International Journal of Pioneering Technology and Engineering

Research Article

Classification of NOx Emission in Marine Engines Utilizing kNN-Based Machine Learning Algorithms

Samet Memiş [a*], **Ramazan Şener** [a]

[a] Department of Marine Engineering, Maritime Faculty, Bandırma Onyedi Eylül University, Bandırma, Balıkesir, Türkiye

Article Info

Article history

Received: 27/10/2024
Revised 1: 18/11/2024
Accepted: 05/12/2024

Keywords:
Marine Engines;
NOx Emission;
Classification;
Machine Learning;
kNN Classifier;
Soft Matrices

ABSTRACT

Marine diesel engines are crucial for powering large vessels in the maritime sector and are known for their efficiency across various industries. However, increasing environmental concerns and stringent regulations targeting air pollutants such as nitrogen oxides (NOx) and particulate matter (PM) have heightened the need for advanced emission control technologies. Addressing this challenge, the study focuses on developing a reliable method to predict NOx emission levels in marine engines, reducing reliance on resource-intensive experimental testing. Leveraging machine learning techniques, particularly k-nearest neighbors (kNN)-based algorithms, the research classifies NOx emissions in marine engines operating under the Reactivity-Controlled Compression Ignition (RCCI) strategy. Comparative performance analysis reveals that the FPFS-kNN algorithm achieves the highest accuracy (90.00%) alongside strong precision (84.23%), recall (82.37%), and F1 score (82.47%). These findings underscore the potential of machine learning in emission prediction and highlight directions for future exploration in this domain.

1. Introduction

Due to their robust performance and efficiency, diesel engines are a backbone across multiple industries, including automotive, railway transportation, power generation, and maritime operations[1]. In the marine sector, these engines play a crucial role in powering ships, driving the need for solutions that enhance operational efficiency while adhering to increasingly stringent environmental regulations [2]. The rising global concern over air pollution and climate change has placed significant emphasis on reducing harmful emissions, mainly nitrogen oxides (NOx) and particulate matter (PM), which are prevalent byproducts of CI engines [3, 4].

In response to these challenges, modern combustion strategies such as Reactivity-Controlled Compression Ignition (RCCI) have emerged, aiming to reduce NOx and PM emissions without compromising engine performance [5]. While these strategies offer promising solutions, the maritime industry faces unique operational challenges that demand precise emission control mechanisms, especially in large marine engines where traditional methods often fail to achieve desired environmental compliance [6]. To address these challenges, machine learning (ML) techniques are increasingly being applied to analyze and control emissions [7]. Among these techniques, the k-nearest neighbors (kNN) algorithm stands out as a promising method for classification. By leveraging historical engine data and operational parameters, kNN-based algorithms can offer accurate predictions of emission levels, allowing for better optimization of combustion processes and adherence to emission regulations.

Using natural gas (NG) and diesel fuels in the RCCI strategy promises remarkable combustion efficiency and minimizes emissions. The synergistic combination of natural gas's low reactivity and diesel fuel's high reactivity governs the combustion process [7]. During the compression stroke, a precise amount of natural gas is introduced into the combustion chamber, creating a uniform charge. This sets the stage for controlled ignition and combustion. Diesel fuel is injected at a carefully timed crank angle to ignite combustion. Implementing the RCCI strategy for marine engines holds significant promise for enhancing efficiency and reducing emissions in maritime transportation [8]. This hybrid approach enables the marine engine to operate with improved fuel efficiency and lower emissions than conventional combustion methods. In maritime operations, where fuel consumption and emissions are significant concerns, RCCI engines offer a compelling solution.

