prevention of illegal traffics at sea* (P. Chaumette, Ed.). https://hal.science/hal-01525298

Chen, X., Wu, S., Liu, Y., Wu, W., & Wang, S. (2022). A patrol routing problem for maritime Crime-Fighting. *Transportation Research Part E: Logistics and Transportation Review*, *168*, 102940. https://doi.org/10.1016/j.tre.2022.102940

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*, *9*, 78368–78381. https://doi.org/10.1109/ACCESS.2021.3084050

Chircop, P. A., Surendonk, T. J., van den Briel, M. H. L., & Walsh, T. (2022). On routing and scheduling a fleet of resource-constrained vessels to provide ongoing continuous patrol coverage. *Annals of Operations Research*, *312*(2), 723–760. https://doi.org/10.1007/s10479-021-04474-6

da Rocha Rei, F. (2016). *Estatística multivariada aplicada à fiscalização da Pesca: Modelação do comportamento das embarcações de pesca para apoio à decisão de fiscalizar* [Universidade Nova de Lisboa]. http://hdl.handle.net/10362/20351

Dästner, K., von Haßler zu Roseneckh-Köhler, B., Opitz, F., Rottmaier, M., & Schmid, E. (2018). Machine Learning Techniques for Enhancing Maritime Surveillance Based on GMTI Radar and AIS. *2018 19th International Radar Symposium (IRS)*, 1–10. https://doi.org/10.23919/IRS.2018.8447961

Davis, T., & Harasti, D. (2020). Predictive modelling of illegal fishing in no-take marine protected areas. *Fisheries Management and Ecology*, *27*(3), 292–301. https://doi.org/10.1111/fme.12412

de Lima Filho, G. M., Passaro, A., Delfino, G. M., & Monsuur, H. (2023). A Novel Bias-TSP Algorithm for Maritime Patrol. *IEEE Access*, *11*, 28190–28198. IEEE Access. https://doi.org/10.1109/ACCESS.2023.3252013

De Wit, J. J., Klein, K. T., & Otten, M. P. (2023). *Compact radar technology development for small maritime unmanned aerial vehicles*. 217–220. https://doi.org/10.23919/EuRAD58043.2023.10289147

Decree Law No. 43/2002, 43/2002, Ministry of National Defence, Republic Diary No. 52/2002, Series 1 of 2002-03-02 1750 (2002). https://diariodarepublica.pt/dr/detalhe/decreto-lei/43-2002-251976

Decree Law No. 185/2014, 185/2014, Ministry of National Defence, Republic Diary No. 250/2014, Series 1 of 2014-12-29 6397 (2014). https://diariodarepublica.pt/dr/detalhe/decreto-lei/185-2014-65983262

Direção-Geral de Recursos Naturais, Segurança e Serviços Marítimos. (2025, September 27). *Zonas Marítimas sob Jurisdição e ou Soberania Nacional—DGRM*. https://www.dgrm.pt/am-ec-zonas-maritimas-sob-jurisdicao-ou-soberania-nacional

Do Nascimento, V. D., Alves, T. A. O., Dutra, D. L. C., & Kundu, S. (2023). A Comparative Study of Fishing Activity Detection Approaches in Maritime Surveillance. *2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE)*, 173–179. https://doi.org/10.1109/CSCE60160.2023.00033

Ezzeddini, L., Ktari, J., Frikha, T., Alsharabi, N., Alayba, A., Alzahrani, A. J., Jadi, A., Alkholidi, A., & Hamam, H. (2024). Analysis of the performance of Faster R-CNN and YOLOv8 in detecting fishing vessels and fishes in real time. *PeerJ Computer Science*, *10*, e2033. https://doi.org/10.7717/peerj-cs.2033

Flanders Marine Institute. (2025, February 5). *Marine Regions*. https://marineregions.org/

Gamage, C., Dinalankara, R., Samarabandu, J., & Subasinghe, A. (2023). A comprehensive survey on the applications of machine learning techniques on maritime surveillance to detect abnormal maritime vessel behaviors. *WMU Journal of Maritime Affairs*, *22*(4), 447–477. https://doi.org/10.1007/s13437-023-00312-7

Ghosh, S., Konugurthi, P. K., Shankar Rao Singupurapu, G., Patel, S., Tammanagari, T., Rao Desu, M., Thakar, L. K., & Ghara, I. (2021). On-Board Ship Detection for Medium Resolution Optical Sensors. *Sensors (Basel, Switzerland)*, *21*(9), 3062.

