A FRAMEWORK FOR MARITIME EMISSION PREDICTION AND OPTIMAL BIRTH SCHEDULING TO REDUCE ENVIRONMENTAL IMPACT

A project report submitted in partial fulfilment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY

Submitted by

G. Ojaswini Sree	(A21126511087)
S. Nivedita	(A21126511121)
K. Hemanth Sai	(A21126511101)
Aakifah	(A21126511066)

Under the Guidance
of
Dr. P. Saritha Hepsibha
Associate Professor

PRAGNANAM BRAHMA
ANITS

DEPARTMENT OF INFORMATION TECHNOLOGY
ANIL NEERUKONDA INSTITUTE OF TECHNOLOGY AND SCIENCES (AUTONOMOUS)
(Permanent Affiliation by Andhra University & Approved by AICTE)
Accredited by NBA (ECE, EEE, CSE, IT, Mech. Civil & Chemical) & NAAC A+
Sangivalasa-531 162, Bheemunipatnam Mandal, Visakhapatnam District

DECLARATION

We hereby declare that the project work entitled "**A Framework for Maritime Emission Prediction and Optimal Berth Scheduling to Reduce Environmental Impact"** submitted to the Anil Neerukonda Institute of Technology and Sciences is a record of an original work done by G.Ojaswini(A21126511087), S.Niveditha(A21126511121), K.Hemanth Sai(A21126511101), Aakifah(A21126511066) under the esteemed guidance of Dr.P.Saritha Hepsibha, Associate Professor, Dept of Information Technology, Anil Neerukonda Institute of Technology and Sciences. And this project work is submitted in the partial fulfillment of the requirements for the award of degree of Bachelor of Technology in Information Technology. This entire project is done with the best of our knowledge and has not been submitted for the award of any other degree in any other universities.

G. Ojaswini Sree (A21126511087)
S. Nivedita (A21126511121)
K. Hemanth Sai (A21126511101)
Aakifah (A21126511066)

ANIL NEERUKONDA INSTITUTE OF TECHNOLOGY AND SCIENCES

(AUTONOMOUS)

(Permanent Affiliation by Andhra University & Approved by AICTE)

Accredited by NBA (ECE, EEE, CSE, IT, Mech. Civil & Chemical) & NAAC A+

Sangivalasa-531 162, Bheemunipatnam Mandal, Visakhapatnam District

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that the project report entitled "**A Framework for Maritime Emission Prediction and Optimal Berth Scheduling to Reduce Environmental Impact**" submitted by G.Ojaswini Sree (A21126511087), S.Nivedita(A21126511121), K.Hemanth Sai (A21126511101), Aakifah (A21126511066) in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Information Technology of Anil Neerukonda Institute of technology and sciences, Visakhapatnam is a record of Bonafide work carried out under my guidance and supervision.

Project Guide

Dr.P.Saritha Hepsibha

Associate Professor

Department of IT

ANITS

Head of the Department

Prof. M. Rekha Sundari

Department of IT

ANITS

ABSTRACT

Despite being essential to international trade, the maritime sector greatly contributes to environmental pollution through carbon emissions and ineffective port operations. Emission forecasting and effective berth scheduling are two significant maritime logistics issues that are addressed in this study using an integrated approach that combines machine learning and optimization algorithms. The goal of the project is to create a predictive model that uses real-time data from environmental variables, fuel consumption, and vessel movements to accurately estimate maritime emissions. To improve the accuracy of emission forecasts, machine learning techniques will be applied to the analysis of both historical and current data. Furthermore, this study provides a framework for berth scheduling optimization that can meet different shipping demands while ensuring lower fuel consumption and fewer delays. Port operations will be streamlined using modern optimization algorithms that maintain both environmental sustainability and economic viability. By integrating these two components, the proposed approach enhances decision-making for maritime stakeholders and makes it possible to create informed scheduling and emission control plans. The results will reduce the carbon footprint of maritime logistics while improving operational effectiveness. This study is a significant step toward sustainable and intelligent maritime transportation, which will benefit the global economy and environment.

ACKNOWLEDGEMENT

We would like to express our deep gratitude to our project guide **Dr.P. Saritha Hepsibha**, Associate Professor, Department of Information Technology, ANITS, for her guidance with unsurpassed knowledge and immense encouragement. We are grateful to **Prof. M. Rekha Sundari,** Head of the Department, Information Technology, for providing us with the required facilities for the completion of the project work.

We are very much thankful to **our Principal, Prof. K. Sri Rama Krishna** and the Management,of ANITS, Sangivalasa, for their encouragement and cooperation to carry out this work.

We express our thanks to all teaching faculty of the Department of IT, whose suggestions during reviews helped us in accomplishment of our project. We would like to thank all non-teaching staff of the Department of IT, ANITS, for providing great assistance in the accomplishment of our project.

We would like to thank our parents, friends, and classmates for their encouragement throughout our project period. At last, but not the least, we thank everyone for supporting us directly or indirectly in completing this project successfully.

PROJECT STUDENTS

G. Ojaswini Sree (A21126511087)

S. Nivedita (A21126511121)

K. Hemanth Sai (A21126511101)

Aakifah (A21126511066)

PUBLICATION DETAILS

Dear Ojaswini,
Greetings and best wishes for the day !!

Our International Conference has accepted your Research Paper entitled **"A Framework for Maritime Emission Prediction and Optimal Berth Scheduling to Reduce Environmental Impact"** with Paper ID **WR-RSET-AHMD-210425-11642**

International Conference on Research in Science, Engineering and Technology

Note : There is no additional Fee for Certificates, Proceeding and Publication.

Online payment link http://wrfer.org/PAYMENT/

OR

Bank Details
EVER LIFE HEALTH PRIVATE LIMITED
Ac/no.: 50200052618798
HDFC Bank, Nayapalli, Bhubaneswar, Odisha, India
IFSC: HDFC0000640
SWIFT CODE: HDFCINBB

Only after the Registration you will get the Acceptance letter and Conference Schedule.

INDEX

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

S. No	Abbreviation	Full Form	Page No
1	ML	Machine Learning	2
2	MILP	Mixed Integer Linear Programming	2
3	CO_2	Carbon Dioxide	3
4	NO_x	Nitrogen Oxides	3
5	SO_x	Sulfur Oxides	3
6	IMO	International Maritime Organization	3
7	SMOTE	Synthetic Minority Over-sampling Technique	35
8	IoU	Intersection over Union	37
9	mAP	Mean Average Precision	37

CHAPTER 1
INTRODUCTION

1. INTRODUCTION

1.1 Introduction:

One essential element of international trade is maritime transportation. Maritime transportation is a vital component of international trade and logistics because it facilitates the movement of goods across borders. However, concerns regarding environmental sustainability are raised by the industry's significant contribution to greenhouse gas emissions, air pollution, and climate change. As ports and shipping operations expand, there is an increasing need for more efficient and sustainable maritime solutions.

One of the biggest environmental problems facing maritime operations is emissions from ships while they are berthing in ports and transiting through them. Ineffective berth scheduling increases carbon emissions by causing long wait times and excessive fuel use. To address this issue, a comprehensive system for forecasting maritime emissions and scheduling berths is required. By combining state-of-the-art data analytics and optimization techniques, this system aims to increase operational efficiency while lowering the environmental impact of maritime logistics.

The economic and environmental advantages of this research make it significant. In addition to lowering congestion and increasing port efficiency, optimized berth scheduling also lowers operating expenses and fuel consumption. Furthermore, precise emissions forecasting makes it possible to adhere to environmental standards, which lessens the negative climate effects of shipping operations.

Therefore, this research focuses on emission forecasting and berth scheduling optimization to improve maritime operations. A Random Forest-based machine learning model is developed to predict shipping emissions, helping ports make proactive decisions for sustainability. Additionally, a Mixed Integer Linear Programming (MILP) model is designed to optimize berth allocation, ensuring efficient port operations while minimizing environmental impact.

1.2 Environmental Challenges in Maritime Logistics:

1. **High Emissions During Ship Berthing and Transit**
    - Ships emit large amounts of greenhouse gases (CO_2, NOx, SOx) while waiting for berths and during manoeuvring in port areas.
    - Inefficient scheduling increases idle time, directly raising fuel consumption and emission levels.
2. **Inefficient Berth Scheduling Leading to Fuel Wastage**
    - Poor coordination in berth allocation results in longer waiting times.
    - Ships burn fuel while idling, significantly contributing to air pollution near ports.
3. **Difficulty in Predicting Emissions Accurately**
    - Real-time variables like weather, sea conditions, and ship load factor affect emission rates, making accurate prediction a challenge.
    - Lack of real-time data integration can reduce the reliability of emission forecasts.
4. **Compliance with Environmental Regulations**
    - Ports and shipping companies need to comply with strict international emission regulations (IMO standards).
    - Failure to manage emissions effectively can lead to penalties and reputational damage.
5. **Limited Infrastructure for Real-Time Monitoring**
    - Many ports lack integrated systems that combine berth scheduling with emission monitoring.
    - Absence of centralized data hinders proactive decision-making.
6. **Balancing Operational Efficiency with Environmental Goals**
    - Optimizing berth schedules for efficiency alone may neglect emission considerations.
    - The challenge lies in finding a balance between minimizing fuel consumption and meeting shipping deadlines.

7. Dynamic Port Conditions and Shipping Demands

- Sudden changes in ship arrivals, port congestion, or environmental factors can disrupt scheduling and increase emissions.
- Adaptive, intelligent systems are required to respond effectively.

1.3 Objective of the Work:

The object of this project is to create a system that improves maritime logistics by accurately predicting ship emissions and optimizing berth scheduling. It aims to use machine learning to forecast emissions based on real-time data, such as ship movements and fuel use. At the same time, it will apply optimization algorithms to reduce ship waiting times and fuel consumption, making port operations more efficient and environmentally friendly.

1.4 Motivation of the Work:

This research is motivated by the need to develop a smarter, more sustainable solution that improves berth scheduling to minimize delays and fuel use while accurately predicting emissions. By combining machine learning for emissions forecasting with optimization algorithms for berth scheduling, this work aims to enhance both operational efficiency and environmental sustainability in maritime logistics.

1.5 Problem Statement:

The maritime industry faces challenges with high carbon emissions and inefficient port operations. Ships often experience long waiting times at ports, leading to increased fuel consumption and pollution. This research aims to solve these issues by developing a system that predicts emissions using machine learning and improves berth scheduling with optimization techniques, reducing delays, fuel use, and emissions.

1.6 Organization of the Thesis:

Chapter 1: Introduction, Covers Environmental Challenges in Maritime Logistics, motivation, problem Statement.

Chapter 2: Literature Survey, Reviews existing research and technologies related to berth scheduling and emissions prediction, Limitations over it.

Chapter 3: Software Requirements discusses about the functional and non- functional requirements of the system.

Chapter 4: System Architecture, describes the overall system design, including key components and interactions.

Chapter 5: System Design & Analysis discusses about UML Diagrams, Provides visual representations of the system through various UML diagrams.

Chapter 6: Implementation & Coding includes the system predicts ship emissions and optimizes berth scheduling using a machine learning model. Results are displayed through a web application for efficient port management and also provide sample code.

Chapter 7: Results & Discussions contains Input and Output, demonstrates how the system processes data and presents results and also provides comparative analysis.

Chapter 8: Conclusion, Summarizes findings, contributions, and potential future improvements.

CHAPTER 2
LITERATURE SURVEY

2.LITERATURE SURVEY

2.1 Introduction

The maritime industry is a crucial component of global trade and transportation. However, it is also a significant contributor to greenhouse gas emissions and air pollution. With increasing pressure from international bodies like the International Maritime Organization (IMO), ports worldwide are seeking sustainable solutions to to Machine learning models, such as Random Forest, have shown great potential in predicting emissions and optimizing port operations. Integrating predictive models with berth scheduling systems can help port authorities make data-driven decisions that balance efficiency with environmental responsibility.

This survey explores existing research in the areas of:

- Emission prediction in maritime logistics.
- Intelligent berth scheduling.
- Machine learning applications in port management.

2.2 Survey of Existing Research

1) Jiang et al. (2024) proposed a Mult objective optimization model for the integrated scheduling of restricted channels, berths, and yards in bulk cargo ports. Recognizing that these port components are interdependent, the study introduced a joint scheduling framework that factors in carbon emissions, tugboat use, berth service types, and yard space allocation. To solve this complex problem, they developed the NSGA-II-DPGR algorithm by enhancing the classic NSGA-II with dual-population and greedy heuristics. The proposed model outperformed traditional First-Come, First-Serve (FCFS) strategies by improving scheduling efficiency, lowering emissions, and better utilizing port resources. Their work showed strong potential for adoption in sea–land integrated scheduling under low-carbon constraints.