The classification problem in ML is encountered in many areas, including NOx emission prediction. For example, machine learning approaches have performed boiler NOx emission concentration prediction [9] and NOx emission predictions in gas turbines [10]. On the other hand, new machine learning approaches have been proposed using mathematical tools dealing with uncertainties, such as fuzzy sets [11], soft sets [12], intuitionistic fuzzy sets [13], picture fuzzy sets [14,15], and hybrid versions. Moreover, new mathematical tools struggling with further uncertainties, e.g., bipolar soft sets [16] and picture fuzzy soft sets [17,18], have been introduced. Although these mathematical tools are successful in modeling uncertainty, they are insufficient in solving problems with large amounts of data and uncertainty in the computer environment. Therefore, using picture fuzzy soft matrices (*pfs*-matrices) [19], fuzzy parameterized fuzzy soft matrices (*fpfs*-matrices) [20], and intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices (*ifpifs*-matrices) [21] – hybrid matrix representations of these mathematical tools – have become compulsory. Recently, the kNN algorithm [22], one of the well-known machine learning algorithms, has been improved by utilizing *fpfs*-matrices [23].

This new algorithm, FPFS-kNN, performs better than the state-of-the-art machine learning algorithms in many machine learning datasets.

In the present paper, we aim to explore the classification of NOx emissions in marine engines using kNN-based machine learning algorithms, including novel FPFS-kNN. Despite existing research studies, we are motivated to utilize state-of-the-art kNN

***Corresponding author:** Samet Memiş
*E-mail address: samettmemis@gmail.com
https://doi.org/10.56158/jpte.2024.97.3.02

algorithms to predict NOx emissions in marine engines operating under the RCCI strategy. The main contributions of the study are highlighted as follows:

- kNN-based classifiers predict the NOx emissions in marine engines.
- In contrast to various studies on NOx emission prediction problems, this paper uses soft matrices-based kNN algorithms to predict NOx emission.
- This research study is one of the pioneer studies predicting NOx emission in marine engines via kNN-based classification algorithms.

The ability to finely tune the combustion process allows for enhanced control over NOx and PM emissions, addressing environmental regulations while maintaining optimal performance [8,24]. The various working conditions of the RCCI engine were analyzed using the experimentally validated computational fluid dynamics (CFD) model. Taking into account the results, the present study classified the NOx emissions of RCCI engines fueled with natural gas/diesel utilizing the well-known and state-of-the-art kNN-based and/or soft-matrices-based machine learning algorithms, such as kNN [22], Fuzzy kNN [25], WkNN [26], PIFWkNN [27], ADAkNN [28], IFROWANN [29], LCkNN [30], GMkNN [31], LMRkNN [32], BM Fuzzy kNN [33], PFS-kNN [19], FPFS-EC [34], FPFS-CMC [35], FPFS-AC [36], IFPIFS-EC [37], IFPIFSC [38], and FPFS-kNN [23], and soft-matrices-based showcased in this study, obtaining crucial NOx emission parameters for the RCCI engine becomes readily predictable, eliminating the need for expensive and time-consuming experimental investigations.

2. Materials and Methods

This section first provides the dataset details concerning NOx emission in marine engines. After that, it details the kNN-based machine learning algorithms. Finally, it provides the performance metrics for classification problems.

2.1. Dataset Setting

A heavy-duty direct injection test engine was modeled. The specifications of the test engine are listed in Table 1. The start of injection [1] timing was changed between 32° to 56° bTDC to optimize the performance and emission for different configurations. The CFD model was performed using CONVERGE software [24]. The computational domain of the engine was designed. With varying operation parameters, the dataset was generated. The detailed model configurations and operation parameters can be found in the ref. [39,40].

Table 1. Caterpillar SCOTE 3401E test engine specifications.

The displacement per cylinder	2.4 liter
Bore × Stroke	137.6 mm
Stroke	165.1 mm
Connecting rod length	261.6 mm
Compression ratio	14.9:1
Engine speed	1300 r/min
Injection	Diesel + Natural gas
Direct injector spray angle	145 deg.