Guida, R., Rodger, M., Bissonauth, V., Soreefan, Z., Hurnath, P., Matthews, M., & Elseoud, A. (2023). Nereus: A Space-Based Maritime Surveillance System for Fisheries Monitoring and Anomaly Detection. *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 2069–2072. https://doi.org/10.1109/IGARSS52108.2023.10282136

Jardim, B., Alpalhão, N., Sarmento, P., & de Castro Neto, M. (2022). The illegal parking score–Understanding and predicting the risk of parking illegalities in Lisbon based on spatiotemporal features. *Case Studies on Transport Policy*, *10*(3), 1816–1826. https://doi.org/10.1016/j.cstp.2022.07.011

Joyce, J. (2003). *Bayes' Theorem*. https://seop.illc.uva.nl/entries/bayes-theorem/

K, S., Vijai, A., & R, S. (2023). Automated Ship Classification and Tracking in Satellite Imagery using Advanced Deep Learning Models. *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–8. https://doi.org/10.1109/ICCCNT56998.2023.10307576

Kirszenblat, D., & Caelli, T. (2022). Mission Path Planning for Optimal Inverse Synthetic Aperture Radar Vessel Imaging. *Military Operations Research*, *27*(4), 69–86.

Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, *18*(17), 1–5.

Li, Y., Zhang, D., Fan, L., Ma, H., & Xu, Z. (2023). Performance Analysis of Ship Target Recognition in Multi-Band Infrared Images Based on Deep Learning. *2023 7th International Conference on Transportation Information and Safety (ICTIS)*, 432–436. https://doi.org/10.1109/ICTIS60134.2023.10243868

Liu, J., Li, J., & Liu, C. (2024). AIS-based kinematic anomaly classification for maritime surveillance. *Ocean Engineering*, *305*, 118026. https://doi.org/10.1016/j.oceaneng.2024.118026

Miguel, B., Simões, P., de Deus, R. G., & Natário, I. (2024). Bayesian Geostatistics Modeling of Maritime Surveillance Data. In O. Gervasi, B. Murgante, C. Garau, D. Taniar, A. M. A. C. Rocha, & M. N. Faginas Lago (Eds.), *Computational Science and Its Applications – ICCSA 2024 Workshops* (pp. 187–205). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-65343-8_12

Morando, E., Daffinà, F. C., Stahl, T., Corvino, M. M., & Pratola, C. (2023). Multi-Sensor Data Fusion to Enhance Maritime Situational Awareness. *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 6829–6831. https://doi.org/10.1109/IGARSS52108.2023.10282392

Moura, R., Pessanha Santos, N., Vala, A., Mendes, L., Simões, P., de Castro Neto, M., & Lobo, V. (2024). Fisheries Inspection in Portuguese Waters from 2015 to 2023. *Scientific Data*, *11*(1), 362.

Mourinha, A. (2012). Monitorização, Controlo e Fiscalização da Pesca. *Maria Scientia*, 83–122.

Navas de Maya, B., Arslan, O., Akyuz, E., Kurt, R. E., & Turan, O. (2022). Application of data-mining techniques to predict and rank maritime non-conformities in tanker shipping companies using accident inspection reports. *Ships and Offshore Structures*, *17*(3), 687–694. https://doi.org/10.1080/17445302.2020.1862530

Newaliya, N., & Singh, Y. (2021). A Review of Maritime Spatio-temporal Data Analytics. *2021 International Conference on Computational Performance Evaluation (ComPE)*, 219–226. https://doi.org/10.1109/ComPE53109.2021.9751726

Pala, A., Zhang, J., Zhuang, J., & Allen, N. (2018). Behavior analysis of illegal fishing in the Gulf of Mexico. *Journal of Homeland Security and Emergency Management*, *15*(1), 20160017.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Pinto, R. (2017). *Análise de dados da fiscalização da pesca–Estudo da evolução e exploração dos recursos piscícolas no ecossistema marinho português* [Escola Naval]. http://hdl.handle.net/10400.26/21064