2) Guo et al. (2021) presented a multi-period coordinated optimization model for berth allocation and yard assignment in container terminals. Recognizing that vessel berthing positions and export container stacking locations significantly influence truck travel distance, the authors formulated an integer programming model aimed at minimizing this distance. They developed a Tabu Genetic Algorithm (TGAIBY) to find optimal berth and yard plans. Their model considered interactions across different time periods and evaluated three scenarios—determinate berth allocation, determinate yard assignment, and coordinated optimization. Numerical experiments demonstrated that the coordinated approach reduced truck travel distances by 2–5% compared to berth-only optimization and by 15% compared to yard-only optimization, thereby enhancing terminal efficiency and sustainability.
3) Li et al. (2024) explored berth and yard scheduling optimization in container ports featuring diagonal yard layouts, such as flying-V and fishbone designs. The study developed a Mixed-Integer Linear Programming (MILP) model aimed at minimizing berthing delays and internal transportation travel distance. A neighborhood search algorithm was implemented to solve the problem, and numerical experiments demonstrated that diagonal layouts can enhance terminal performance. The paper provided a novel framework for evaluating how geometric yard design impacts scheduling efficiency. Sensitivity analyses showed how the angle of cross lanes and the yard's width-to-depth ratio affect total travel distance, supporting the use of alternative yard layouts in modern ports.
4) Lin and Wang (2025) proposed the Voting-BRL model, an ensemble machine learning method that combines Bayesian Ridge Regression and Lasso Regression to predict ship CO_2 emissions and fuel consumption. The model leverages Analysis of Variance (ANOVA) for feature selection and integrates the predictions of its sub-models through a voting mechanism. Using four years of real-world maritime data from the THETIS-MRV platform, the model achieved a high predictive performance with an R^2 of 0.9981 and RMSE of 8.53, outperforming traditional models such as XGB Regressor. The Voting-BRL model was especially effective in handling high-dimensional and noisy maritime datasets. Its ability to balance

accuracy, generalizability, and interpretability makes it a powerful tool for emission management and sustainable port operations.

5) Ma et al. (2019) proposed an integrated model for berth allocation and yard planning in container terminals with discontinuous berth layouts. Recognizing the limitations of traditional discrete and continuous layout models, the authors introduced a multi-continuous berth layout approach, supported by a Mixed-Integer Linear Programming (MILP) model. To improve computational efficiency, they developed a Guided Neighborhood Search (GNS) algorithm, which significantly reduced the runtime. The model aimed to maximize berth space utilization, reduce vessel waiting times, and minimize the container travel distance between berth and yard. Experimental results showed the integrated model outperformed existing methods in terms of both efficiency and practicality, making it suitable for real-world port operations with irregular berth configurations.

2.3 Limitations Surveyed Papers

1) While Jiang et al. (2024) provided a robust model for integrated scheduling and demonstrated improvements over existing FCFS strategies, the implementation remains theoretical. The model's real-world applicability depends on the accuracy and availability of detailed, real-time data across all three components—channels, berths, and yards. Additionally, the NSGA-II-DPGR algorithm, though efficient, may face scalability challenges in extremely large ports with high traffic variability and operational uncertainties. Furthermore, the feasibility of integrating such models with legacy port management systems was not explored.

2) While Guo et al. (2021) effectively demonstrated the advantages of coordinated berth and yard planning, their study focused solely on the export container loading process, omitting the unloading process, which could significantly impact yard congestion and truck routing. The model's effectiveness may also decrease in real-world terminals with dynamic operational factors like unplanned vessel delays or equipment breakdowns. Additionally, although the Tabu Genetic Algorithm showed good performance in simulations, its computational complexity and scalability were not

tested in ultra-large terminals or under high-traffic conditions, which may affect its practical deployment.

3) Although Li et al. (2024) introduced an innovative approach to berth and yard scheduling in diagonal layouts, their study was limited to theoretical models and simulation-based validation. The implementation of these yard designs in real-world ports may face practical challenges such as equipment compatibility, reconfiguration costs, and operational training. Additionally, the proposed MILP model's computational scalability for very large terminals was not addressed. The influence of external disruptions (e.g., weather delays, vessel congestion) was also not integrated, potentially limiting the model's robustness under dynamic port conditions.

4) While the Voting-BRL model by Lin and Wang (2025) shows exceptional predictive accuracy, its reliance on high-quality, structured data from sources like THETIS-MRV may limit its generalizability in regions or ports where such data is unavailable or incomplete. Additionally, the ensemble structure—though powerful—adds computational overhead, which could be a barrier for real-time applications in resource-constrained settings. Although more interpretable than deep learning models, the combination of Bayesian and Lasso regressions may still pose challenges in terms of clear operational explainability for non-technical stakeholders. Lastly, the model primarily focuses on CO_2 emissions and does not address other pollutants or multimodal shipping scenarios.

5) While Ma et al. (2019) addressed the problem of discontinuous berth layouts effectively, their model's applicability may be constrained in terminals with high operational variability or real-time uncertainty (e.g., sudden delays, equipment failure). The assumption of pre-known vessel arrival times and container demands might limit the model's flexibility in dynamic environments. Additionally, although the GNS algorithm enhances optimization speed, its performance for large-scale terminals or highly congested ports with many operational constraints was not extensively tested. Finally, environmental considerations such as emissions and energy use were not explicitly integrated into the model's objectives.

CHAPTER 3
SOFTWARE REQUIREMENTS

3. SOFTWARE REQURIREMENTS

3.1 System requirements:

This section outlines the Software Requirement Specification (SRS) essential for the successful development and implementation of the system. The analysis phase plays a crucial role in understanding both user expectations and technical demands, ensuring the solution is efficient, scalable, and reliable. The SRS defines all functional and non-functional requirements, serving as a roadmap for the development team to design a system that integrates smoothly with existing infrastructure, meets performance standards, and supports future enhancements.

3.2 Functional Requirements:

Ship Data Management: The system must collect and store real-time data from ships, including arrival time, size, fuel consumption, and cargo details.

Emission Prediction: The system must predict ship emissions using a Random Forest model based on real-time and historical data.

Berth Scheduling: The system must use MILP to create an optimized berth schedule, minimizing waiting time and fuel consumption.

Real-Time Updates: The system should update berth schedules in real-time when new ship data is received or conditions change.

User Interface: The system should provide an interactive web interface for port authorities to monitor ship status, manage schedules, and view emission forecasts.

Data Storage: The system must store ship data, berth schedules, and emission predictions in a centralized database.

3.3 Non- Functional Requirements:

Security: Ship data, berth schedules, and emission records should be encrypted to prevent unauthorized access.

Scalability: The system should handle a large number of ships and berth requests efficiently without performance degradation.

Reliability: The system should provide accurate berth scheduling and emission predictions with minimal errors or failures.

Usability: The web application should have a simple, user-friendly interface for port authorities to manage schedules and monitor emissions easily.

Performance: The system should process ship data and update berth schedules and emission predictions in real-time.

Availability: The system should be available 24/7 to handle ship arrivals and schedule changes without downtime.

3.4 Software Requirements

To develop and deploy the system, the following software components are required:

1. Frontend:
    - Technology: HTML, CSS
2. Backend:
    - Technology: Flask
    - Database: Sqlite for structured data storage
3. Machine Learning Module:
    - Programming Language: Python
    - Libraries: Scikit-learn, Pandas for ML-based emissions prediction

3.5 Hardware Requirements

1. **Processor (CPU)**
    - Minimum: Intel Core i5 (8th Gen or above) / AMD Ryzen 5
    - Recommended: Intel Core i7 / AMD Ryzen 7 for faster data processing and real-time operations
    - Justification: Required for efficient handling of machine learning tasks and optimization algorithms like MILP.

2. **RAM (Memory)**
    - Minimum: 8 GB
    - Recommended: 16 GB or higher

- Justification: To manage large datasets, especially historical and real-time ship data processing.

3. **Storage**

- Minimum: 256 GB SSD or HDD
- Recommended: 512 GB SSD or higher
- Justification: For storing the dataset (1,000,000 entries), machine learning models, system files, and application data.

4. **Graphics Processing Unit (GPU)**

- Optional but recommended: NVIDIA GTX 1650 or higher
- Justification: To accelerate training of the Random Forest model and improve efficiency in real-time data visualization (if using advanced models or future deep learning extensions).

5. **Network Connectivity**

- Stable internet connection
- Minimum: 10 Mbps download and upload
- Justification: Required for real-time data flow between the web application, database, and port systems.

6. **Display and Peripherals**

- Full HD Monitor (1080p)
- Keyboard and Mouse for development and operational control
- Justification: For better visualization of real-time monitoring dashboards and manual intervention by port authorities.

CHAPTER 4
SYSTEM ARCHITECTURE

4.SYSTEM ARCHITECTURE

4.1 Architecture of The Proposed System:

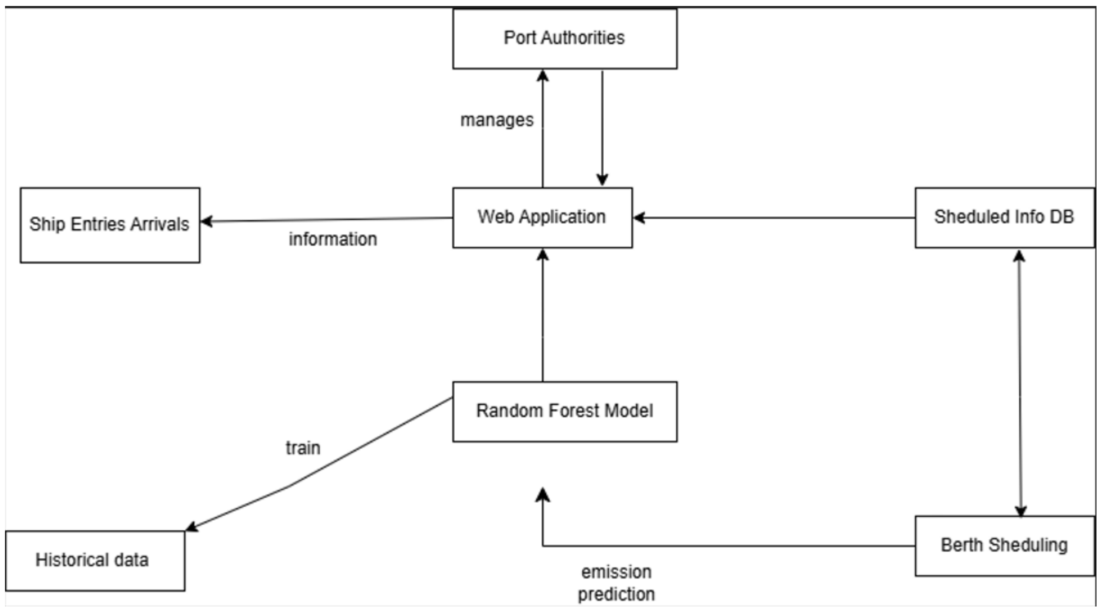


Figure 4.1 Architecture Diagram of Port Emission Prediction and Berth Scheduling System

The proposed architecture integrates machine learning for emission forecasting with MILP-based berth scheduling through a centralized web application. As shown in Figure 1, our system collects real-time ship data, such as vessel type, size, fuel usage, arrival time, and cargo details. A web application that updates and retrieves berth schedules from a database process this data. We employ the Random Forest model, which helps port authorities make better decisions by analyzing historical and current data.

Simultaneously, berth scheduling is optimized through the use of Mixed Integer Linear Programming (MILP), which lowers fuel consumption and ship waiting times . Lastly, the web application allows for real-time schedule monitoring and modification, improving the efficiency of port operations. The key components as shown in Figure 4.1 are:

Ship Arrivals:

Ships offer up-to-date information on cargo details, estimated arrival time, fuel consumption, emission levels, vessel type, size, and weight. The web application processes this data to maintain the emissions models and berth scheduling current. Quick reactions to modifications in ship traffic and port conditions are made possible by the system's assurance of a smooth and accurate information flow. The system can improve berth assignments and forecast emissions by incorporating this real-time data, which will ultimately increase port efficiency and lessen its environmental impact.

Web Application:

By processing real-time ship data and enabling communication between the berth scheduling engine, the emission prediction model, and the ship arrival system, the web application acts as the system's main control center. It serves as the port authorities' interface, enabling them to keep an eye on and modify schedules as necessary. Port operators can react swiftly to changes in ship arrivals and environmental conditions thanks to the web application's processing of real-time data and generation of actionable insights.

Scheduled Data (DB):

In order to maintain consistency between ship arrival times and berth availability, the database keeps track of ship data and berth schedules. Berth allocations are continuously optimized thanks to its support for dynamic updates based on current ship status and port traffic. Efficient information retrieval and storage are made possible by the database, which offers a solid basis for scheduling and emission prediction. This guarantees that port operations continue to be effective and flexible in the face of unforeseen delays or changes in ship arrivals.

Random Forest Model (Emission Prediction):

The system uses a Random Forest model to predict emissions based on historical data and real-time ship movement information. The model considers vessel size, speed, fuel type, and environmental factors to generate accurate emission forecasts. The output from this model is displayed through the web application, allowing port authorities to monitor and manage emissions more effectively. This predictive capability ensures compliance with environmental regulations and helps reduce the ecological impact of port operations.

Berth Scheduling (MILP):

Mixed Integer Linear Programming (MILP) is used to optimize berth scheduling in order to minimize ship waiting times, improve berth utilization, and lower emissions and fuel consumption. The model enhances efficiency and sustainability by dynamically modifying schedules in response to port traffic and real-time ship arrivals.