The dataset is converted to a three-class dataset according to NOx emission rates (NOxER) by utilizing the following conditions:

i. If $\mathrm{NOxER} \leq 0.141$, then the class of the case is 1

ii. If $0.141 < \mathrm{NOxER} \leq 0.421$, then the class of the case is 2

iii. If $\mathrm{NOxER} > 0.421$, then the class of the case 3

Table 2. Details of the dataset

	Input				Output	
Case No	**IPR (bar)**	**TFM (mg)**	**ER (%)**	**SOI (CA bTDC)**	**NOx (g/kWh)**	**NOx Class**
1	1.67	72.80	0	32	0.928	3
2	2.03	89.00	0	32	0.914	3
3	2.41	111.03	0	32	0.739	3
4	2.79	128.76	0	32	0.744	3
5	1.67	69.81	5	32	0.619	3
6	2.03	85.20	5	32	0.609	3
7	2.41	106.13	5	32	0.685	3
8	2.79	122.97	5	32	0.627	3
9	1.67	69.27	10	32	0.619	3
10	2.03	81.40	10	32	0.645	3
11	2.41	101.23	10	32	0.545	3
12	2.79	117.18	10	32	0.515	3
13	1.67	63.83	15	32	0.619	3
14	2.03	77.60	15	32	0.576	3
15	2.41	96.32	15	32	0.370	2
16	2.79	111.39	15	32	0.344	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
171	2.41	101.23	10	56	0.019	1
172	2.79	117.18	10	56	0.016	1
173	1.67	63.83	15	56	0.005	1
174	2.03	77.60	15	56	0.003	1
175	2.41	96.32	15	56	0.007	1
176	2.79	111.39	15	56	0.005	1
177	1.67	60.84	20	56	0.003	1
178	2.03	73.80	20	56	0.002	1
179	2.41	91.42	20	56	0.002	1
180	2.79	105.61	20	56	0.001	1

2.2. kNN-Based Machine Learning Algorithms

The kNN classifier is a prevalent classification technique owing to its simplicity and ease of implementation. Fuzzy kNN assigns a weight to each nearest neighbor based on a fuzzy membership degree, influencing the membership of the test sample. WkNN has been introduced to classify data utilizing a weighting factor dependent on and independent of the test sample. Nonetheless, WkNN necessitates the optimization of supplementary data-dependent parameters beyond k, which may overlook the influence of representative features within the data. The PIFW-kNN method has been introduced to resolve these problems.

The LCkNN is engineered for large datasets, partitioning the data through k-means clustering. ADAkNN has been introduced by addressing two concerns: selecting the number of nearest neighbors and optimizing the k value based on a given dataset.

Recent kNN-based classifiers, including the GMkNN, LMRkNN, and Fuzzy kNN classifier utilizing the BM-Fuzzy kNN, have been enhanced by applying diverse mean operators. GMkNN computes multi-local mean vectors for a specified query point within each class based on its class-specific k-nearest neighbors. In contrast, LMRkNN identifies the categorical k-nearest neighbors of a given query sample and assigns the class of the query sample through a linear combination of the categorical k-local mean vectors.

Thanks to redefined picture fuzzy soft matrices (*pfs*-matrices) because of some inconsistencies resulting from Cuong's definition of picture fuzzy sets, several distance measures of *pfs*-matrices are introduced. A new kNN-based classifier, namely the Picture Fuzzy Soft k-Nearest Neighbor (PFS-kNN) classifier, has been proposed, utilizing Minkowski's metric over *pfs*-matrices to find the k-nearest neighbor. A novel kNN-based classifier, FPFS-kNN, has gained prominence in recent years. Contemporary kNN classifiers determine the nearest neighbors utilizing a metric function. The nearest neighbors may vary if an alternative metric function is employed. The impact of the selected metrics on classification may differ depending on the datasets. FPFS-kNN evaluates the five pseudo-metrics over *fpfs*-matrices to address these limitations and computes the k-nearest neighbors for each pseudo-metric. Unlike conventional kNN classifiers, FPFS-kNN considers the influence of parameters (attributes) on classification and generates a feature weight vector for each dataset utilizing Pearson, Spearman, or Kendall correlation coefficients. In contrast to kNN-based machine learning algorithms employed *fpfs*-matrices, there are different approaches based on *fpfs*-matrices and intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices (*ifpifs*-matrices), such as FPFS-EC, FPFS-CMC, FPFS-AC, IFPIFS-EC, and IFPIFSC. Their working principles rely on distance and similarity measures of *fpfs*-matrices and *ifpifs*-matrices.

