TIJER || ISSN 2349-9249 || © March 2024, Volume 11, Issue 3 || www.tijer.org

# A Comprehensive Analysis of Marine Life Pollution Using Machine Learning Techniques on historical Pollutant data.

**Dr. Santosh Kumar Singh [1], Anjali Singh[2] and Srishti Dubey[3]**

[1] H.O.D(IT), [2,3] PG Student

[1,2,3] Department of Information Technology, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai-401107, Maharashtra, India

**Abstract.** This study addresses the urgent and multifaceted issue of marine pollution, focusing on plastic waste accumulation and oil spills, particularly highlighting the diverse impacts of hydraulic, diesel, and bilge oils. Leveraging comprehensive datasets and advanced analytical techniques, including machine learning algorithms such as Multi-layer Perceptron (MLP) and Gradient Boosting, the research aims to uncover hidden patterns, assess pollution severity, and predict future trends. By analyzing global shares of plastics emitted to the ocean and assessing the impact of marine events across various regions and vessel types, the study provides actionable insights for policymakers and stakeholders. Results indicate the effectiveness of both MLP and Gradient-boosting classifiers in predictive modeling, with the MLP classifier demonstrating higher accuracy and the Gradient-boosting classifier offering a robust approach with its ensemble learning technique. The findings underscore the critical need for evidence-based strategies to mitigate marine pollution and safeguard ocean health for future generations.

**Keywords:** Marine pollution, Machine learning techniques, Oil spills, Hydraulic oil, Diesel oil, Bilge oil, Plastic pollutants, Historical data, Environmental challenges, Data-driven approaches

## 1.Introduction

Marine pollution is a pressing environmental issue with far-reaching consequences for ecosystems, biodiversity, and human well-being. Among the myriad pollutants threatening our oceans, plastic waste and oil spills stand out as particularly pervasive and damaging. Plastic debris, ranging from microscopic particles to large floating masses, poses a persistent threat to marine life through ingestion, entanglement, and the release of harmful chemicals. Similarly, oil spills, whether from maritime accidents, industrial activities, or illegal dumping, can have devastating effects on marine habitats, wildlife populations, and coastal communities.

In the context of oil spills, it is essential to recognize the diversity of oils involved and their respective impacts. Hydraulic oil, commonly used in machinery and hydraulic systems onboard ships, can contaminate marine environments during accidents or leaks. Diesel oil, a prevalent fuel for marine vessels, is another significant contributor to oil pollution, known for its toxicity and persistence in the environment. Additionally, bilge oil, a mixture of various substances including lubricants, fuel residues, and wastewater, poses unique challenges due to its complex composition and potential for widespread contamination.

To address the complex challenges posed by marine pollution, there is a growing need for comprehensive and data-driven approaches. Traditional methods of analysis often struggle to cope with the vast and heterogeneous nature of environmental datasets. However, recent advancements in machine learning offer promising solutions by enabling the extraction of valuable insights from large volumes of historical data.

This study aims to undertake a comprehensive analysis of marine pollution, focusing on both plastic waste accumulation and oil spill incidents, with specific emphasis on hydraulic, diesel, and bilge oils. By leveraging machine learning techniques, such as regression, classification, and clustering, we seek to uncover hidden patterns, identify pollution sources, and predict future trends.

Through the integration of diverse datasets spanning different marine environments and time periods, we aim to provide a nuanced understanding of pollution dynamics and their ecological consequences.

The findings of this research hold significant implications for environmental policy, conservation efforts, and sustainable resource management. By providing actionable insights into the drivers and impacts of marine pollution, we hope to empower policymakers, regulators, and stakeholders to enact evidence-based measures for pollution prevention and mitigation. Ultimately, our collective efforts are essential for preserving the health and resilience of marine ecosystems, ensuring their continued vitality for generations to come.