Port Authorities:

The web application is used by port authorities to control emissions and ship schedules. With the use of real-time data and forecasts, they can modify ship priorities and berth assignments. This aids in their decision-making to increase efficiency and lower fuel and emission levels.

4.2 Conclusion:

The analysis phase plays a crucial role in defining the system's structure, functionality, and technical requirements. By understanding user needs and specifying software requirements, we ensure that the proposed system aligns with operational goals and industry standards.

The software requirements outline the technologies needed for frontend development, backend processing, database management, machine learning, and system deployment.

With a well-defined analysis, the system is designed to enhance efficiency, reduce ship waiting times, and support sustainable maritime operations. This structured approach ensures a smooth development process, minimizing challenges during implementation while maximizing the system's impact on port management.

CHAPTER 5
SYSTEM DESIGN & ANALYSIS

5. SYSTEM DESIGN & ANALYSIS

5.1 Introduction:

The design phase is crucial in transforming system requirements into a structured blueprint for development. It defines the architecture, user interface, data flow, and interactions between different components of the system. A well-structured design ensures scalability, maintainability, and efficiency in the implementation phase.

The proposed system for berth scheduling and emissions prediction is designed using a modular approach, ensuring each component functions independently while integrating seamlessly. The design process focuses on:

System Architecture: Defines the overall structure, including frontend, backend, and database layers.

User Interface Design: Ensures an intuitive and user-friendly experience for port authorities and logistics managers.

Data Flow and Processing: Outlines how input data is processed, stored, and analysed for decision-making.

Algorithm Design: Describes the implementation of machine learning models for emissions prediction and AI-based berth scheduling.

This section provides a detailed breakdown of the system's architecture, user interface design, and functional components, ensuring an optimized, user-friendly, and efficient port management solution.

5.2 UML Design:

In this project, UML diagrams are used to represent the berth scheduling and emissions prediction system, ensuring clarity in system design and implementation. The key UML diagrams used are:

5.2.1 Use case Diagram:

Represents the interactions between users and the system, showing different functionalities and user roles. It helps in understanding system requirements and how users engage with various features.

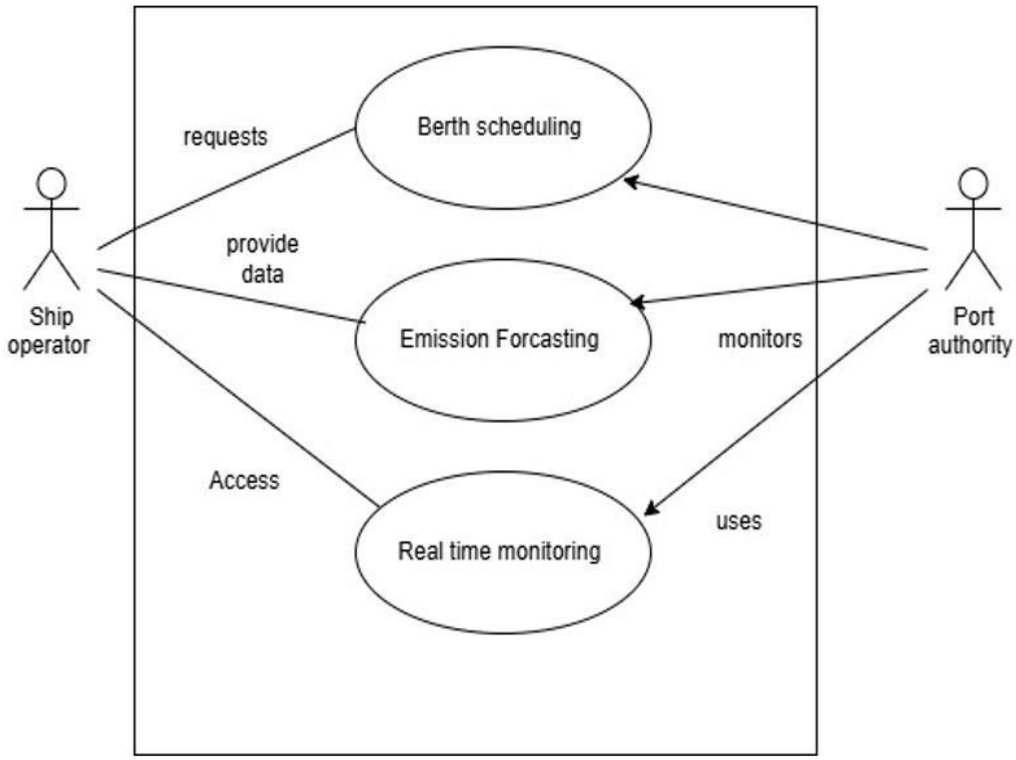


Figure 5.1 Use Case Diagram of Port Emission Forecasting and Berth Scheduling System

Actors:

1. **Ship Operator**
   - Provides data required for emission forecasting.
   - Requests berth scheduling for docking.
   - Accesses real-time monitoring for tracking information.
2. **Port Authority**
   - Monitors emission forecasting to ensure environmental regulations are met.
   - Uses real-time monitoring for tracking ships and port operations.
   - Receives berth scheduling requests to manage ship docking.

Responsibilities:

1. **Berth Scheduling**
   - The ship operator requests berth scheduling to secure docking space.
   - The port authority handles and approves berth scheduling requests.
2. **Emission Forecasting**
   - The ship operator provides data related to emissions.
   - The port authority monitors emissions to ensure compliance with regulations.
3. **Real-Time Monitoring**
   - The ship operator accesses real-time data for tracking ship status.
   - The port authority uses real-time monitoring to oversee ship activities and ensure smooth port operations.

Process Flow

- The ship operator interacts with the system to manage berth scheduling, emission forecasting, and real-time monitoring.
- The port authority oversees and monitors these processes for efficient port management.
- The system ensures smooth port operations while keeping emissions in check.

5.2.2 Class Diagram:

Class Diagram represents the static structure of a system by showing its classes, attributes, methods, and relationships. It defines how different entities (classes) interact with each other. This diagram is widely used in software development for system modeling and database design

.

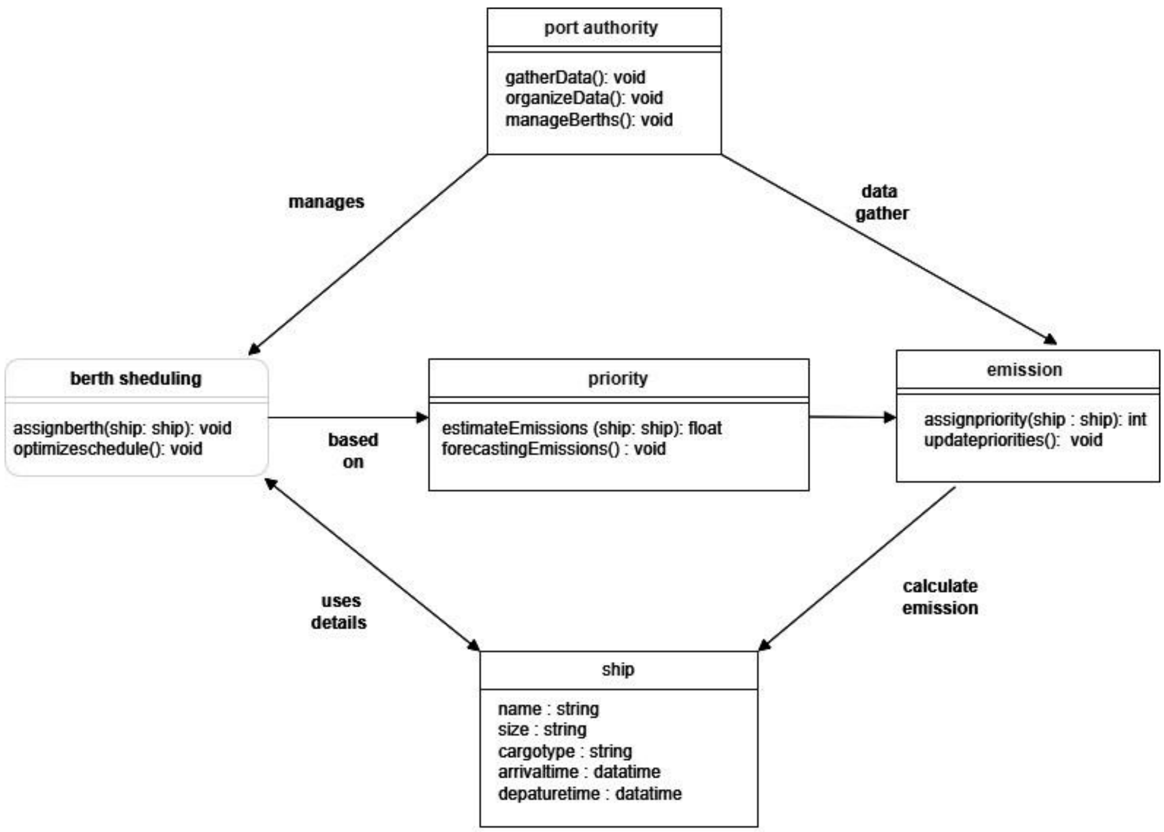


Figure 5.2 UML Class Diagram for Port Berth Scheduling and Emission Management System

Classes and Their Relationships

Port Authority

- gatherData: Collects relevant data from ships
- organizeData: Organizes collected information
- manageBerths: Handles berth allocation for ships

Ship

- name: Name of the ship
- size: Size of the ship
- cargotype: Type of cargo it carries
- arrivaltime: Arrival time at the port
- departuretime: Departure time from the port

Berth Scheduling

- assignberth: Assigns a berth to a ship
- optimizeschedule: Optimizes the scheduling of ships

Emission

- assignpriority: Assigns a priority based on emissions
- updatepriorities: Updates priority levels based on emissions

Priority

- estimateEmissions: Estimates emissions for a ship
- forecastingEmissions: Predicts future emissions

Process Flow:

- Port Authority collects data and manages berth assignments
- Ships provide details like arrival, departure, cargo, and size
- Berth Scheduling assigns berths based on priority and optimizes schedules
- Emission Calculation determines emission levels and assigns ship priority
- Priority class helps in forecasting emissions and influencing berth scheduling

5.2.3 Activity Diagram:

An Activity Diagram is a flowchart that represents the step-by-step execution of a process. It includes actions, decisions, and parallel flows, showing how tasks progress from start to finish. This diagram helps in understanding system behaviour and identifying process optimizations.

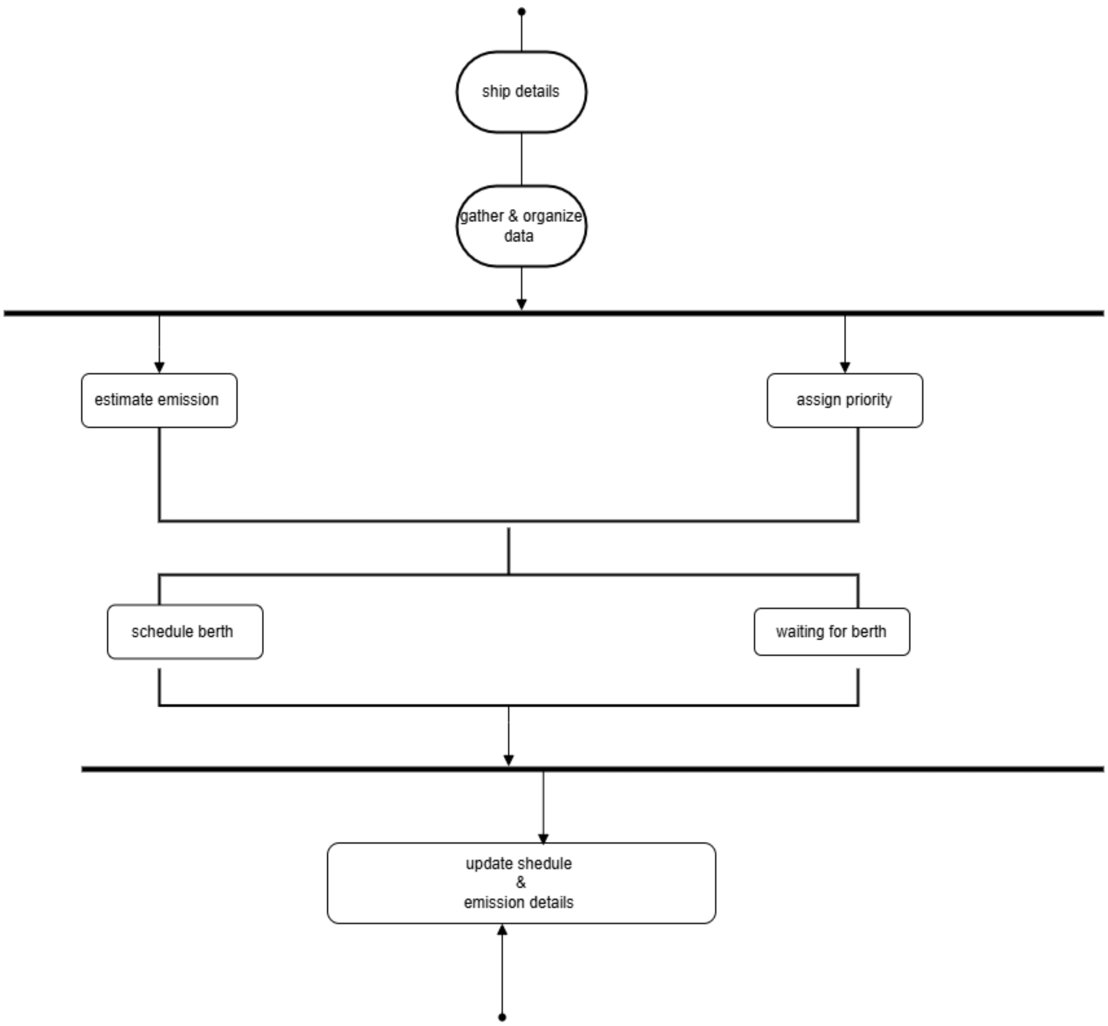


Figure 5.3 UML Activity Diagram for Port Berth Scheduling and Emission Management System

This activity diagram represents the workflow for managing ship arrivals, estimating emissions, and scheduling berths efficiently. It illustrates the sequence of activities, decision points, and interactions between different processes in the system.