2.3. Performance Metrics

This subsection provides well-known performance metrics [41,42] such as accuracy (Acc) (Eq. 1), precision (Pre) (Eq. 2), recall (Rec) (Eq. 3), F1 score (F1) (Eq. 4) to compare the performance of ML algorithms. Their mathematical notations are as follows.

n samples to be classified $X = \{x_1, x_2, \dots, x_n\}$ let it be denoted as. $\mathbb{Y} = \{\mathbb{Y}_1, \mathbb{Y}_2, \dots, \mathbb{Y}_n\}$ the correct classes of these instances, $\hat{\mathbb{Y}} = \{\hat{\mathbb{Y}}_1, \hat{\mathbb{Y}}_2, \dots, \hat{\mathbb{Y}}_n\}$ let l denote the predicted classes of these samples and l the total number of classes.

The dataset herein contains three classes; therefore, for the class i, true positive (TP_i), true negative (TN_i), false positive (FP_i) and false negative (FN_i) values are defined in Eq. 5-8.

$$\mathrm{Acc}\ (\mathbb{Y}, \hat{\mathbb{Y}}) := \frac{1}{l}\sum_{i=1}^{l} \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \tag{1}$$

$$\mathrm{Pre}\ (\mathbb{Y}, \hat{\mathbb{Y}}) := \frac{1}{l}\sum_{i=1}^{l} \frac{TP_i}{TP_i + FP_i} \tag{2}$$

$$\mathrm{Rec}\ (\mathbb{Y}, \hat{\mathbb{Y}}) := \frac{1}{l}\sum_{i=1}^{l} \frac{TP_i}{TP_i + FN_i} \tag{3}$$

$$\mathrm{F1}\ (\mathbb{Y}, \hat{\mathbb{Y}}) := \frac{1}{l}\sum_{i=1}^{l} \frac{2TP_i}{2TP_i + FP_i + FN_i} \tag{4}$$

such that

$$TP_i := \left|\{x_t \mid i \in \mathbb{Y}_t \wedge i \in \hat{\mathbb{Y}}_t, 1 \le t \le n\}\right| \tag{5}$$

$$TN_i := \left|\{x_t \mid i \notin \mathbb{Y}_t \wedge i \notin \hat{\mathbb{Y}}_t, 1 \le t \le n\}\right| \tag{6}$$

$$FP_i := \left|\{x_t \mid i \notin \mathbb{Y}_t \wedge i \in \hat{\mathbb{Y}}_t, 1 \le t \le n\}\right| \tag{7}$$

$$FN_i := \left|\{x_t \mid i \in \mathbb{Y}_t \wedge i \notin \hat{\mathbb{Y}}_t, 1 \le t \le n\}\right| \tag{8}$$

3. NOx Classification Results and Discussion

The performance metrics of Acc, Pre, Rec, F1, and running time (RT) presented in Subsection 2.3 compare the classification outcomes of the machine learning algorithms discussed here. The simulation results are obtained using MATLAB R2018b software on a laptop with an Intel Core i7-13700H processor operating at 2.4 GHz and 32 GB of RAM. All classification algorithms underwent training and evaluation through 5-fold cross-validation [43]. The average performance results and runtimes for five iterations were documented in the 5-fold cross-validation. This procedure was subsequently executed ten times. The results for Acc, Pre, Rec, F1, and RT from ten iterations were acquired. The performance results derived from 5-fold cross-validation with ten repetitions (In a total of 50 runs) are displayed in Table 3.

Table 3. Comparison of the machine learning algorithms herein.

Algorithm	Acc	Pre	Rec	F1	RT
kNN	85.0370	76.5967	75.7321	75.5058	0.0432
Fuzzy kNN	85.7037	78.1775	76.9924	76.6984	0.0013
WkNN	84.0741	75.0823	73.9085	73.7963	0.0022
PIFWkNN	86.8889	82.4913	76.3576	75.4953	59.1177
ADAkNN	86.0000	79.5354	75.2102	76.1173	0.3018
IFROWANN	70.2963	53.6792	57.7436	55.3479	0.5839
LCkNN	79.7037	71.3209	63.7779	64.0482	0.0213
GMKNN	86.4444	79.2705	76.9456	77.0659	0.0039
LMRKNN	81.7778	72.0605	71.0445	70.4544	0.0033
BM-Fuzzy KNN	82.8889	73.1721	70.2706	70.4366	0.0051
PFS-kNN	77.7778	67.6419	69.4784	65.4102	0.0528
FPFSEC	89.1111	82.6937	81.2435	81.1894	0.0095
FPFSCMC	87.6296	80.4511	78.8561	79.0443	0.1031
FPFSAC	89.1111	82.6937	81.2435	81.1894	0.1078
IFPIFSEC	83.9259	74.1498	74.7646	73.6565	0.0396
IFPIFSC	84.0741	74.3125	75.0053	73.8688	0.1309
FPFS-kNN	90.0000	84.2272	82.3713	82.4680	0.0898

The best performance is shown in red.

Table 3 presents a comparative analysis of various kNN-based machine learning algorithms for classifying NOx emissions in marine engines. The algorithms are evaluated based on their Acc, Pre, Rec, F1, and RT, offering a detailed insight into each model's performance across these metrics.

Among the tested algorithms, FPFS-kNN achieves the highest Acc rate at 90.00%, with a Pre rate of 84.23%, Rec rate of 82.37%, and an F1 rate of 82.47%, indicating robust classification performance. This algorithm's high values in all performance metrics highlight its reliability in accurately distinguishing NOx emission levels, making it the most suitable model among the alternatives for this application. Furthermore, its relatively low runtime of 0.0898 seconds suggests that FPFS-kNN is both practical and efficient, contributing to its feasibility for real-time or near-real-time applications in a marine setting.

GMKNN and FPFSEC also perform well, with Acc rates of 86.44% and 89.11%, respectively. However, while these algorithms demonstrate strong classification capabilities, they fall short of FPFS-kNN's performance across multiple metrics. Specifically, GMKNN, though accurate, has slightly lower Pre and Rec rates than FPFS-kNN. Conversely, IFROWANN, with an Acc rate of 70.30%, exhibits a significantly lower performance, reflected in its low Pre, Rec, and F1 rates, indicating that this algorithm may not be as reliable for NOx emission classification tasks.

In terms of runtime, WkNN and GMKNN are the fastest, with RT of 0.0022 and 0.0039 seconds, respectively, but their lower F1 rates (73.80% for WkNN and 77.07% for GMKNN) suggest a trade-off between speed and classification effectiveness. On the other hand, algorithms like PIFWkNN, with a notably high RT of 59.1177 seconds, demonstrate that high RT does not necessarily correlate with better classification performance, as PIFWkNN's F1 rate (75.50%) remains below that of FPFS-kNN.

Overall, the results in Table 3 highlight FPFS-kNN as the optimal algorithm for NOx emission classification, given its high Acc, Pre, Rec, F1, and manageable RT balance. This balance of effectiveness and efficiency underscores FPFS-kNN's potential utility for practical applications in monitoring and managing NOx emissions in marine engines.

Fig. 1 visually compares the performance metrics (accuracy, precision, recall, F1 score, and runtime) for various kNN-based algorithms in classifying NOx emissions in marine engines. This graphical analysis complements the detailed tabular results in Table 3 by offering an intuitive understanding of the algorithms' strengths and trade-offs.

As highlighted in Table 3, FPFS-kNN achieves the highest values across performance metrics: accuracy (90.00%), precision (84.23%), recall (82.37%), and F1 score (82.47%). In Fig. 1, these metrics are visually prominent, with FPFS-kNN consistently outperforming other algorithms, underscoring its robustness and classification reliability. Its runtime (RT) of 0.0898 seconds, while not the fastest, remains within a practical range, suggesting a favorable balance between effectiveness and efficiency (Fig. 1).

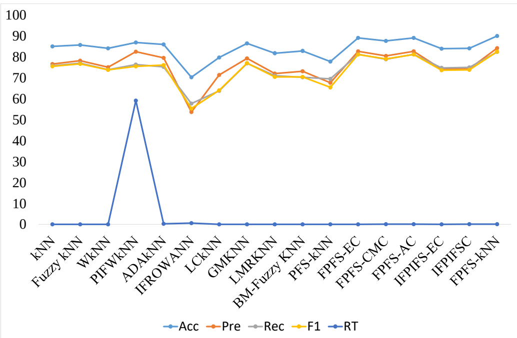

Fig. 1. Visual comparison of the performance results of the algorithms

Other notable performers include GMKNN and FPFSEC, which achieve competitive accuracy rates of 86.44% and 89.11%, respectively. However, as Fig. 1 illustrates, their precision, recall, and F1 scores fall short of FPFS-kNN's values, indicating slightly less balanced classification performance. Additionally, GMKNN boasts a notably low runtime (0.0039 seconds), making it a strong candidate for scenarios where computational speed is critical. However, this comes with a trade-off in overall performance (Fig. 1).

Conversely, IFROWANN demonstrates a significantly lower performance, with accuracy, precision, recall, and F1 score markedly below the leading algorithms, as reflected in both Table 3 and Fig. 1. This result suggests that IFROWANN may not be suitable for NOx emission classification tasks. Similarly, PIFWkNN exhibits the highest runtime (59.1177 seconds) without a corresponding improvement in classification metrics, highlighting inefficiency (Fig. 1).

Overall, Fig. 1 reinforces the conclusion that FPFS-kNN is the optimal algorithm for NOx emission classification, offering superior accuracy and balanced performance across all metrics while maintaining reasonable computational efficiency. These insights, in conjunction with Table 3, affirm FPFS-kNN's potential for real-time or near-real-time applications in marine settings.

The box-plot analysis provides detailed evaluations of the classification performance of various kNN-based algorithms on NOx emission prediction in marine diesel engines. Regarding accuracy (Fig. 2a), FPFS-kNN demonstrates superior performance with the highest median value, indicating consistently strong predictions. Its narrow interquartile range reflects remarkable stability compared to other algorithms, such as ADAkNN and IFROWANN, which exhibit wider spreads and lower medians, suggesting more significant prediction variability. Outliers in particular algorithms, including LCkNN and BM Fuzzy kNN, reveal scenarios of underperformance that may undermine their reliability in practical applications.

Similarly, the F1 score analysis (Fig. 2b) underscores the robust classification capability of FPFS-kNN, as it achieves high median and compact variability, maintaining a delicate balance between precision and recall. In contrast, algorithms such as FPFS-CMC and WkNN display lower medians and broader spreads, indicating challenges in achieving consistent performance across scenarios. Notably, ADAkNN and IFROWANN perform competitively in the F1 score, though their broader distributions highlight inconsistencies under varying conditions.