Fig.1 Oil Spill and Marine Plastic Pollution

## 2.Literature review

The analysis of twenty peer-reviewed articles focused on marine plastic pollution across 31 islands in the Atlantic Ocean and Caribbean Sea, revealing a predominant emphasis on microplastics (>5 mm), particularly on beaches. Despite their biodiversity significance, these islands have been relatively understudied, with research showing a recent increase in publications, mostly from the 2000s onwards. The sources of microplastics varied between the North/South Atlantic and Caribbean Sea, with marine-based sources prevailing in the former and land-based sources more common in the latter. Monteiro RC, do Sul JA, Costa MF. Plastic pollution in islands of the Atlantic Ocean. Environmental Pollution. 2018 Jul 1; 238:103-10. [1]

The review paper highlights various aspects of plastic pollution in coastal and marine environments, emphasizing the widespread distribution of microplastics in water, sediment, and biota. It outlines the detrimental ecological and socio-economic effects of plastic pollution, including entanglement, ingestion, habitat disruption, and negative impacts on tourism, fishery, shipping, and human health. Practical approaches such as 3Rs (Reduce-Recycle-Reuse), awareness campaigns, and policy interventions are recommended to mitigate plastic pollution and safeguard marine ecosystems. Thushari GG, Senevirathna JD. Plastic pollution in the marine environment. Heliyon. 2020 Aug 1;6(8).[2]

In areas where plastic pollution accumulates and natural mineralization processes are slow, the resulting negative outcomes are deemed practically irreversible. Potential impacts include disruptions to carbon and nutrient cycles, alterations in habitats across various ecosystems, and adverse effects on endangered or keystone species. The rational approach to address this global threat is to swiftly reduce plastic emissions by curbing the consumption of virgin plastic materials and implementing internationally coordinated waste management strategies. Ballerini T, Pen JR, Andrady A, Cole M, Galgani F, Kedzierski M, Maria LP, ter Halle A, van Arkel K, Zettler E, Amaral-Zettler L. *Plastic pollution in the ocean: what we know and what we don't know about* (Doctoral dissertation, Plastic and Ocean Platform; The camp).[3]

The origin of the term 'plastic' can be traced back to the Ancient Greek word 'plastikos', meaning shape able or mouldable. This characteristic has greatly contributed to its widespread adoption and exponential increase in global usage in recent decades. However, it also poses significant challenges for detecting, identifying, and quantifying plastic pollution in marine ecosystems. A review of current methods in marine plastic assessment highlights the pressing need for standardized methodologies and techniques worldwide to enable accurate evaluations and comparisons of plastic pollution and deepen our understanding of its impact on marine organisms and ecosystems. Vered G, Shenkar N. Monitoring plastic pollution in the oceans. Current Opinion in Toxicology. 2021 Sep 1; 27:60-8.[4]

The study conducted manipulative experiments in outdoor mesocosms to investigate the interactive effects of plastic pollution, ocean warming, and acidification on macrophyte detrital decomposition. Results revealed that high levels of plastic pollution significantly reduced the decomposition rate of kelp and eelgrass, while higher seawater temperatures increased decomposition rates. However, ocean acidification did not have a significant impact on macrophyte decomposition or nutrient liberation. MacLeod M, Arp HP, Tekman MB, Jahnke A. The global threat from plastic pollution. Science. 2021 Jul 2;373(6550):61-5.[5]

The provides an overview of the facts and effects of oil spills in oceans, highlighting the significant environmental and economic impacts on marine ecosystems, local economies, and coastal communities. It emphasizes the ongoing risk of oil spills despite technological advancements and regulatory measures, underscoring the need for improved response strategies and policies to mitigate future incidents and minimize their consequences. Zhang B, Matchinski EJ, Chen B, Ye X, Jing L, Lee K. Marine oil spills—oil pollution, sources and effects. In World seas: an environmental evaluation 2019 Jan 1 (pp. 391-406). Academic Press. [6]

Over the past twelve years, there has been a significant increase in the consumption of petroleum products in industrially developed nations, accompanied by notable environmental disasters such as the Torrey Canyon wreck and the Santa Barbara oil blow-out. These incidents have raised public awareness about the consequences of oil consumption and prompted increased interest in oil pollution and environmental issues among biologists, technologists, and the general public. Despite efforts by activists and legislation initiated since the 1920s, international regulations governing tanker traffic remain inadequate to address the scale of oil pollution in coastal waters. Nelson-Smith A. Oil pollution and marine ecology. London: Elek; 1972. [7]