1. **Ship Details**
   - The process begins with collecting ship details, such as name, size, cargo type, arrival time, and departure time.
   - This information is crucial for making informed decisions regarding berth scheduling and emission forecasting.
2. **Gather and Organize Data**
   - After receiving the ship details, the system processes and organizes them.
   - This step ensures that relevant data is structured for further analysis.
3. **Estimating Emissions**
   - The first major branch of the workflow focuses on emission calculations.
   - Using ship-related data, fuel consumption patterns, and environmental conditions, an estimation of emissions is calculated.
   - This step helps in monitoring and controlling pollution levels.
4. **Assigning Priority**
   - The second branch of the process assigns a priority level to the ship based on factors such as arrival time, cargo type, and environmental impact.
   - Higher priority may be given to ships with perishable goods, urgent cargo, or lower emissions.
5. **Scheduling a Berth**
   - Once emissions are estimated and priority is assigned, the ship is scheduled for a berth if available.
   - The berth scheduling is optimized to reduce waiting time and fuel consumption.
6. **Waiting for a Berth**
   - If no berth is available, the ship waits until a suitable berth is free.
   - The system continuously monitors berth availability and assigns one as soon as possible.

7. **Updating Schedule and Emission Details**
    - The final step updates the schedule to reflect the ship's docking time and emission details.
    - These updates help in maintaining accurate records for future optimization and decision-making.

5.2.4 Component Diagram:

A component diagram is a type of UML diagram that illustrates the components of a system and their relationships. It shows how different components interact and communicate with each other, focusing on system architecture. This diagram helps to visualize the high-level structure and organization of the system.

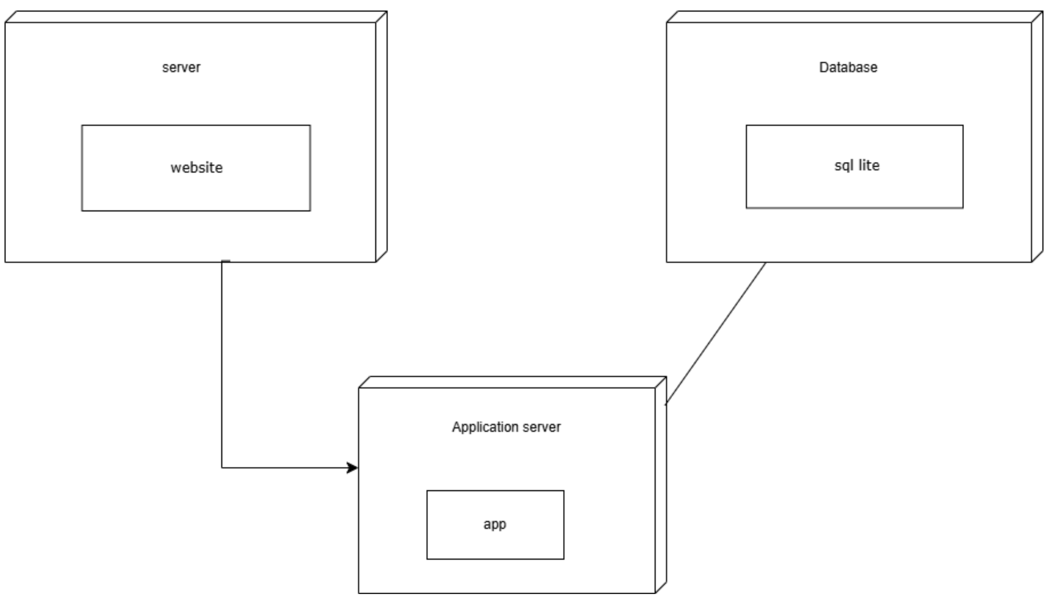


Figure 5.4 UML Component Diagram for Port Berth Scheduling and Emission Management System

This component diagram illustrates the structure of a system consisting of a website, an application server, and a database. It highlights how different components interact to support system functionality.

1. **Server**

   Hosts the website that serves as the user interface for accessing the system.

   Handles incoming client requests and forwards them to the appropriate components for processing.

2. **Website**

   The front-end interface where users interact with the system.

   Sends requests to the application server for processing and retrieves data from the database when needed.

3. **Application Server**

   Acts as an intermediary between the website and the database.

   Processes requests, runs business logic, and manages communication between the website and the database.

4. **App**

   A backend application running on the application server that handles system operations.

   Ensures data is retrieved and processed correctly before sending responses to the website.

5. **Database**

   Stores and manages system data using SQLite.

   Provides data retrieval and storage services to the application server for efficient processing.

5.2.5 Sequence Diagram:

A sequence diagram is a type of UML diagram that models the interactions between objects in a system over time. It shows the sequence of messages exchanged between objects or components in a particular scenario. The diagram visually represents the flow of control and data between the system's parts, helping to understand and analyze the behavior of a system's interactions.

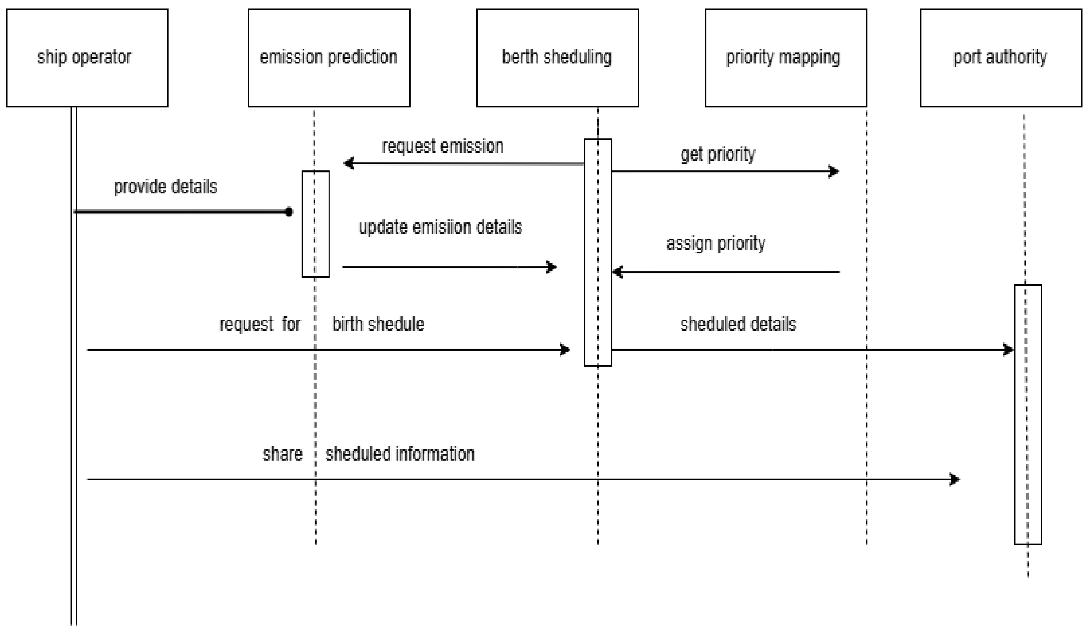


Figure 5.5 UML Sequence Diagram for Port Berth Scheduling and Emission Management System

1. **Ship Operator**
    - Provides ship details to the system for processing.
    - Initiates the request for berth scheduling.
    - Receives scheduled berth information.
2. **Emission Prediction**
    - Receives ship details from the ship operator.
    - Requests emission data based on provided information.
    - Updates the emission details for further processing.
3. **Berth Scheduling**
    - Processes the berth scheduling request from the ship operator.
    - Requests priority mapping to determine berth allocation.
    - Updates the schedule based on priority mapping.
4. **Priority Mapping**
    - Receives priority request from berth scheduling.
    - Assigns priority based on predefined criteria.
    - Sends scheduled details back to berth scheduling.
5. **Port Authority**
    - Receives scheduled berth details from the berth scheduling system.
    - Maintains overall port operations based on the provided schedule.

Process Flow

- The ship operator provides details to initiate the process.
- The emission prediction module estimates emissions and updates details.
- The berth scheduling system requests priority from priority mapping.
- The priority mapping module assigns priority and sends scheduled details back.
- The berth scheduling system shares the final schedule with the port authority.

CHAPTER 6
IMPLEMENTATION & CODING

6. IMPLEMANTATION AND CODING

6.1 Ship Details Form:

The system features a comprehensive web-based form designed for port authorities or ship operators to input crucial ship-related data. This information is essential for real-time berth scheduling, emission forecasting, and historical data analysis.

6.1.1 Purpose:

The Ship Details Form ensures accurate data capture that feeds into backend algorithms responsible for optimizing berth allocation and predicting emissions based on operational parameters.

6.1.2 Form Description:

The form is embedded within a responsive HTML interface and consists of multiple input fields grouped by categories. It allows users to submit ship-related information through a user-friendly interface. Upon submission, the data is stored in a structured SQL database (ship_infoo table), where it is further utilized by scheduling algorithms and machine learning models.

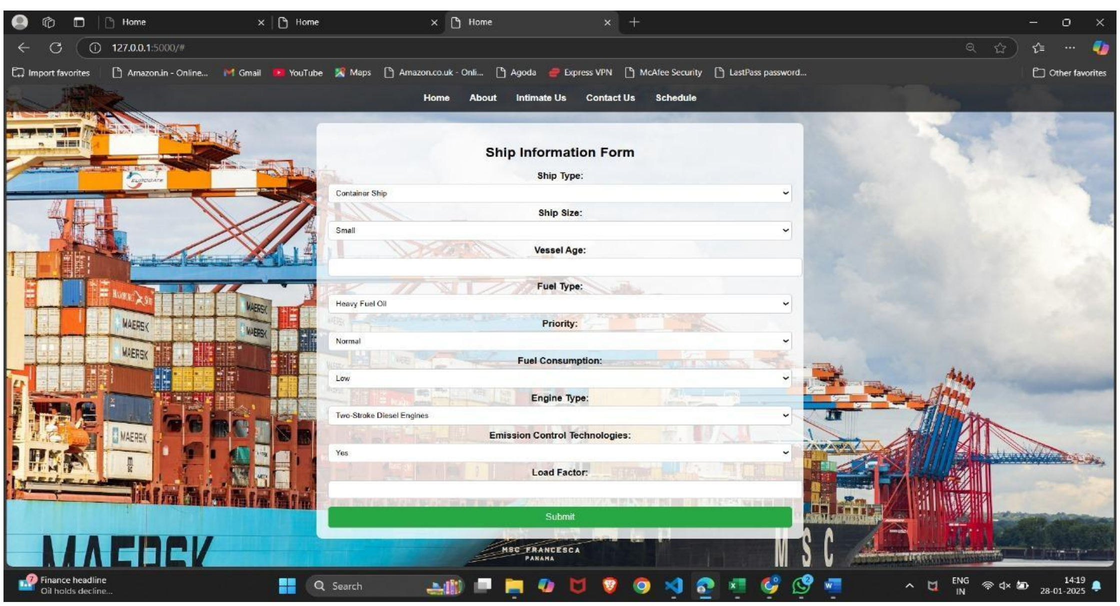


Figure 6.1 Ship Information Form in the Port Management System

6.1.3 Key Input Fields:

As shown in Figure 6.1, the respective fields include the following:

1. Ship Type: Identifies the category of the vessel.

2. Ship Size: Indicates the size class of the ship.

3. Vessel Age: Records the age of the vessel in years.

4. Fuel Type: Specifies the type of fuel used by the ship.

5. Priority: Determines the urgency or priority status of the ship.

6. Fuel Consumption: Represents emission indicator level.

7. Engine Type: Indicates the configuration of the ship's engine.

8. Emission Control Technologies: Flags whether the ship uses emission-reducing technologies.

9. Load Factor: Numeric input indicating how much load the ship is carrying.

6.2 Emission Forecasting:

We developed a machine learning-based predictive model to estimate ship emissions using Random Forest, analyzing both historical and real-time data for accurate predictions.