Pohontu, A., Deliu, A. D., & Vertan, C. (2023). Ship Type Classification: A Handwriting Signature Verification Approach for Maritime Trajectories. *2023 8th International Symposium on Electrical and Electronics Engineering (ISEEE)*, 110–115. https://doi.org/10.1109/ISEEE58596.2023.10310540

Qian, Y., Chen, Q., Qian, W., Yu, X., Ren, K., Wan, M., Kong, X., & Wang, F. (2023). Combinatorial detection of ship and its wake via skeleton theory based on joint feature knowledge association. *Optik*, *288*, 171177. https://doi.org/10.1016/j.ijleo.2023.171177

Ramongassie, S., Taveneau, N., Calmettes, T., Richard, J., Challamel, R., Autran, O., Foix, V., & Durand, P. (2010). RADAR and AIS sensors constellation for global maritime surveillance. *2010 IEEE International Geoscience and Remote Sensing Symposium*, 3793–3796. https://doi.org/10.1109/IGARSS.2010.5651146

Ribeiro, C. V., Paes, A., & Oliveira, D. D. (2023). AIS-based maritime anomaly traffic detection: A review. *Expert Systems with Applications*, *231*, 120561. https://doi.org/10.1016/j.eswa.2023.120561

Rodrigues, T. A. P. (2021). *Planeamento de Ações de Fiscalização Marítima Suportadas por Análise de Dados Georreferenciados em Larga Escala*. http://hdl.handle.net/10400.26/38084

Saager, M., Behrendt, M., & Baber, P. (2021). Towards Improving Situation Awareness of Maritime Field Operators Using Augmented Reality. *HCI International 2021 - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence*, 337–349. https://doi.org/10.1007/978-3-030-90963-5_25

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Scikit-learn. (2025a, February 14). *1.9. Naive Bayes—Scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/naive_bayes.html#categorical-naive-bayes

Scikit-learn. (2025b, February 14). *CategoricalNB — scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/generated/sklearn.naive_bayes.CategoricalNB.html

Scikit-learn. (2025c, March 12). *1.1. Linear Models—Scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/linear_model.html#regularized-logistic-loss

Scikit-learn. (2025d, March 13). *1.6. Nearest Neighbors—Scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/neighbors.html

Scikit-learn. (2025e, March 13). *DecisionTreeClassifier—Scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html#sklearn.tree.DecisionTreeClassifier

Scikit-learn. (2025f, March 13). *LogisticRegression—Scikit-learn 1.6.1 documentation*. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html

Scikit-learn. (2025g, June 11). *5.2. Permutation feature importance—Scikit-learn 1.7.0 documentation*. https://scikit-learn.org/stable/modules/permutation_importance.html

Scikit-learn. (2025h, June 11). *Common pitfalls in the interpretation of coefficients of linear models*. Scikit-Learn. https://scikit-learn/stable/auto_examples/inspection/plot_linear_model_coefficient_interpretation.html

Scikit-learn. (2025i, June 11). *Feature importances with a forest of trees—Scikit-learn 1.7.0 documentation*. https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html

Şengül, B., Yılmaz, F., & Uğurlu, Ö. (2023). Safety–Security Analysis of Maritime Surveillance Systems in Critical Marine Areas. *Sustainability*, *15*(23), 16381. https://doi.org/10.3390/su152316381

Soldi, G., Gaglione, D., Forti, N., Simone, A. D., Daffinà, F. C., Bottini, G., Quattrociocchi, D., Millefiori, L. M., Braca, P., Carniel, S., Willett, P., Iodice, A., Riccio, D., & Farina, A. (2021). Space-Based Global Maritime Surveillance. Part I: Satellite Technologies. *IEEE Aerospace and Electronic Systems Magazine*, *36*(9), 8–28. IEEE Aerospace and Electronic Systems Magazine. https://doi.org/10.1109/MAES.2021.3070862

Vala, A. F. (2023). *Modelação e Análise Geoespacial das infrações pesqueiras na costa portuguesa* [Escola Naval]. http://hdl.handle.net/10400.26/48999

Vujović, Ž. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, *12*(6), 599–606.