Oil spills at sea primarily impact surface-associated species, and the damage caused by oil and dispersants after the Torrey Canyon incident was less severe than initially anticipated. While slicks can naturally disappear, factors such as bacterial decomposition and zooplankton ingestion of oil droplets may not significantly contribute to this process in marine environments. Additionally, on shores affected by oil spills, excess detergent has been found to cause more damage than oil itself, with repopulation following expected sequences but many areas remaining abnormal. Spooner M. Some ecological effects of marine oil pollution. In International Oil Spill Conference 1969 Dec 1 (Vol. 1969, No. 1, pp. 313-316). American Petroleum Institute.[8]

In their innovative approach, the researchers address the pressing issue of plastic waste by proposing a method to repurpose discarded HDPE bottles into high-performance oil-sorbent films. This solution not only provides a sustainable alternative to traditional disposal methods but also offers practical benefits, such as increased oil absorption capacity. By utilizing waste materials to create value-added products, they contribute to both environmental sustainability and efficient oil spill remediation efforts. Saleem J, Ning C, Barford J, McKay G. Combating oil spill problem using plastic waste. waste management. 2015 Oct 1;44:34-8.[9]

In light of the recent oil spill and plastic debris incident in the Fernando de Noronha Archipelago, environmental experts express concern over the ongoing challenges facing marine conservation efforts in Brazil. They emphasize the critical need for effective maritime surveillance systems to detect and respond to environmental accidents promptly. Additionally, they highlight the vulnerability of unique marine ecosystems, such as seagrass beds and mangroves, to the growing impacts of plastic pollution and accidents related to the fossil fuel industry and shipping activities. This event underscores the urgency for improved environmental monitoring and enforcement measures to protect biodiversity and ecosystem services in the South Atlantic region. Magalhaes KM, Rosa Filho JS, Teixeira CE, Coelho-Jr C, Lima MC, Souza AM, Soares MO. Oil and plastic spill: 2021 as another challenging year for marine conservation in the South Atlantic Ocean. Marine Policy. 2022 Jun 1;140:105076[10].

## 3.Problem Statement

i. Marine pollution, characterized by plastic waste accumulation and oil spills, poses a significant threat to ocean ecosystems globally.
ii. Plastic pollution, originating from various sources including improper waste disposal and industrial runoff, harms marine life through ingestion, entanglement, and habitat degradation.
iii. Oil spills, whether accidental or intentional, introduce toxic substances into marine environments, leading to acute and long-term ecological damage, including disruption of food chains and loss of biodiversity.

iv. The diverse array of oils involved in spills, such as hydraulic, diesel, and bilge oils, presents unique challenges for mitigation efforts due to their distinct chemical properties and ecological impacts.

v. Traditional approaches to addressing marine pollution often lack scalability and sophistication, hindering effective analysis of vast and heterogeneous datasets.

vi. The complexity of marine pollution dynamics, compounded by interconnected ecosystems and human activities, underscores the need for comprehensive and data-driven approaches to mitigation and prevention.

vii. There is a critical need to leverage advanced analytical techniques, such as machine learning, to analyse historical data on plastic pollution and oil spills, to gain deeper insights and develop evidence-based strategies for sustainable marine resource management.

viii. Interdisciplinary collaboration, innovative technologies, and concerted action at local, regional, and global levels are essential to address the multifaceted challenges of marine pollution and protect ocean health for future generations.

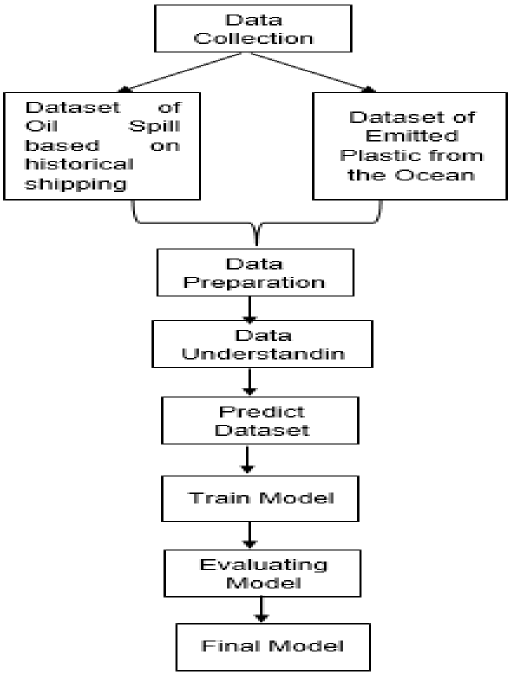
## 4.Methodology

The dataset provides a comprehensive view of marine activities and pollution incidents, encompassing event dates, geographical regions (such as Cairns, Brisbane, and Gladstone), vessel types (ranging from "Commercial" to "Defence"), maritime areas (including "Port" and "Port Limits"), specific event locations, and pollution severity ratings. It offers valuable insights into the temporal and geographical distribution of marine events, vessel classifications, and the environmental impact of incidents. Severity ratings for common pollutant categories like Bilge (0.0), Diesel (0.5), and Hydraulic Oil (0.1) aid in assessing and addressing the ecological consequences of marine pollution. Additionally, the dataset is crucial for studying the spread of events over time and geography, understanding various vessel types, and evaluating the environmental repercussions of pollution incidents. Moreover, the context highlights the significant impact certain countries, such as the Philippines, India, and Malaysia, have on marine pollution through their respective shares of global plastics emitted to the ocean. For instance, in 2019, India contributed 52.2% of global plastic waste to the ocean, while the Philippines and Malaysia accounted for 18.5% and 8.2%, respectively. This underscores the urgent need for global cooperation to address the environmental challenge posed by plastic waste pollution. Unfortunately, information regarding China and Indonesia's shares is not provided, limiting a comprehensive understanding of the issue.

### 4.1 Implementation

To analyse the dataset comprehensively, we will employ Python programming language along with libraries such as Pandas, Matplotlib, and Seaborn for data manipulation and visualization. This will enable us to perform Exploratory Data Analysis (EDA) on the provided information, including event dates, geographical regions, vessel types, maritime areas, pollution severity ratings, and global plastics emission shares. Additionally, we'll utilize statistical techniques and visualizations to gain insights into the temporal and geographical distribution of marine events, assess the impact of pollution incidents, and understand the severity of different pollutant categories. Moreover, we'll create specific visualizations such as histograms, and pie charts to effectively communicate our findings.

## 4.2 Flowchart



## 4.3 Analysing the Dataset

In 2019, the share of global plastics emitted to the ocean by country and the shipping of pollutant classes were as follows: The Philippines, India, Malaysia, China, and Indonesia were the top five contributors to ocean plastic pollution, with India being the most significant, accounting for 52.2% of the global plastic waste in the ocean. In terms of pollutant transportation, there were five classes, with Class 1 having the highest share (76.7%). Class 2 and Class 3 pollutant transportation made up 12.6% and 10.7%, respectively, while Class 4 had no contribution (0.0%). Lastly, Class 5 pollutant transportation accounted for a minimal share of 1.0%. This information highlights the urgent need for these countries and the global community to address plastic waste management and pollutant transportation to protect the environment.

Fig.2 Analyse the Data of Pollutants and Plastic Emitted to the Ocean

This given fig.3 the graph shows the top five Asian countries that contribute to the global share of plastics emitted into the ocean. The Philippines is at the top with a 50% share, followed by China with 35%, Indonesia with 25%, India with 15%, and Malaysia with 10%. The other countries not listed in the top five contribute to the remaining 10% of global plastics emitted to the ocean. This data highlights the significant impact of these five countries on ocean pollution and the urgent need for measures to reduce plastic waste. It is important to note that the percentages do not add up to 100, as each country's contribution is calculated independently and not relative to the others.

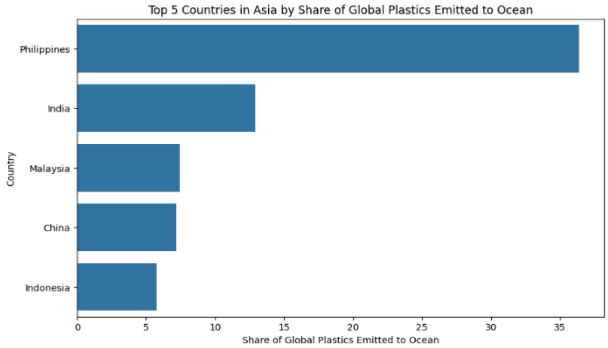


Fig.3 Analysis of the Asia Top 5 Countries' Plastic Emitted to the Ocean

## 5. Algorithms

### 5.1 Multi-layer Perceptron (MLP)

Multilayer Perceptron is a type of artificial neural network (ANN) that is widely used in machine learning for tasks such as classification and regression. MLPs are composed of multiple layers of interconnected neurons, each layer consisting of one or more nodes (also known as neurons) that perform computations on input data. The layers typically include an input layer, one or more hidden layers, and an output layer. MLPs are powerful models capable of learning complex patterns and relationships in data. Still, they also require careful tuning of hyperparameters, such as the number of hidden layers, the number of nodes in each layer, and the choice of activation functions, to achieve optimal performance. Additionally, they can be prone to overfitting if not regularized properly. However, with proper training and tuning, MLPs can achieve state-of-the-art performance on a wide range of machine-learning tasks.

### 5.2 Gradient Boosting Algorithm

Gradient Boosting is a powerful machine learning algorithm that works by combining the predictions of multiple weak learners, typically decision trees, into a strong predictor. It operates in an iterative manner, where each new weak learner focuses on the mistakes made by the previous ones. At each step, the algorithm calculates the gradient of the loss function concerning the predictions and fits a weak learner to the residuals. By sequentially minimizing the errors, Gradient Boosting constructs a robust model capable of making accurate predictions on a wide range of datasets.

## 6.Result and discussion

The analysis of the provided dataset revealed crucial insights into marine pollution and its contributors. In 2019, the top five contributors to global plastic pollution in the ocean were identified: India, the Philippines, Malaysia, China, and Indonesia. India stood out as the most significant contributor, accounting for 52.2% of the plastic waste in the ocean, followed by the Philippines (18.5%) and Malaysia (8.2%). Further examination of pollutant transportation highlighted those Class 1 pollutants constituted the highest share (76.7%), followed by Class 2 (12.6%) and Class 3 (10.7%), while Class 4 had no contribution and Class 5 accounted for a minimal share (1.0%). These findings underscore the urgent need for collaborative efforts among these countries and the global community to address plastic waste management and pollutant transportation, emphasizing the importance of implementing effective measures to mitigate environmental degradation caused by marine pollution. Additionally, machine learning algorithms, including the MLP Classifier and Gradient Boosting Classifier, showcased strong predictive capabilities in analyzing marine pollution data, providing valuable tools for assessing and addressing ecological consequences.

## 7.Conclusion

The MLP (Multi-Layer Perceptron) Classifier is a type of neural network model that consists of multiple layers of nodes, each connected to the next layer. It utilizes backpropagation to optimize weights and biases during training, aiming to minimize error and improve performance. With an accuracy of 87.5%, the MLP Classifier demonstrates strong predictive capability in this context. On the other hand, the Gradient Boosting Classifier is an ensemble learning technique that builds a strong predictive model by combining multiple weak learners sequentially. It fits new models to the residuals of previous models, with each subsequent model focusing on the errors of the previous ones. Despite a slightly lower accuracy of 85.42%, the Gradient Boosting Classifier still performs well, indicating its effectiveness in capturing complex patterns within the data. Overall, both algorithms showcase their strengths in predictive modeling, with the MLP Classifier demonstrating higher accuracy but the Gradient Boosting Classifier providing a robust alternative with its ensemble learning approach.

## 8.Future Prediction

In Fig. 4, the graph depicts both historical and forecasted data for a specific variable over a period from 72 to 80, with the y-axis representing values from 0 to 10 and the x-axis denoting time. Three lines are presented: actual (historical data), forecast (predicted values for future time periods), and a confidence interval indicating the range within which the actual value is likely to fall with a certain level of confidence. The actual values rise from 72 to around 76 before slightly declining, while the forecast predicts a continued decrease in the future, with the confidence interval widening as time progresses, suggesting less certainty in the forecasted values. Fig. 5 illustrates the growth in yearly plastic production from 1960 to 2040, with the x-axis showing years and the y-axis displaying plastic production in millions of metric tons. The graph demonstrates a clear upward trend in plastic production, particularly accelerating after 2000, as indicated by the trendline. Additionally, a forecast predicts further growth in plastic production beyond 2040. This synthesis highlights both the observed increase in plastic production and the forecasted continuation of this trend, underscoring the environmental concerns associated with escalating plastic usage and disposal.

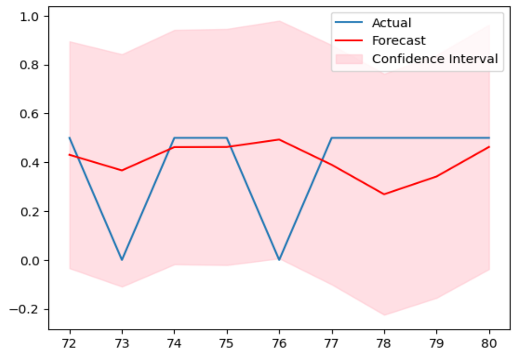


Fig.4 SARIMAX Model

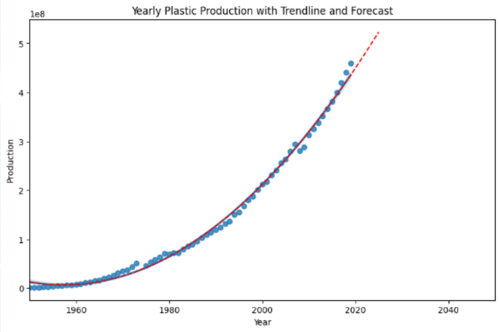


Fig.5 Yearly Plastic Prediction

## References

1. Monteiro RC, do Sul JA, Costa MF. Plastic pollution in islands of the Atlantic Ocean. Environmental Pollution. 2018 Jul 1; 238:103-10.
2. Thushari GG, Senevirathna JD. Plastic pollution in the marine environment. Heliyon. 2020 Aug 1;6(8).
3. Ballerini T, Pen JR, Andrady A, Cole M, Galgani F, Kedzierski M, Maria LP, ter Halle A, van Arkel K, Zettler E, Amaral-Zettler L. *Plastic pollution in the ocean: what we know and what we don't know about* (Doctoral dissertation, Plastic and Ocean Platform; The camp).
4. Vered G, Shenkar N. Monitoring plastic pollution in the oceans. Current Opinion in Toxicology. 2021 Sep 1;27:60-8.
5. MacLeod M, Arp HP, Tekman MB, Jahnke A. The global threat from plastic pollution. Science. 2021 Jul 2;373(6550):61-5.
6. Zhang B, Matchinski EJ, Chen B, Ye X, Jing L, Lee K. Marine oil spills—oil pollution, sources and effects. InWorld seas: an environmental evaluation 2019 Jan 1 (pp. 391-406). Academic Press.
7. Nelson-Smith A. Oil pollution and marine ecology. London: Elek; 1972.
8. Spooner M. Some ecological effects of marine oil pollution. InInternational Oil Spill Conference 1969 Dec 1 (Vol. 1969, No. 1, pp. 313-316). American Petroleum Institute.
9. Saleem J, Ning C, Barford J, McKay G. Combating oil spill problem using plastic waste. waste management. 2015 Oct 1;44:34-8.
10. Magalhaes KM, Rosa Filho JS, Teixeira CE, Coelho-Jr C, Lima MC, Souza AM, Soares MO. Oil and plastic spill: 2021 as another challenging year for marine conservation in the South Atlantic Ocean. Marine Policy. 2022 Jun 1;140:105076