6.2.1 Data Preparation:

To ensure data quality and improve model accuracy, several preprocessing steps were applied to the dataset, which consisted of 100,000 records. The numerical features—such as vessel age, fuel consumption, and load factor—were normalized, bringing all feature values to a consistent scale. This step is essential for preventing large-scale features from dominating the model’s learning process. Additionally, outliers were detected and removed to avoid skewing the model’s performance and introducing bias. Missing values were handled carefully: depending on their frequency and relevance, some were imputed using statistical techniques, while others were removed if deemed too sparse or non-impactful.

A major challenge in the dataset was class imbalance—particularly, the underrepresentation of the "High" emission class. To overcome this, the Synthetic Minority Over-sampling Technique (SMOTE) was employed. SMOTE generates artificial data points for minority classes by interpolating between existing examples, allowing the model to learn more equitable decision boundaries. This technique enhanced the model's ability to generalize across all emission categories and significantly improved both fairness and predictive power.

6.2.2 Model Training:

The model was trained using the Random Forest Classifier, a robust ensemble learning method known for its ability to handle complex datasets and reduce overfitting. The dataset, which contained 100,000 records, was split into 80% for training and 20% for validation to assess the model's ability to generalize to unseen data. The input features included key ship parameters such as ship size, fuel type, vessel age, fuel consumption level, engine type, emission control technology, and load factor. These features were carefully selected based on domain expertise and their significant influence on ship emissions.

The model training process involved several phases: data preparation, feature selection, resampling with SMOTE to address class imbalance, model training, prediction, and performance evaluation. During training, the Random Forest algorithm builds multiple decision trees, each trained on a bootstrapped subset of the training data. At every split in a decision tree, a random subset of features is selected to promote diversity among trees and prevent overfitting. This randomness ensures that the model captures a wide range of patterns and relationships in the data.

During the prediction phase, each tree in the forest independently predicts the emission category for a given input, and the final output is determined by majority voting. This ensemble strategy enhances the model's accuracy, robustness, and stability.

The model was fine-tuned using the following hyperparameters:

- n_estimators = 50: Number of trees in the forest
- max_depth = 10: Maximum depth of each decision tree
- min_samples_split = 7: Minimum number of samples required to split a node
- min_samples_leaf = 1: Minimum number of samples required at a leaf node
- class_weight = 'balanced': Automatically adjusts weights to handle class imbalance
- random_state = 42: Ensures reproducibility

These hyperparameters were selected to optimize performance while ensuring efficient training. As a result, the Random Forest model achieved high predictive accuracy and maintained balanced classification performance across emission levels, successfully supporting sustainable decision-making in maritime operations.

6.2.3 Model Integration with Application Logic:

In the backend application, user inputs are collected through a web form and mapped to the same structure used during model training. Features like fuel type and engine type are converted to numerical values consistent with the model's expectations. Input data is carefully aligned with the training feature set to ensure compatibility.

Once the input is processed, the model is loaded, and a prediction is generated. This emission prediction is stored in the database and used to support decision-making in the berth allocation process. Ships predicted to have higher emissions may be deprioritized to reduce environmental impact, while cleaner ships are given preference.

6.2.4 Evaluation Metrics:

To evaluate the performance of the model, several metrics were used to assess its ability to predict ship emissions and optimize berth schedules. The evaluation highlights the model's precision, accuracy, and overall effectiveness in making predictions.

- Mean Average Precision (mAP):
  - mAP@50: Measures the model's accuracy at an Intersection over Union (IoU) threshold of 0.5, ensuring accurate predictions even with moderate overlap.
  - mAP@50-95: Assesses the model's precision across a range of IoU thresholds (from 0.50 to 0.95), providing a more comprehensive evaluation of its prediction capabilities.
- Intersection over Union (IoU): IoU measures the overlap between predicted and actual outcomes (such as emissions and berth assignments). The model achieved an IoU threshold of 0.7, ensuring strong alignment between predicted and actual outcomes.
- Accuracy: The model achieved an overall accuracy of 98%, showing it can reliably predict ship emissions and berth schedules. This high accuracy suggests that the model minimizes false positives and false negatives effectively.
- Precision and Recall:
  - Precision: The model's precision scores were 1.00, 0.98, and 0.96 across different classes, indicating a very low rate of false positives.
  - Recall: With recall scores of 0.98, 0.99, and 0.99, the model successfully captures almost all true positive cases, ensuring very few positive cases are missed.
- F1-Score: The model achieved F1-scores of 0.99, 0.99, and 0.97 across different classes, providing a balanced measure of its precision and recall. These high scores indicate that the model maintains a strong balance between detecting true positives and avoiding false positives.
- Confusion Matrix Analysis: The confusion matrix reveals the model's performance in classifying ship emissions and berth assignments into different categories (High, Moderate, and Low). The majority of predictions align with the diagonal of the matrix, indicating high classification accuracy and minimal misclassification. The normalized confusion matrix provides insights into the model's class-wise accuracy, identifying areas that may need further optimization.

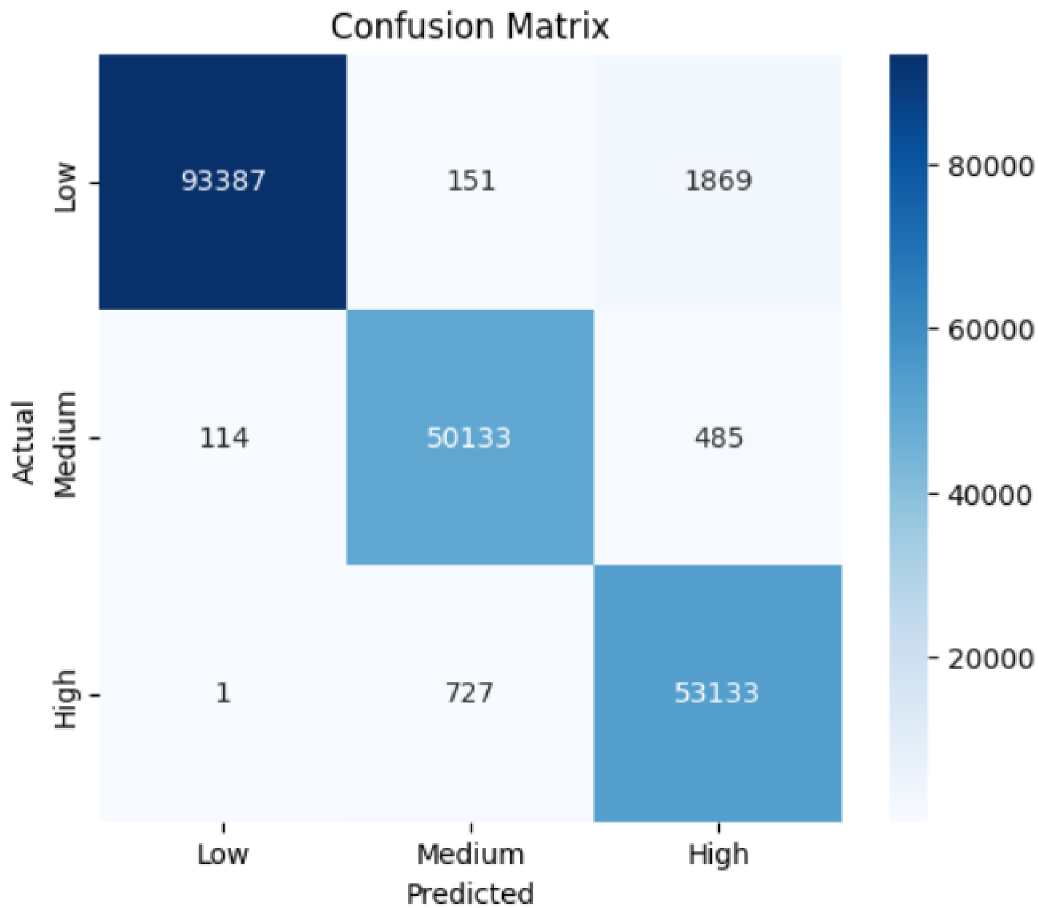


Figure 6.2: Confusion Matrix of the Random Forest Model

As shown in figure 6.2, The confusion matrix illustrates the model's accuracy in classifying ship emissions into High, Moderate, and Low categories. Most predictions align with the diagonal, reflecting high classification accuracy and minimal misclassification.

6.3 Berth Scheduling:

Efficient berth scheduling is critical for minimizing ship wait times, optimizing port operations, and reducing environmental emissions. In our system, we implemented a priority-based dynamic berth allocation mechanism that considers both environmental and operational factors.

6.3.1 System Overview:

The berth scheduling system operates with a limited number of berths and a waiting lobby. Each berth can accommodate one ship at a time, and berthing duration varies depending on the ship's size. The system continuously monitors berthing status and dynamically assigns berths to waiting ships based on a combination of ship priority, emission level, and arrival time.

6.3.2 Berthing Lobby vs. Waiting Lobby:

- **Berthing Lobby**: Ships currently occupying a berth.
- **Waiting Lobby**: Ships that have arrived but are waiting for an available berth.

Each incoming ship is evaluated for berth availability. If no berth is available, the ship is placed into the waiting lobby until a berth is freed.

6.3.3 Berthing Time Allocation:

Ships are assigned a berthing time based on their size:

- Small: 8 hours
- Medium: 16 hours
- Large: 24 hours

6.3.4 Berth Assignment:

When a ship arrives at the port, the system first checks for available berths. If a berth is available, the ship is immediately assigned to it based on current availability. If no berth is free at the time, the ship is placed into a waiting lobby. Once a ship's allocated berthing time comes to an end, it departs, and the berth is marked as available. As soon as a berth becomes vacant, the system automatically scans the waiting lobby, sorts the ships based on their priority level and arrival time, and assigns the berth to the highest-priority ship in the queue.

After a ship is assigned to a berth, the system updates the corresponding berthing information in the database. This includes the berth number, from time, and to time. These details are later retrieved and shown in the user interface, ensuring users have access to real-time berth assignment status.

6.4 Sample code:

App.py

```
from flask import Flask, request, render_template
import joblib
import pandas as pd
import sqlite3
from db import create_connection, create_table
from Milpp import ship_arrives
from datetime import datetime

app = Flask(__name__)

# Load trained model
rf_classifier = joblib.load('trained_model.pkl')

# Read dataset to understand expected features
data = pd.read_csv("emissions_dataset2.csv")
X = data.drop(columns=['emissions'])  # Remove target column
expected_columns = list(X.columns)  # Get the column names

@app.route('/submit', methods=['POST'])
def submit():
    conn = create_connection()
    if conn is not None:
        create_table(conn)
        cursor = conn.cursor()

        try:
            # Get form data
            ship_type = request.form['ship-type']
            ship_size = int(request.form['ship-size'])
```

```
        vessel_age = int(request.form['vessel-age'])
        fuel_type = request.form['fuel-type']
        fuel_consumption = request.form['fuel-consumption']
        engine_type = request.form['engine-type']
        emission_control = request.form['emission-control']
        load_factor = float(request.form['load-factor'])
        priority = request.form['priority']

        # Convert fuel consumption to numeric
        fuel_consumption_mapping = {
            "Low": 0,
            "Medium": 1,
            "High": 2
        }
        fuel_consumption = fuel_consumption_mapping.get(fuel_consumption, -1)
        if fuel_consumption == -1:
            return "Error: Invalid fuel consumption value", 400

        # Log received data
        app.logger.info(f"Received Form Data: Ship Type: {ship_type}, Ship Size:
{ship_size}, Vessel Age: {vessel_age}, Fuel Type: {fuel_type}, Fuel Consumption:
{fuel_consumption}, Engine Type: {engine_type}, Emission Control:
{emission_control}, Load Factor: {load_factor}")

        # Insert into database
        cursor.execute('''
            INSERT INTO ship_infoo (ship_type, ship_size, vessel_age, fuel_type,
fuel_consumption, engine_type,
                                emission_control_technologies, load_factor)
            VALUES (?, ?, ?, ?, ?, ?, ?, ?)
        ''', (ship_type, ship_size, vessel_age, fuel_type, fuel_consumption,
engine_type, emission_control, load_factor))
```

```
last_id = cursor.lastrowid
app.logger.info(f"Inserted Ship ID: {last_id}")

# Map fuel type to numeric
fuel_type_mapping = {
    "Heavy Fuel Oil": 0,
    "Diesel": 1,
    "LNG": 2,
    "Biofuel": 3
}
fuel_type_numeric = fuel_type_mapping.get(fuel_type, -1)
if fuel_type_numeric == -1:
    conn.rollback()
    return "Error: Unknown fuel type", 400

# Map Engine Type to Numeric Values
engine_type_mapping = {
    "Two-Stroke Diesel Engines": 0,
    "Four-Stroke Diesel Engines": 1,
    "Gas Turbine Engines": 2,
    "Electric Propulsion": 3
}
engine_type_numeric = engine_type_mapping.get(engine_type, -1)
if engine_type_numeric == -1:
    conn.rollback()
    return "Error: Invalid engine type", 400

# Convert Emission Control to Numeric
emission_control_mapping = {
    "Yes": 1,
```

```
    "No": 0
}
emission_control_numeric = emission_control_mapping.get(emission_control,
-1)
if emission_control_numeric == -1:
    conn.rollback()
    return "Error: Invalid emission control value", 400

# Prepare input data for prediction
input_values = {
    "Ship Size": ship_size,
    "Vessel Age": vessel_age,
    "Fuel Consumption": fuel_consumption,
    "Engine Type": engine_type_numeric,
    "Emission Control Tech": emission_control_numeric,
    "Load Factor": load_factor,
}

# Ensure input values match model's expected features
input_data = pd.DataFrame([input_values])
for col in expected_columns:
    if col not in input_data:
        input_data[col] = 0

# Reorder columns to match model training order
input_data = input_data[expected_columns]

# Make prediction
predicted_emissions = int(rf_classifier.predict(input_data)[0])

# Update database with prediction
cursor.execute('''
```

```
        UPDATE ship_infoo
        SET Emissions = ?
        WHERE id = ?
    ''', (predicted_emissions, last_id))

    conn.commit()
    app.logger.info(f"Prediction Updated for Ship ID: {last_id}")

    # Determine emission level and ship size
    emission_levels = {
        0: "Low",
        1: "Moderate",
        2: "High"
    }
    sizeOfShip_levels = {
        2: "Large",
        0: "Small",
        1: "Medium"
    }
    sizeOfShip = sizeOfShip_levels.get(ship_size)
    emission_level = emission_levels.get(predicted_emissions, "Unknown")

    # Call ship_arrives function
    ship_arrives(int(last_id), sizeOfShip, priority, emission_level)

    return render_template('check.html', status_message="Check your ship
status.", ship_id=last_id, emission_level=emission_level)

except Exception as e:
    conn.rollback()
    app.logger.error(f"Database Error: {e}")
```

```
            return "Error processing data", 500

        finally:
            conn.close()

    else:
        return 'Error: Unable to connect to the database'

@app.route('/check')
def check():
    ship_id = request.args.get('ship_id')
    emission_level=request.args.get('emission_level')
    # Connect to the database
    conn = create_connection()
    cursor = conn.cursor()

    # Retrieve the ship information from the database based on the ship_id
    cursor.execute('''
        SELECT Berth, from_time, to_time
        FROM ship_infoo
        WHERE id = ?
    ''', (ship_id,))
    ship_info = cursor.fetchone()  # Fetch one row

    # Close the database connection
    conn.close()

    # Check if ship_info is not None (i.e., ship_id exists in the database)
    if ship_info:
        berth, from_time, to_time = ship_info

        # app.logger.info(f"Ship {ship_id} has been assigned to Berth {berth} from
```

{from_time_str} to {to_time_str}.")

```
        if berth == 0:
            # If the ship is in the waiting lobby, display a message
            return render_template('check.html', status_message="Your ship is in the
waiting lobby.", ship_id=ship_id)
        else:
            # If the ship has been allocated a berth, form the message
            # Fix by adding:
            from_time_str = datetime.fromtimestamp(from_time).strftime('%Y-%m-%d
%H:%M:%S')
            to_time_str = datetime.fromtimestamp(to_time).strftime('%Y-%m-%d
%H:%M:%S')

            return render_template('result.html', berth=berth, from_time=from_time_str,
to_time=to_time_str, emission_level=emission_level)



    else:
        # If ship_id does not exist in the database, display an error message
        return render_template('check.html', status_message="Ship ID not found in the
database.",ship_id= ship_id)

@app.route('/result')
def result():
    ship_id = request.args.get('ship_id')
    emission_level = request.args.get('emission_level')

    app.logger.info(f"Received ship_id: {ship_id}, emission_level: {emission_level}")

    conn = create_connection()
```

```
    cursor = conn.cursor()

    cursor.execute('''
        SELECT Berth, from_time, to_time FROM ship_infoo WHERE id = ?
    ''', (ship_id,))
    ship_info = cursor.fetchone()
    conn.close()

    if ship_info:
        berth, from_time, to_time = ship_info

        from_time_str = datetime.fromtimestamp(float(from_time)).strftime('%Y-%m-
%d %H:%M:%S') if from_time else "N/A"
        to_time_str = datetime.fromtimestamp(float(to_time)).strftime('%Y-%m-%d
%H:%M:%S') if to_time else "N/A"

        print(f"Retrieved from DB -> Berth: {berth}, From: {from_time_str}, To:
{to_time_str}")  # Debugging

        return render_template('result.html',
                               berth = str(berth) if berth else "Your Ship is Assigned to Waiting
Area. We will notify you when it is available",
                               from_time=from_time_str,
                               to_time=to_time_str,
                               emission_level=emission_level)

    print("No berthing details found for this ship.")  # Debugging
    return render_template('result.html',
                           berth="Your Ship is Assigned to Waiting Area. We will notify you
when it is available",
                           from_time="N/A",
                           to_time="N/A",
```

```
                    emission_level=emission_level)

@app.route('/')
def home():
    return render_template('home.html')

@app.route('/about')
def about():
    return render_template('about.html')

@app.route('/intimate')
def intimate():
    return render_template('intimate.html')

@app.route('/contactus')
def contactus():
    return render_template('contactus.html')

@app.route('/form')
def form():
    return render_template('form.html')

if __name__ == '__main__':
    app.run(debug=True)
```

CHAPTER 7
RESULTS AND DISCUSSIONS

7.COMPARITIVE ANALYSIS

7.1 Comparative Analysis:

To evaluate the efficiency of the Random Forest model, we conducted experiments using a proprietary dataset collected from maritime operations. The dataset consists of 100,000 samples with 9 features, which was split into 80% for training and 20% for testing. We applied Random Forest, Logistic Regression, and Gradient Boosting to predict the emission levels and fuel usage patterns. The outcomes of this comparison are outlined in Table 7.1.

Table 7.1: Comparison of Machine Learning Models for Emission Forecasting.

Model	Precision	Recall	F1-Score	Accuracy	Inference Time-(ms)
Logistic Regression	69.2%	70.1%	69.2%	70.1%	25
Gradient Boosting	92.1%	91.8%	91.9%	91.7%	55
Random Forest	98.0%	98.1%	98.0%	98.0%	45

From the Table 7.1, Our approach utilizes the Random Forest model, which outperformed better than other logistic regression and gradient boosting, achieving 98% accuracy with a precision of 98% and a recall of 98.1%. Its fast inference time of 45 ms makes it suitable for real-time prediction of ship emissions and berth scheduling.

7.2 Visualization of Output Screens:

Case 1: Ship Assigned to Berth:

This output shows a ship that has been successfully assigned a berth.

Details:

- **Berth Number**: 3
- **From Time**: 2025-04-11 20:05:40
- **To Time**: 2025-04-12 04:05:40
- **Emission Level**: Moderate

Figure 7.1 Berthing Details Interface of the Port Management System

Figure 7.1 indicates that the ship has been allocated to berth number 3, with the docking period scheduled from April 11, 2025, 8:05 PM to April 12, 2025, 4:05 AM. During this time, the emission level for the operation is classified as "Moderate," indicating a balanced impact on the environment based on the ship's fuel consumption and other factors during the docking period.

Case 2: Ship Assigned to Waiting Area:

This output shows a ship that has not yet been assigned a berth and is currently waiting.

Details:

- **Status**: Your ship is assigned to the waiting area. We will notify you when a berth becomes available.
- **Emission Level**: Low

Figure 7.2 Berthing Status Interface Display for Waiting Area Assignment

From Figure 7.2, the ship is not yet assigned to a specific berth due to unavailability. While waiting, the emission level is classified as "Low," indicating minimal environmental impact, as the ship is not actively docked or consuming fuel at the port. The system will notify the ship when a berth becomes available for scheduling.

CHAPTER 8
CONCLUSION

8. CONCLUSION

This project effectively combines machine learning and optimization tools to improve maritime operations, with a focus on both emission forecasting and berth scheduling optimization. A predictive system utilizing a Random Forest-based model can accurately calculate expected emissions for shipping, allowing ports to utilize a data-driven approach to reduce emissions. The berth scheduling optimization is conducted with Mixed Integer Linear Programming (MILP) to minimize waiting time and fuel consumption for ships, while minimizing emissions through optimized scheduling of berths. The predictive and machine learning tools, combined with real-time data availability, improve the efficiency of port operations while providing a management framework for port authorities to make adaptive and iterative decisions. This framework combines predictive analytics and intelligent scheduling to help avoid delays, fuel costs, and emissions, while improving sustainability and efficiency in the maritime logistics sector.

Future enhancements could develop deep-learning models that would improve the accuracy of predictions for emissions and to optimize berth assignment and scheduling on more complicated patterns. In addition, to optimize vessel location and berth assignments even more, future upgrades could include real-time satellite data to update vessel tracking and berth assignments. Further updates could also include an ability to implement multi-port coordination by linking environmental vessels' operation and scheduling with respect to multiple ports where vessels may need to change to a different port for berthing. Future upgrades would also include establishing itself with other blockchain as data management for vessels at sea could further help with increased transparency and security and efficiency with the exchange of maritime data. With all these enhancements built into a model, a more intelligent, scalable, and sustainable port management system would assist all parties involved.

REFERENCES

1. M. Mansoursamaei, M. Moradi, G. González-Ramírez, & & E. Lalla-Ruiz. (2023), "Machine learning for promoting environmental sustainability in ports", Artificial Intelligence Approaches for Green Transportation Planning.
2. Mingyuan Yue, Yubing Wang, Siqing Guo, Lei Dai, & Hao Hu (2024), "A multi-objective optimization study of berth scheduling considering shore side electricity supply", Transportation Research Part E: Logistics and Transportation Review.
3. Lorenz Kolley, Nicolas Rückert, Marvin Kastner, Carlos Jahn, & Kathrin Fischer (2022), "Robust berth scheduling using machine learning for vessel arrival time prediction", Flexible Services and Manufacturing Journal.
4. Yuquan Du, Qiushuang Chen, Xiongwen Quan, Lei Long, & Richard Y.K. Fung (2011), "Berth allocation considering fuel consumption and vessel emissions", Transportation Research Part E: Logistics and Transportation Review.
5. T. Fletcher, Vikram Garaniya, Shuhong Chai, Rouzbeh Abbassi, Hongyang Yu, Thuy Chu Van, Richard J. Brown, & Faisal Khan (2018), "An application of machine learning to shipping emission inventory", The International Journal of Maritime Engineering.
6. Carlos D. Paternina-Arboleda, Dayana Agudelo-Castañeda, Stefan Voß, & Shubhendu Das, "Towards Cleaner Ports: Predictive Modeling of Sulfur Dioxide Shipping Emissions in Maritime Facilities Using Machine Learning", Sustainability in Logistics and Supply Chain Management.
7. Wei Guo, Minghao Ji, & Hai Zhu (2021), "Multi-Period Coordinated Optimization on Berth Allocation and Yard Assignment in Container Terminals Based on Truck Route", IEEE Access.
8. Fabregat, A., García-Martínez, A., Rizza, U., & Millán, & M. M. (2021), "Impact of port and ship traffic on urban air quality: A case study", Atmospheric Environment.
9. Wenxin Xie, Yong Li, Yang Yang, Peng Wang, Zhishan Wang, Zhaoxuan Li, Qiang Mei, & Yaqi Sun, (2023), "Maritime greenhouse gas emission estimation and

forecasting through AIS data analytics: a case study of Tianjin port in the context of sustainable development”, Frontiers in Marine Science.

10. Houjun Lu, & Xiao Lu (2025), “Joint Optimization of Berths and Quay Cranes Considering Carbon Emissions: A Case Study of a Container Terminal in China”, Journal of Marine Science and Engineering.
11. Zhihui Hu, Yongxin Jing, Qinyou Hu, Sukanta Sen, Tianrui Zhou, & Mohd Tarmizi Osman (2019), "Prediction of Fuel Consumption for Enroute Ship based on Machine Learning", IEEE Access.
12. Yinchen Lin, & Chuanxu Wang (2025), “Prediction of ship CO2 emissions and fuel consumption using Voting-BRL model”, Sustainability.
13. Haolin Li, Jiajing Gao, Lu Zhen, & Xueting He (2024), “Berth and yard scheduling optimization for a port with a diagonal yard layout”, Flexible Services and Manufacturing Journal.
14. Yonggai Dai, Zongchen Li, & Boyu Wang (2023), “Optimizing Berth Allocation in Maritime Transportation with Quay Crane Setup Times Using Reinforcement Learning”, ResearchGate.
15. N. V. Chawla, K. W. Bowyer, L. O. Hall, & W. P. Kegelmeyer (2002), “SMOTE: Synthetic minority over-sampling technique”, Journal of Artificial Intelligence Research
16. H.L. Ma, S.H. Chung., H.K. Chan, & Li Cui (2017),” An integrated model for berth and yard planning in container terminals with multi-continuous berth layout”, Annals of Operations Research.
17. Xing Jiang, Ming Zhang, Jiahui Shi, & Weifeng Li (2024), “Optimization of integrated scheduling of restricted channels, berths, and yards in bulk cargo ports considering carbon emissions”, ResearchGate.
18. Pietukhov, R., Ahtamad, M., Faraji-Niri, M., & El-Said, T. (2023), “A hybrid forecasting model with logistic regression and neural networks for improving key performance indicators in supply chains”, Supply Chain Analytics.
19. Mohamed M. Ahmed, & Mohamed Abdel-Aty (2013), “Application of stochastic gradient boosting technique to enhance reliability of real-time risk assessment: Use of automatic vehicle identification and remote traffic microwave sensor data.”

PLAGARISAM REPORT

DrillBit

The Report is Generated by DrillBit Plagiarism Detection Software

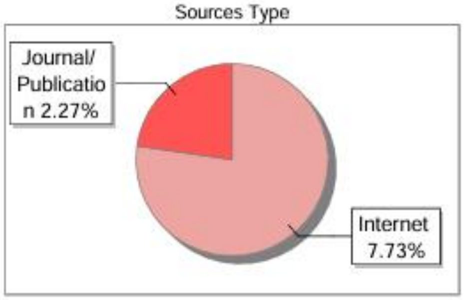
Submission Information

Author Name	P. Saritha Hepsibha
Title	A Framework for Maritime Emission Prediction and Optimal Berth Scheduling to Reduce Environmental Impact
Paper/Submission ID	3448152
Submitted by	sureshs.nice@gmail.com
Submission Date	2025-03-29 13:49:26
Total Pages, Total Words	6, 3165
Document type	Article

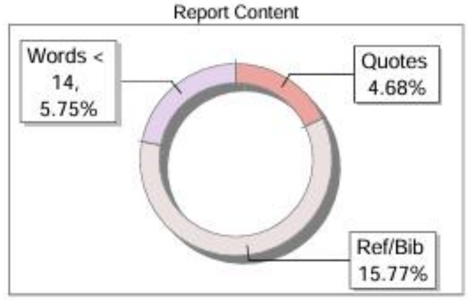
Result Information

Similarity **10 %**

Sources Type



Report Content



Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Source: Excluded < 14 Words	Not Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

A Unique QR Code use to View/Download/Share Pdf File

DrillBit

DrillBit Similarity Report

SIMILARITY %	MATCHED SOURCES	GRADE
10	23	A

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	thatware.co	1	Internet Data
2	Thesis Submitted to Shodhganga Repository	1	Publication
3	fastercapital.com	1	Internet Data
4	frontiersin.org	1	Internet Data
5	frontiersin.org	<1	Internet Data
6	www.rees-journal.org	<1	Internet Data
7	aissmsioit.org	<1	Publication
8	www.doaj.org	<1	Publication
9	www.bengalurutechsummit.com	<1	Internet Data
10	jutif.if.unsoed.ac.id	<1	Internet Data
11	www.researchgate.net	<1	Internet Data
12	eprints.uai.ac.id	<1	Internet Data
13	imagebuildingmedia.com	<1	Internet Data
14	macabacus.com	<1	Internet Data

PAPER COMMUNICATED

A Framework for Maritime Emission Prediction and Optimal Berth Scheduling to Reduce Environmental Impact

[1] Gollangi Ojaswini Sree, [2] Samuroutu Niveditha, [3] Kurmapu Hemanth Sai, [4] Aakifah
[1,2,3,4] Department of Information Technology, Anil NeeruKonda Institute of Technology and Sciences, Visakhapatnam, Andhra Pradesh.

Email: [1]ojaswinisree@gmail.com, [2] nivedithasamuroutu@gmail.com,[3] hemanthsai8003@gmail.com, [4] aakifah968@gmail.com

Abstract: Despite being essential to international trade, the maritime sector greatly contributes to environmental pollution through carbon emissions and ineffective port operations. Emission forecasting and effective berth scheduling are two significant maritime logistics issues that are addressed in this study using an integrated approach that combines machine learning and optimization algorithms. The goal of the project is to create a predictive model that uses real-time data from environmental variables, fuel consumption, and vessel movements to accurately estimate maritime emissions. To improve the accuracy of emission forecasts, machine learning techniques will be applied to the analysis of both historical and current data. Furthermore, this study provides a framework for berth scheduling optimization that can meet different shipping demands while ensuring lower fuel consumption and fewer delays. Port operations will be streamlined using modern optimization algorithms that maintain both environmental sustainability and economic viability. By integrating these two components, the proposed approach enhances decision-making for maritime stakeholders and makes it possible to create informed scheduling and emission control plans. The results will reduce the carbon footprint of maritime logistics while improving operational effectiveness. This study is a significant step toward sustainable and intelligent maritime transportation, which will benefit the global economy and environment.

Index terms: Maritime Industry, Port Operations, Berth Assigning, Emission, Sustainability.

1. Introduction

One essential element of international trade is maritime transportation. Maritime transportation is a vital component of international trade and logistics because it facilitates the movement of goods across borders. However, concerns regarding environmental sustainability are raised by the industry's significant contribution to greenhouse gas emissions, air pollution, and climate change. As ports and shipping operations expand, there is an increasing need for more efficient and sustainable maritime solutions [1].

One of the biggest environmental problems facing maritime operations is emissions from ships while they are berthing in ports and transiting through them. Ineffective berth scheduling increases carbon emissions by causing long wait times and excessive fuel use [2]. To address this issue, a comprehensive system for forecasting maritime emissions and scheduling berths is required. By combining state-of-the-art data analytics and optimization techniques, this system aims to increase operational efficiency while lowering the environmental impact of maritime logistics [3].

The economic and environmental advantages of this research make it significant. In addition to lowering congestion and increasing port efficiency, optimized berth scheduling also lowers operating expenses and fuel consumption [4]. Furthermore, precise emissions forecasting makes it possible to adhere to environmental standards, which lessens the negative climate effects of shipping operations.

Therefore, this research focuses on emission forecasting and berth scheduling optimization to improve maritime operations. A Random Forest-based machine learning model is developed to predict shipping emissions, helping ports make proactive decisions for sustainability. Additionally, a Mixed Integer Linear Programming (MILP) model is designed to optimize berth allocation, ensuring efficient port operations while minimizing environmental impact. The rest of this paper is organized as follows: Section 2 provides a review of existing literature, identifying key insights. Section 3 presents the system architecture and our strategic approach. Section 4 assesses the effectiveness of the proposed methodologies. Finally, the study concludes with key findings.

2. Literature Review

In recent years, there has been a growing interest in applying machine learning techniques to various challenges in maritime operations, particularly in reducing emissions, optimizing port logistics, and improving air quality. Several studies have explored how machine learning models can be leveraged to enhance environmental sustainability by forecasting and mitigating shipping emissions. Fletcher et al. developed predictive models using machine learning to create a comprehensive shipping emission inventory [5]. Similarly, Paternina-Arboleda et al. assessed sulfur dioxide emissions at maritime facilities, contributing to the development of greener port initiatives [6]. Expanding on this, Peng et al. introduced a machine learning model that incorporated green port activities, focusing on predicting energy consumption at ports, thus demonstrating the potential of data-driven sustainability approaches [7].

Various studies have focused on using machine learning to streamline port operations and optimize scheduling strategies. Guo et al introduced a multi-period coordinated optimization model for yard and berth assignment in container terminals, demonstrating how machine learning can enhance port logistics [8]. Beyond emissions reduction and air quality management, machine learning has also been applied to maritime infrastructure optimization, ship scheduling, and terminal management. Christiansen and Fagerholt emphasized the importance of robust ship scheduling strategies across different time frames [9].

These studies demonstrate the role of machine learning in improving maritime logistics and environmental outcomes by integrating predictive analytics, real-time data, and optimization techniques.

3. Methodology:

3.1 Architecture:

The proposed architecture integrates machine learning for emission forecasting with MILP-based berth scheduling through a centralized web application. As shown in Figure 1, our system collects real-time ship data, such as vessel type, size, fuel usage, arrival time, and cargo details. A web application that updates and retrieves berth schedules from a database process this data. We employ the Random Forest model, which helps port authorities make better decisions by analyzing historical and current data [10]

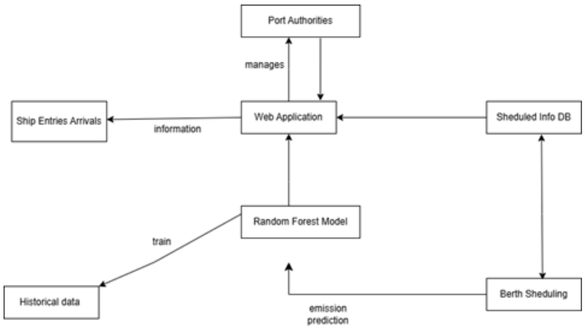


Figure 1: Architectural Flow of the Maritime Emission Prediction and Scheduling System

Simultaneously, berth scheduling is optimized through the use of Mixed Integer Linear Programming (MILP), which lowers fuel consumption and ship waiting times [10]. Lastly, the web application allows for real-time schedule monitoring and modification, improving the efficiency of port operations. The key components as shown in Figure 1 are:

Ship Arrivals: Ships offer up-to-date information on cargo details, estimated arrival time, fuel consumption, emission levels, vessel type, size, and weight. The web application processes this data to maintain the emissions models and berth scheduling current. Quick reactions to modifications in ship traffic and port conditions are made possible by the system's assurance of a smooth and accurate information flow. The system can improve berth assignments and forecast emissions by incorporating this real-time data, which will ultimately increase port efficiency and lessen its environmental impact [11].

Web Application: By processing real-time ship data and enabling communication between the berth scheduling engine, the emission prediction model, and the ship arrival system, the web application acts as the system's main control center. It serves as the port authorities' interface, enabling them to keep an eye on and modify schedules as necessary. Port operators can react swiftly to changes in ship arrivals and environmental conditions thanks to the web application's processing of real-time data and generation of actionable insights [12].

Scheduled Data (DB): In order to maintain consistency between ship arrival times and berth availability, the database keeps track of ship data and berth schedules. Berth allocations are continuously optimized thanks to its support for dynamic updates based on current ship status and port traffic.

Efficient information retrieval and storage are made possible by the database, which offers a solid basis for scheduling and emission prediction. This guarantees that port operations continue to be effective and flexible in the face of unforeseen delays or changes in ship arrivals.

Random Forest Model (Emission Prediction):

The system uses a Random Forest model to predict emissions based on historical data and real-time ship movement information. The model considers vessel size, speed, fuel type, and environmental factors to generate accurate emission forecasts. The output from this model is displayed through the web application, allowing port authorities to monitor and manage emissions more effectively. This predictive capability ensures compliance with environmental regulations and helps reduce the ecological impact of port operations [13].

Berth Scheduling (MILP): Mixed Integer Linear Programming (MILP) is used to optimize berth scheduling in order to minimize ship waiting times, improve berth utilization, and lower emissions and fuel consumption. The model enhances efficiency and sustainability by dynamically modifying schedules in response to port traffic and real-time ship arrivals [14].

Port Authorities: The web application is used by port authorities to control emissions and ship schedules. With the use of real-time data and forecasts, they can modify ship priorities and berth assignments. This aids in their decision-making to increase efficiency and lower fuel and emission levels [15].

3.2 Implementation:

3.2.1 Software Requirements:

The model was implemented using Python with Scikit-learn for training and deploying the Random Forest model for emission forecasting and Flask for creating a lightweight API to handle model predictions and scheduling updates.

3.2.2 Dataset Description:

The dataset used for training and testing consists of 1,000,000 entries with 9 features. It includes details such as ship type, size, fuel type, vessel age, fuel consumption, engine type, emission control technology, load factor, and emission scores.

3.2.2 Emission Forecasting Using Machine Learning:

We developed a machine learning-based predictive model to estimate ship emissions using Random Forest, analyzing both historical and real-time data for accurate predictions.

3.2.2.1 Data Preparation:

To ensure accuracy and reliability we have normalized numerical features, removed outliers, and handled missing values to ensure data accuracy. Additionally, Synthetic Minority Over-sampling Technique (SMOTE) was applied to balance class distribution, improving the model’s fairness and predictive performance [16], where SMOTE is a method used to address class imbalance in datasets.

3.2.2.2 Model Training:

The Random Forest model was trained using an 80-20 data split, with 80% of the data for training and 20% for validation. The input data included ship type, fuel type, vessel age, engine type, fuel consumption, load factor, and emission control technology. The model was designed to predict ship emissions, categorizing them as Low, Medium, or High based on these factors. It follows a bootstrapping technique, creating multiple decision trees using different random subsets of data to improve generalization. At each split, a random subset of features was chosen to reduce overfitting. Hyperparameter tuning was done by adjusting the number of trees (50), maximum depth (10), and minimum samples per split (7) to ensure stability and reliable decision-making. During prediction, each tree provides an output, and the final result is obtained through majority voting for classification or averaging for regression, making the model more accurate and robust.

3.2.2.3 Mathematical Approach Of Random Forest Model:

Random Forest handles data by creating bootstrapped subsets, where the dataset is randomly sampled with replacement to generate multiple training sets. At each tree node, a random subset of features is selected for splitting, ensuring diversity among trees. The best feature is chosen based on Gini Impurity(G) for classification, which measures the likelihood of an incorrect classification if a random sample were classified according to the distribution of classes in a node.

It is calculated as:

$$G = 1 - \sum_{i=0}^{c} P_i^2 \qquad (1)$$

From equation (1) P_i represents the probability of a class in the node, and i represents each class, c represents number of classes. A lower Gini Impurity value indicates a purer node with fewer mixed classes.

For regression tasks, the Mean Squared Error (MSE) is minimized:

$$MSE = \frac{1}{n}\Sigma(y_i - \hat{y}_i)^2 \qquad (2)$$

From Equation (2) where y_i the actual value, and $\hat{y}_i$ is the predicted value, n is number of samples and I is index of each data point..

3.2.2.4 Prediction and Evaluation:

Emission levels were predicted using fuel consumption, and vessel size. Emissions were categorized by the model as High, Moderate, and Low orders. Evaluation was performed using delicacy, perfection, recall, and F1- score, with the model achieving 98% delicacy and all other criteria exceeding 96%, indicating high prophetic trustability and balanced performance.

3.2.3 Berth Scheduling Optimization Using MILP:

Berth scheduling can be optimized using Mixed Integer Linear Programming (MILP) by minimizing waiting times and fuel consumption while ensuring fair berth allocation. Optimizing berth can greatly increase port efficiency and lower emission [17].

3.2.3.1 Scheduling and Optimization:

The MILP model assigns ships to berths using decision variables such as berth assignment, start time, and waiting time. The objective function minimizes total ship waiting time and fuel consumption, balancing operational efficiency and sustainability. Constraints ensure berth capacity limits, prioritize high-ranking ships, and favor vessels with lower emissions.

Ships are categorized based on size, priority level, fuel consumption, and emissions. A priority score is assigned, favouring emergency cases and eco-friendly vessels. Ships are allocated berths immediately if available; otherwise, they enter a waiting queue, ordered by priority and emission index [18].

As shown in figure 3 it confirms that that a ship has been successfully assigned to a berth.

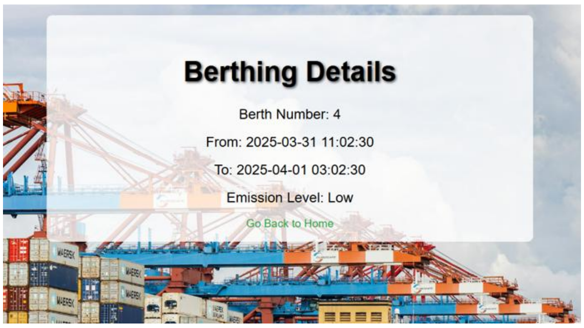


Figure 3: Berthing Details

4. Results and Discussions:

4.1. Performance Metrics and Evaluation:

After training, the machine learning model for ship emission forecasting was evaluated using multiple quantitative metrics. The testing confirmed that the model generalizes well to new data and accurately predicts ship emissions under varying port conditions.

4.2 Evaluation Metrics:

To assess the model's performance, the following metrics were applied:

- **Mean Average Precision (mAP):**
    - mAP@50: Measures the model's accuracy at an Intersection over Union (IoU) threshold of **0.5**, confirming that the model accurately predicts ship emissions and berth schedules when moderate overlap is allowed.
    - mAP@50-95: Evaluates the model's precision across a range of IoU thresholds (0.50 to 0.95), providing a comprehensive assessment of its detection and prediction capabilities.
- **Intersection over Union (IoU):** IoU measures the overlap between predicted and actual ship emissions and berth assignments. The model was tuned to achieve an IoU threshold of **0.7**, ensuring high alignment between predicted and actual outcomes.
- **Accuracy:** The model achieved an overall accuracy of 98%, indicating high consistency in predicting ship emissions and berth schedules based on real-time data and historical patterns. High accuracy reflects the model's ability to minimize both false positives and false negatives.

- **Precision and Recall:**
    - Precision: Measures the ratio of rightly predicted ship emissions and berth assignments to total predictions. The model demonstrated a high precision of 1.00, 0.98, and 0.96 across different classes, suggesting minimal false positive rates.
    - Recall: Measures the ratio of correctly predicted positive cases to total actual positive cases. The recall scores of 0.98, 0.99, and 0.99 indicate that the model successfully captures utmost true positive cases, ensuring that few positive cases are missed.
- **F1-Score:** The F1-score is the harmonious mean of precision and recall, giving a balanced measure of the model's accuracy. The model achieved F1-scores of 0.99, 0.99, and 0.97 across different classes, showing that it maintains a good balance between precision and recall.
- **Confusion Matrix Analysis:** The confusion matrix illustrates the model's classification performance across different ship emission levels and berth categories. High true positive rates and low misclassification rates confirm the model's ability to differentiate between different ship and berth classes accurately. The normalized confusion matrix provides insight into class-wise accuracy, helping to identify potential misclassification patterns for further optimization.

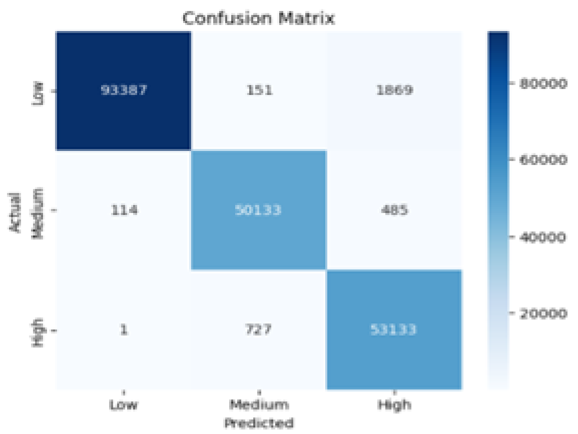


Figure 2: Confusion Matrix of the Random Forest Model for Emission Classification.

As shown in figure 2, The confusion matrix illustrates the model's accuracy in classifying ship emissions into High, Moderate, and Low categories. Most predictions align with the diagonal, reflecting high classification accuracy and minimal misclassification.

4.3 Comparison with Existing Models:

To evaluate the efficiency of the Random Forest model, we conducted experiments using a proprietary dataset collected from maritime operations. The dataset consists of 100,000 samples with 9 features, which was split into 80% for training and 20% for testing. We applied Random Forest, Logistic Regression, and Gradient Boosting to predict the emission levels and fuel usage patterns. The outcomes of this comparison are outlined in Table 1.

Table 1: Comparison of Machine Learning Models for Emission Forecasting.

Model	Precisi-on	Recall	F1-Score	Accur-acy	Inferen-ce Time-(ms)
Logistic Regression	69.2%	70.1%	69.2%	70.1%	25
Gradient Boosting	92.1%	91.8%	91.9%	91.7%	55
Random Forest	98.0%	98.1%	98.0%	98.0%	45

From the Table 1, Our approach utilizes the Random Forest model, which outperformed better than other logistic regression [19] and gradient boosting [20], achieving 98% accuracy with a precision of 98% and a recall of 98.1%. Its fast inference time of 45 ms makes it suitable for real-time prediction of ship emissions and berth scheduling.

5. Conclusion:

This study focused on the integration of machine learning and optimization techniques to enhance maritime operations through emission forecasting and berth scheduling optimization. A Random Forest-based predictive model was developed to estimate shipping emissions, enabling ports to make informed decisions that contribute to environmental sustainability. Additionally, a Mixed Integer Linear Programming (MILP) model was employed to optimize berth scheduling, reducing ship waiting times, fuel consumption, and overall emissions.

The experimental results demonstrated that the Random Forest model outperformed traditional methods such as Logistic Regression and Gradient Boosting in emission prediction, achieving an accuracy of 98% with a precision of 98% and a recall of 98.1%. The berth scheduling system, optimized through MILP, efficiently allocated berths based on ship priority, size, and emission levels. This approach not only improved port efficiency but also contributed to minimizing environmental impact by reducing idle times and unnecessary fuel consumption. The integration of real-time data through a web application provided a dynamic decision-making framework for port authorities, enhancing operational flexibility and responsiveness.

This research demonstrates that machine learning and optimization can improve maritime logistics by making port operations more efficient and eco-friendly. By combining predictive analytics with intelligent scheduling, the system reduces delays, fuel consumption, and emissions. Future improvements could include deep learning for better predictions and real-time satellite data for more accurate berth allocations. Overall, this approach supports sustainable and data-driven maritime operations while ensuring compliance with global regulations.

6. References:

[1] M. Mansoursamaei, M. Moradi, G. González-Ramírez, & & E. Lalla-Ruiz. (2023), "Machine learning for promoting environmental sustainability in ports", Artificial Intelligence Approaches for Green Transportation Planning.

[2]Mingyuan Yue, Yubing Wang, Siqing Guo, Lei Dai, & Hao Hu (2024), "A multi-objective optimization study of berth scheduling considering shore side electricity supply", Transportation Research Part E: Logistics and Transportation Review.

[3] Lorenz Kolley, Nicolas Rückert, Marvin Kastner, Carlos Jahn, & Kathrin Fischer (2022), "Robust berth scheduling using machine learning for vessel arrival time prediction", Flexible Services and Manufacturing Journal.

[4] Yuquan Du, Qiushuang Chen, Xiongwen Quan, Lei Long, & Richard Y.K. Fung (2011), "Berth allocation considering fuel consumption and vessel emissions", Transportation Research Part E: Logistics and Transportation Review.

[5] T. Fletcher, Vikram Garaniya, Shuhong Chai, Rouzbeh Abbassi, Hongyang Yu, Thuy Chu Van, Richard J. Brown, & Faisal Khan (2018), "An application of machine learning to shipping emission inventory", The International Journal of Maritime Engineering.

[6] Carlos D. Paternina-Arboleda, Dayana Agudelo-Castañeda, Stefan Voß, & Shubhendu Das, "Towards Cleaner Ports: Predictive Modeling of Sulfur Dioxide Shipping Emissions in Maritime Facilities Using Machine Learning", Sustainability in Logistics and Supply Chain Management.

[7] Yun Peng, Huakun Liu, Xiangda Li, Jian Huang, & Wenyuan Wang (2020), "Machine learning method for energy consumption prediction of ships in port considering green ports", Energy Reports.

[8] Wei Guo, Minghao Ji, & Hai Zhu (2021), "Multi-Period Coordinated Optimization on Berth Allocation and Yard Assignment in Container Terminals Based on Truck Route", IEEE Access.

[9] Fabregat, A., García-Martínez, A., Rizza, U., & Millán, & M. M. (2021), "Impact of port and ship traffic on urban air quality: A case study", Atmospheric Environment.

[10] Wenxin Xie, Yong Li, Yang Yang, Peng Wang, Zhishan Wang, Zhaoxuan Li, Qiang Mei, & Yaqi Sun, (2023), "Maritime greenhouse gas emission estimation and forecasting through AIS data analytics: a case study of Tianjin port in the context of sustainable development", Frontiers in Marine Science.

[11] Houjun Lu, & Xiao Lu (2025), "Joint Optimization of Berths and Quay Cranes Considering Carbon Emissions: A Case Study of a Container Terminal in China", Journal of Marine Science and Engineering.

[12] Zhihui Hu, Yongxin Jing, Qinyou Hu, Sukanta Sen, Tianrui Zhou, & Mohd Tarmizi Osman (2019), "Prediction of Fuel Consumption for Enroute Ship based on Machine Learning", IEEE Access.

[13] Yinchen Lin, & Chuanxu Wang (2025), "Prediction of ship CO2 emissions and fuel consumption using Voting-BRL model", Sustainability.

[14] Haolin Li, Jiajing Gao, Lu Zhen, & Xueting He (2024), "Berth and yard scheduling optimization for a port with a diagonal yard layout", Flexible Services and Manufacturing Journal.

[15] Yonggai Dai, Zongchen Li, & Boyu Wang (2023), "Optimizing Berth Allocation in Maritime Transportation with Quay Crane Setup Times Using Reinforcement Learning", ResearchGate.

[16] N. V. Chawla, K. W. Bowyer, L. O. Hall, & W. P. Kegelmeyer (2002), "SMOTE: Synthetic minority over-sampling technique", Journal of Artificial Intelligence Research

[17] H.L. Ma, S.H. Chung., H.K. Chan, & Li Cui (2017)," An integrated model for berth and yard planning in container terminals with multi-continuous berth layout", Annals of Operations Research.

[18] Xing Jiang, Ming Zhang, Jiahui Shi, & Weifeng Li (2024), "Optimization of integrated scheduling of restricted channels, berths, and yards in bulk cargo ports considering carbon emissions", ResearchGate.

[19] Pietukhov, R., Ahtamad, M., Faraji-Niri, M., & El-Said, T. (2023), "A hybrid forecasting model with logistic regression and neural networks for improving key performance indicators in supply chains", Supply Chain Analytics.

[20] Mohamed M. Ahmed, & Mohamed Abdel-Aty (2013), "Application of stochastic gradient boosting technique to enhance reliability of real-time risk assessment: Use of automatic vehicle identification and remote traffic microwave sensor data.", Transportation Research Record: Journal of the Transportation Research Board.