Wang, Y., Rajesh, G., Mercilin Raajini, X., Kritika, N., Kavinkumar, A., & Shah, S. B. H. (2021). Machine learning-based ship detection and tracking using satellite images for maritime surveillance. *Journal of Ambient Intelligence and Smart Environments*, *13*(5), 361–371. https://doi.org/10.3233/AIS-210610

Wen, C.-H., Hsu, P.-Y., Wang, C., & Wu, T.-L. (2012). Identifying Smuggling Vessels with Artificial Neural Network and Logistics Regression in Criminal Intelligence Using Vessels Smuggling Case Data. In J.-S. Pan, S.-M. Chen, & N. T. Nguyen (Eds.), *Intelligent Information and Database Systems* (pp. 539–548). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28490-8_56

Xu, J., Song, R., & Cao, H. (2023). Multi-USV Target Search Algorithm Based on Markov Prediction Model in Ocean Environment. *2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, 1192–1197. https://doi.org/10.1109/ICMSP58539.2023.10170774

Xu, Y., Li, H., & Jin, K. (2017). Resource Management and Decision Support System for Border Patrol. *American Journal of Data Mining and Knowledge Discovery*, *12*(4), Article 4. https://doi.org/10.11648/j.ajdmkd.20170203.12

Yan, Z., Song, X., & Yang, L. (2023). Research on AIS Data Aided Ship Classification in Spaceborne SAR Images. *Proceedings of the 2022 11th International Conference on Computing and Pattern Recognition*, 179–185. https://doi.org/10.1145/3581807.3581833

Yeh, I.-C., & Lien, C. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, *36*(2, Part 1), 2473–2480. https://doi.org/10.1016/j.eswa.2007.12.020

Zhang, H. (2004). The optimality of naive Bayes. *Aa*, *1*(2), 3.

Zhu, T., Xiao, Y., & Zhang, H. (2023). Maritime patrol tasks assignment optimization of multiple USVs under endurance constraint. *Ocean Engineering*, *285*, 115445. https://doi.org/10.1016/j.oceaneng.2023.115445

APPENDIX A – ETHICS COMMITTEE REPORT

"Dear André Nunes,

Dear professor Bruno Jardim,

Thank you for submitting your Research Ethics Checklist for review. After careful consideration, we confirm that your study meets the ethical requirements necessary for approval, provided that the following key conditions are met:

Your research involves analyzing data provided by the Marinha, including information on date, location, vessel type, and inspection outcomes. While you currently have access to this data through your professional role, ethical approval is contingent upon obtaining formal written authorization from the relevant authority confirming:

- Explicit Permission for Research Use: A document stating that the Marinha authorizes the use of this data for academic purposes, specifying any restrictions or conditions.
- Compliance with Data Protection Regulations: Ensure adherence to national laws governing the use of sensitive or operational data, particularly regarding security and confidentiality.
- Clarification of Data Sensitivity: Given that the dataset relates to maritime law enforcement and potential illegal activities, additional precautions may be required to prevent the misuse or unintended disclosure of sensitive information.

Also, as the study aims to apply data mining techniques to predict the probability of illegality in maritime areas, we recommend:

- Risk of Misuse: Predictive models must be carefully designed to avoid misinterpretation or unjust profiling of specific vessels or areas.
- Operational & Security Implications: Since the data relates to maritime enforcement, ensure that findings do not compromise law enforcement strategies or national security concerns.
- Transparency & Accountability: Clearly define how predictions will be validated and interpreted to avoid ethical dilemmas in decision-making based on AI-driven insights.

Taking these suggestions into account, we confirm that you may proceed with the study, as we do not foresee any significant ethical concerns with the project in its current form.

Project No.: **DSCI2025-3-51801**

Project Title: **Applying Machine Learning Models for Decision Support in Maritime Surveillance: Predicting Illegal Activities in National Sovereignty Areas**

Principal Researcher: **André Francisco Taveira Seixas Nunes**

according to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered **APPROVED** on 14/03/2025.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change;

- Any unforeseen events or unexpected developments that merit notification;

- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 14/03/2025

NOVA IMS Ethics Committee

ethicscommittee@novaims.unl.pt

Cristina Oliveira

Gestora executiva do centro de investigação MagIC | *Executive manager of the Information Management Research Center (MagIC)*

Find out more about our research at https://magic.novaims.unl.pt/en/

Team member of RM Roadmap - Co-creating the future of Research Management (https://rmroadmap.eu/)

https://orcid.org/0000-0002-0887-7961"

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa