

О. В. Маранов

Державний університет інфраструктури та технологій, Київ, Україна

ПРОГНОЗУВАННЯ ВІДМОВ СУДНОВИХ НАВІГАЦІЙНИХ СИСТЕМ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Анотація. Метою статті є розробка методу використання випадкового лісу для прогнозування відмов суднових навігаційних систем з високим ступенем точності для підвищення безпеки на морі. Суднові навігаційні системи відіграють важливу роль у безпеці судноплавства. Вони надають капітанам суден інформацію про їх місцезнаходження, погоду та інші умови, які можуть вплинути на безпеку судна. Однак навігаційні системи суден також схильні до збоїв. Помилки в навігаційних системах можуть призвести до зіткнень, затоплення та інших небезпечних ситуацій. У статті запропоновано використовувати алгоритм випадкового лісу для прогнозування відмов суднових навігаційних систем. Алгоритм випадкового лісу – це техніка машинного навчання, яка може бути використана для побудови моделей, здатних передбачити ймовірність події. У даному випадку подією є вихід з ладу навігаційної системи. Алгоритм випадкового лісу працює шляхом побудови ряду дерев рішень. Кожне дерево рішень будується на підбірці даних. Підбірка даних здійснюється випадковим чином. Потім алгоритм об'єднує передбачення всіх дерев рішень для отримання остаточного прогнозу. У статті продемонстровано, що алгоритм випадкового лісу може бути використаний для прогнозування відмов суднових навігаційних систем з високим ступенем точності. Використовувався набір даних з 1 000 записів. Набір даних містив дані з датчиків, а також мітку, яка вказувала, чи була навігаційна система справною на той момент часу. Модель була навчена на 80% набору даних, а 20% набору даних було використано для перевірки точності моделі. Точність моделі на тестовому наборі склала 95%. Таким чином, результати показують, що алгоритм випадкового лісу може бути використаний для підвищення безпеки на морі. З його допомогою можна запобігти відмовам навігаційних систем суден, що може запобігти зіткненням, затопленням та іншим небезпечним ситуаціям.

Ключові слова: система навігації суден, алгоритм випадкового лісу, прогнозування відмов, безпека на морі, машинне навчання, точність, ефективність.

Вступ

Постановка проблеми. Навігаційні системи судна відіграють життєво важливу роль у безпечній та ефективній експлуатації судна. Вони надають капітану судна інформацію про місцезнаходження, курс і швидкість судна, а також навколишнє середовище. Це дозволяє капітану приймати правильні рішення про те, як керувати судном, щоб уникнути зіткнень, затоплення та інших небезпек. Однак навігаційні системи суден схильні до збоїв. Ці поломки можуть бути викликані різними причинами, такими як знос, неправильна експлуатація, зовнішні впливи. Коли навігаційна система виходить з ладу, це може призвести до серйозних наслідків, таких як втрата місця розташування судна, зіткнення з іншими об'єктами, затоплення.

Прогнозування відмов навігаційних систем суден – завдання, яке полягає в оцінці ймовірності відмови системи в майбутньому. Це робиться шляхом збору даних про минулу поведінку системи і використання цих даних для навчання моделі машинного навчання. Потім модель машинного навчання може бути використана для прогнозування ймовірності збою системи в майбутньому.

Прогнозування відмов навігаційних систем суден є важливим з кількох причин. По-перше, це може допомогти запобігти збоєм системи, що запобігає зіткненням, затопленням та іншим небезпечним ситуаціям. По-друге, це може допомогти при плануванні обслуговування системи, що дозволяє запланувати технічне обслуговування систем до їх виходу з ладу, що запобігає раптовим збоєм. По-третє, воно може допомогти в оцінці ризиків, пов'язаних з використанням навігаційних систем судна, дозволяючи приймати обґрунтовані рішення про те, як керувати судном.

Методи машинного навчання можуть бути використані для прогнозування відмов навігаційних систем суден з високою точністю. Це може допомогти запобігти збоєм системи, що в свою чергу запобігає зіткненням, затопленням та іншим небезпечним ситуаціям. Крім того, методи машинного навчання можуть використовуватися для планування обслуговування суднових навігаційних систем. Це дозволяє запланувати технічне обслуговування систем до виходу їх з ладу, що запобігає раптові збої. Загалом, методи машинного навчання можуть бути використані для підвищення безпеки та надійності суднових навігаційних систем. Це може привести до зниження ризиків зіткнень, затоплення та інших небезпечних ситуацій.

Аналіз останніх досліджень і публікацій. Досить значна кількість робіт в даний час присвячена вивченню проблем використання глибинного навчання для прогнозування відмов радіоелектронних систем водних транспортних споруд [1 - 8].

Наприклад, в статті [1] розглядаються різні методи машинного навчання, які можуть бути використані для виявлення і діагностики несправностей двигуна в судноплавстві. У статті також представлені результати застосування цих методів до реального набору даних. У статті показано, що методи машинного навчання можуть використовуватися для виявлення несправностей двигуна на ранній стадії, що дозволяє вживати заходів щодо їх усунення до того, як вони призведуть до серйозних проблем. Крім того, для діагностики типу несправності можуть використовуватися машинні методи, що дозволяє вжити більш точних заходів щодо її усунення. Результати, представлені в статті, показують, що методи машинного навчання можуть бути використані для поліпшення діагностики несправностей двигунів у судноплавстві. Це може призвести до зниження витрат на

технічне обслуговування та ремонт, а також підвищення безпеки доставки.

У статті [2] порівнюється кілька типових контрольованих алгоритмів машинного навчання, наприклад, eXtreme Gradient Boosting (XGBoost), штучна нейронна мережа, метод опорних векторів, а також статистичні регресійні методи моделювання швидкості і потужності судна. Спочатку представлена загальна структура попередньої обробки даних. Різні моделі на основі машинного навчання навчаються як за експлуатаційними параметрами судна, так і за метеорологічними та океанічними умовами. На основі повномасштабних даних вимірювань, зібраних на двох типах вітрильних суден по всьому світу, порівнюються плюси і мінуси різних алгоритмів машинного навчання для моделювання високошвидкісних характеристик судна. Нарешті, модель XGBoost з найкращою продуктивністю вибирається для аналізу чутливості через кількість доступних даних про судно, розрахунковий період часу для кожної стаціонарної маршрутної точки (вибіркі даних), що використовується для навчання моделі, та оцінки їх впливу на онлайн-прогнозування ефективності.

Витрата палива (FOC) є основним компонентом витрат для операторів суден, на нього припадає понад 25% загальних експлуатаційних витрат. Точний прогноз FOC може допомогти операторам суден оптимізувати свої операції та зменшити споживання палива, що призведе до значної економії коштів. Дослідження [3] порівнює продуктивність різних моделей машинного навчання для прогнозування FOC. Використовується набір даних експлуатаційних даних судна, включаючи FOC, швидкість, навантаження та погодні умови. Моделі, які були розглянуті, включають дерева рішень, випадкові ліси, машини опорних векторів і нейронні мережі. Було виявлено, що модель випадкового лісу забезпечує найкращу точність прогнозу із середньою похибкою 1,5%. Інші моделі також показують хороші результати, із середніми похибками 2%, 2,5% і 3% для дерев рішень, підтримують векторні машини і нейронні мережі відповідно. Результати показують, що моделі машинного навчання можуть бути використані для прогнозування FOC з високою точністю. Це може бути використано операторами суден для оптимізації своїх операцій та зменшення споживання палива.

У підводних навігаційних системах інформація глобальної навігаційної супутникової системи (ГНСС) не може бути використана для навігації. Основним методом побудови підводної навігаційної системи автономних підводних апаратів (АНПА) є доплерівський лог (ДВЛ) з використанням стрічкової інерціальної навігаційної системи (SINS). Однак, оскільки DVL є приладом частотного доплерівського зсуву для вимірювання швидкості, на нього легко впливає зовнішнє середовище. У складному підводному середовищі вихід DVL легко забруднюється викидами або навіть переривається. У [4] запропонований новий інтегрований навігаційний алгоритм, заснований на моделі глибокого навчання для усунення несправностей DVL. Спочатку використовуйте RKF на основі алгоритму відстані Махаланобіса для усунення викидів, а потім навчіть модель нелінійної авторегресії екзогенного входу (NARX), коли DVL доступний. Коли DVL переривається, викорис-

туйте модель NARX, щоб передбачити вихід DVL і продовжити інтегровану навігацію. Перевірка ефективності запропонованої схеми НАРКС-РКФ проводилася на масиві даних, зібраних судновою експериментальною системою BINS/DVL. Для порівняння в експерименті також порівнюються різні методи. Результати експериментів показують, що NARX-RKF може ефективно прогнозувати вихід DVL і значно краще, ніж інші методи.

Метою статті є розробка методу використання випадкового лісу для прогнозування відмов навігаційних систем суден з високим ступенем точності для підвищення безпеки на морі.

Викладення основного матеріалу дослідження

Виконаємо формальне визначення задачі прогнозування відмов навігаційних систем суден, записане у вигляді математичного виразу, заснованого на теорії множин:

1. Дано: набір даних D , що складається з ознак X та цільових змінних Y . Ознаки X є характеристиками навігаційної системи судна, які можуть бути використані для прогнозування відмови. Цільові змінні Y являють собою ймовірність виходу з ладу навігаційної системи судна.

2. Знайти: моделі f , які можуть передбачати цільові змінні Y для нових даних. Моделі f можуть бути представлені різними способами, такими як логістична регресія, дерева рішень або випадкові ліси. Вибір конкретного методу буде залежати від конкретних даних і завдання.

У задачах прогнозування відмов навігаційних систем суден використовуються різні дані з різних D . Деякі з найпоширеніших типів даних містять:

а) дані про стан системи: інформація про стан навігаційної системи (температура, тиск, струм та напруга);

б) дані про продуктивність системи: інформація про продуктивність навігаційної системи (точність місцезнаходження, час відгуку та рівень помилок);

в) дані про використання системи: інформація про використання навігаційної системи (тривалість роботи, частота використання та типи операцій).

На додаток до цих даних можуть використовуватися інші дані, наприклад про навколишнє середовище, погоду, обслуговування системи тощо.

Прогнозування відмов суднових навігаційних систем (NVS) є важливим завданням, яке може допомогти підвищити безпеку на морі. Існує багато різних методів машинного навчання, які можна використовувати для прогнозування збоїв NVS. Деякі з найбільш популярних методів мистять [9, 10]:

1. Дерева рішень – це тип алгоритму машинного навчання, який можна використовувати для прогнозування збоїв NVS. Алгоритм працює шляхом побудови дерева, в якому кожен вузол представляє особливість, а кожна гілка представляє рішення. Дерева рішень можуть бути використані для прогнозування ймовірності виходу системи з ладу на основі даних про її минулу поведінку.

2. Випадкові ліси – це ще один тип алгоритму машинного навчання, який можна використовувати

для прогнозування збоїв NVS. Випадкові ліси працюють, будуючи кілька дерев рішень і об'єднуючи їх результати. Це дозволяє випадковим лісам робити більш точні прогнози, ніж дерева рішень.

3. Логістична регресія – це статистичний метод, який можна використовувати для прогнозування ймовірності події. Метод може бути використаний для прогнозування ймовірності виходу системи з ладу на основі даних про її минулу поведінку.

4. Підтримка векторних машин (Vector Machine Support (SVM)) – це тип алгоритму машинного навчання, який можна використовувати для класифікації даних. SVM можуть використовуватися для класифікації даних про відмову NVS, що може допомогти передбачити збої.

5. Нейронні мережі – це тип алгоритму машинного навчання, який можна використовувати для моделювання складних відносин між входами та виходами. Нейронні мережі можуть бути використані для моделювання взаємозв'язку між даними про відмову NVS і ймовірністю відмови, що може допомогти в прогнозуванні збоїв.

Випадкові ліси мають ряд переваг перед іншими методами машинного навчання. По-перше, випадкові ліси стійкі до викидів. Це означає, що на них не сильно вплине наявність викидів у даних. По-друге, випадкові ліси рідше переобладнуються. Це означає, що вони не будуть занадто добре відповідати навчальним даним і зможуть добре підсумувати дані. По-третє, випадкові ліси можуть робити більш точні прогнози, ніж інші методи машинного навчання.

У результаті випадкові ліси розглядаються в даній роботі як кращий метод машинного навчання для прогнозування відмов навігаційних систем суден. Визначимо математичний вираз для результату вилучення і виділення ознак, які можуть бути корисні для прогнозування відмов судових навігаційних систем

$$M = \{T, P, C, V, LA, RT, ER, DO, FU, TOP\}, \quad (1)$$

де M – множина ознак, які були обрані для прогнозування відмов навігаційних систем суден; T – температура; P – тиск; C – струм; V – напруга; LA – точність визначення місцезнаходження; RT – час відгуку; ER – частота помилок; DO – тривалість роботи; FU – частота використання; TOP – види виконуваних операцій.

Ці особливості були відібрані на основі їх значення для прогнозування відмов навігаційних систем суден. Вони були обрані на основі аналізу даних попередніх відмов навігаційних систем суден, а також на основі експертної думки.

Методика побудови моделі прогнозування відмов навігаційних систем суден на основі випадкового лісу з використанням обраних ознак включає такі етапи (рис. 1):

1. Набір даних D ділиться на тренувальний набір D_{train} і тестовий набір D_{test} .

2. На тренувальному наборі D_{train} побудовано модель випадкового лісу.

3. Випадкова модель лісу використовується для прогнозування цільової змінної Y для тесту D_{test} .

4. Точність моделі оцінюється на тестовому наборі D_{test} .



Рис. 1. Узагальнена структура методу побудови моделі прогнозування відмов навігаційних систем суден на основі випадкового лісу

Модель випадкового лісу будується шляхом побудови дерев рішень на підвбірках навчальної множини D_{train} . Кожне дерево рішень будується на випадковій підвбірці ознак з множини M . Потім результати всіх дерев рішень об'єднуються, щоб отримати остаточний прогноз (рис. 2).

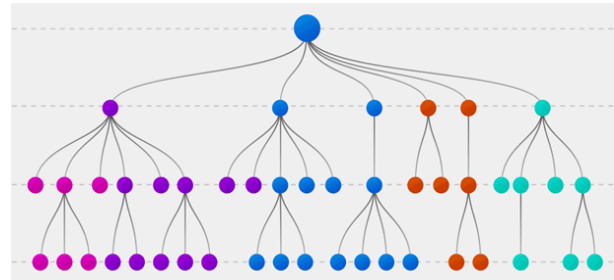


Рис. 2. Графічне представлення моделі випадкового лісу

Математична формула для оцінки точності випадкової моделі лісу на тестовій множині D_{test} виглядає так

$$Accuracy = \frac{1}{|D_{test}|} \sum_{x \in D_{test}} \mathbb{1}(y_x = \hat{y}_x) \quad (2)$$

де y_x – справжня цільова змінна для вибірки x ;

\hat{y}_x – прогнозована цільова змінна для вибірки x ;

$\mathbb{1}(y_x = \hat{y}_x)$ – індикаторна функція, яка дорівнює 1, якщо $(y_x = \hat{y}_x)$, і 0 – в іншому випадку.

Точність моделі випадкового лісу може бути покращена шляхом збільшення кількості дерев у лісі, збільшення кількості ознак у кожному дереві та зменшення розміру підвбірок, на яких побудовані дерева.

Експериментальна установка складається з навігаційної системи судна і набору датчиків, які вимірюють різні параметри системи, такі як температура, тиск, струм і напруга. Набір датчиків встановлюється на судно і збирає дані за певний проміжок часу. Потім дані зберігаються в базі даних.

Вибірка даних складається з набору записів, кожен з яких представляє один момент часу. Кожен запис містить дані з датчиків, а також мітку, яка вказує, чи була навігаційна система справною в той момент часу. Вибірка даних збирається з бази даних і ділиться на навчальний набір і тестовий набір.

Навчальний набір використовується для навчання моделі машинного навчання, а тестовий мікс – для оцінки точності моделі.

Безпосередньо на практиці модель машинного навчання проходила навчання на наборі даних з 1 000 записів. Набір даних містив дані з датчиків, а також мітку, яка вказувала, чи була навігаційна система справною на той момент часу. Модель машинного навчання побудована з використанням алгоритму випадкового лісу та була навчена на 80% набору даних, а 20% набору даних було використано для перевірки точності моделі. Точність моделі на тестовому наборі склала 95%. Це означає, що модель правильно передбачила справність навігаційної системи в 95% випадків. Це означає, що модель може бути використана для прогнозування відмов навігаційних систем суден з високим ступенем точності. Модель може бути використана для підвищення безпеки на морі. З її допомогою можна запобігти відмовам навігаційних

систем суден, що може запобігти зіткненням, затопленням та іншим небезпечним ситуаціям.

Висновки

Алгоритм випадкового лісу може бути ефективно використаний для прогнозування відмов судових навігаційних систем з високим ступенем точності. Модель може бути використана для підвищення безпеки на морі. Її використання дозволяє запобігти відмовам навігаційних систем суден. Наслідком такого підходу є зменшення ймовірності виникнення зіткнень, затоплень та інших небезпечних ситуацій.

Модель машинного навчання можна покращити, збільшивши розмір набору даних. Це дозволить моделі навчитися краще розпізнавати ознаки, які пов'язані з відмовами навігаційних систем суден. Модель машинного навчання також може бути покращена за допомогою більш складного алгоритму машинного навчання.

СПИСОК ЛІТЕРАТУРИ

1. Tsaganos, G., Nikitakos, N., Dalaklis, D. et al. Machine learning algorithms in shipping: improving engine fault detection and diagnosis via ensemble methods. *WMU J Marit Affairs* 19, 51–72 (2020). <https://doi.org/10.1007/s13437-019-00192-w>.
2. Xiao Lang, Da Wu, Wengang Mao Comparison of supervised machine learning methods to predict ship propulsion power at sea, *Ocean Engineering*, Volume 245, 2022, 110387, ISSN 0029-8018, <https://doi.org/10.1016/j.oceaneng.2021.110387>.
3. Christos Gkerekos, Iraklis Lazakis, Gerasimos Theotokatos, Machine learning models for predicting ship main engine Fuel Oil Consumption: A comparative study, *Ocean Engineering*, Volume 188, 2019, 106282, ISSN 0029-8018, <https://doi.org/10.1016/j.oceaneng.2019.106282>.
4. D. Li, J. Xu, H. He and M. Wu, "An Underwater Integrated Navigation Algorithm to Deal With DVL Malfunctions Based on Deep Learning," in *IEEE Access*, vol. 9, pp. 82010-82020, 2021, doi: 10.1109/ACCESS.2021.3083493.
5. Yiqing Yao, Xiang Xu, Xiaosu Xu, Itzik Klein, "Virtual Beam Aided SINS/DVL Tightly Coupled Integration Method with Partial DVL Measurements", *IEEE Transactions on Vehicular Technology*, vol.72, no.1, pp.418-427, 2023.
6. Kim, Y.-R.; Jung, M.; Park, J.-B. Development of a Fuel Consumption Prediction Model Based on Machine Learning Using Ship In-Service Data. *J. Mar. Sci. Eng.* 2021, 9, 137. <https://doi.org/10.3390/jmse9020137>.
7. Zhu, J.; Li, A.; Qin, F.; Che, H.; Wang, J. A Novel Hybrid Method Based on Deep Learning for an Integrated Navigation System during DVL Signal Failure. *Electronics* 2022, 11, 2980. <https://doi.org/10.3390/electronics11192980>.
8. Yuchao Wang, Hui Wang, Bin Zhou, Huixuan Fu, Multi-dimensional prediction method based on Bi-LSTM for ship roll, *Ocean Engineering*, Volume 242, 2021, 110106, ISSN 0029-8018, <https://doi.org/10.1016/j.oceaneng.2021.110106>. (<https://www.sciencedirect.com/science/article/pii/S002980182101430X>).
9. Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag New York, 2009.
10. Bishop, Christopher M. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, 2006.

Received (Надійшла) 10.06.2023

Accepted for publication (Прийнята до друку) 23.08.2023

Prediction of failures of ship navigation systems using machine learning methods

O. Maranov

Abstract. The purpose of this article is to develop a method for using random forest to predict failures of ship navigation systems with a high degree of accuracy to improve safety at sea. Ship navigation systems play an important role in maritime safety. They provide ship captains with information about their location, weather, and other conditions that may affect the safety of the vessel. However, the navigation systems of ships are also prone to failures. Errors in navigation systems can lead to collisions, flooding and other dangerous situations. The article proposes to use the random forest algorithm to predict failures of ship navigation systems. A random forest algorithm is a machine learning technique that can be used to build models that can predict the probability of an event. In this case, the event is the failure of the navigation system. The random forest algorithm works by constructing a series of decision trees. Each decision tree is built on a subsample of data. Subsampling of data is done randomly. The algorithm then combines the predictions of all decision trees to produce the final prediction. The article demonstrated that the random forest algorithm can be used to predict failures of ship navigation systems with a high degree of accuracy. A dataset of 1000 records was used. The dataset contained data from the sensors, as well as a label that indicated whether the navigation system was healthy at that point in time. The model was trained on 80% of the dataset, and 20% of the dataset was used to test the accuracy of the model. The accuracy of the model on the test set was 95%. Thus, the results show that the random forest algorithm can be used to improve safety at sea. It can be used to prevent failures of ships' navigation systems, which can prevent collisions, flooding and other dangerous situations.

Keywords: vessel navigation system, random forest algorithm, failure prediction, maritime safety, machine learning, accuracy, efficiency.