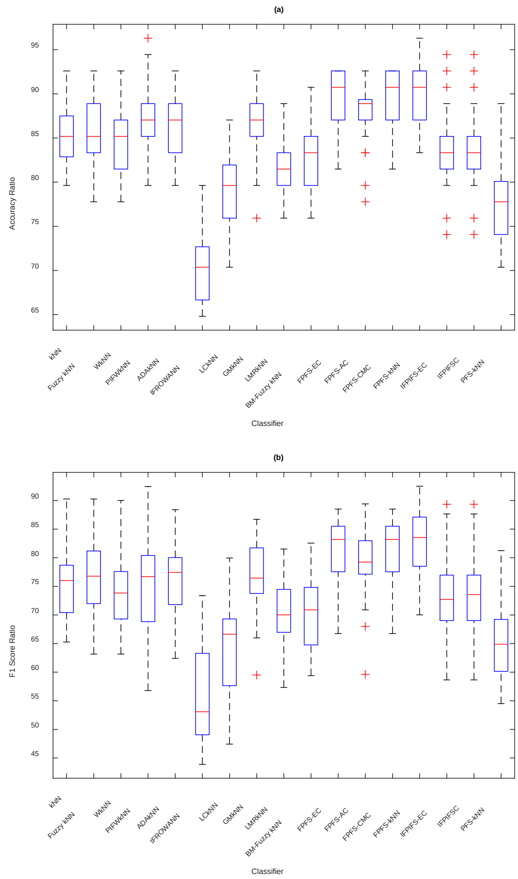

Fig. 2. Box-plot analysis of the performance result derived from 50 runs, i.e., 5-fold cross-validation and ten runs.

FPFS-kNN emerges as the most robust and reliable algorithm, excelling across accuracy (Fig. 2a) and F1 score (Fig. 2b) metrics. However, the comparative analysis also highlights the potential strengths of algorithms like ADAkNN and IFROWANN, which, despite some variability, show promise in specific scenarios. These findings suggest FPFS-kNN as the leading candidate for NOx emission prediction while emphasizing the importance of further refinement for competing algorithms to enhance their stability and generalizability.

4. Conclusion

The principal objective of contemporary engines is to attain optimal performance while complying with progressively rigorous emission regulations. The dual-fueled RCCI engine is distinguished by its exceptional capacity to produce minimal soot and NOx emissions while achieving high thermal efficiency. This study demonstrates that the meticulous adjustment of critical engine parameters, including SOI, EGR, fuel mass, and intake pressure, is essential for regulating exhaust emissions. RCCI engines are adaptable to function under diverse loads and conditions, presenting a promising pathway to attain environmental sustainability and engine efficiency across various applications.

Although some studies have presented a systematic approach to increase further engine performance and various methods, such as soft decision-making [28], have been used to determine the optimum parameters, it is crucial to be able to estimate NOx emission values according to the optimum input parameters determined before an experiment is performed. To this end, this study presents a systematic approach for estimating NOx emissions in ship engine engines using state-of-the-art machine learning algorithms. Simulation results show that the most suitable model for estimating NOx emissions in ship engine engines is the FPFS-kNN approach.

This study provides significant validation for the potential of FPFS-kNN to be successfully applied in different emission estimation scenarios. While the results demonstrate robust performance, further enhancements can be pursued by exploring advanced fuzzy hybrid models [44]. Future studies can investigate the integration of FPFS-kNN with complex picture fuzzy sets [45], bipolar complex fuzzy sets [46], bipolar soft sets [16], interval-valued intuitionistic fuzzy parameterized interval-valued intuitionistic fuzzy soft sets/matrices [47,48], and picture fuzzy parameterized picture fuzzy soft sets [49]. Such extensions may offer improved classification accuracy and flexibility in handling uncertainties and multi-dimensional data, enhancing the model's applicability in more complex and diverse scenarios. These directions represent promising avenues for advancing the current research and addressing the open challenges in emission estimation.

Declaration of conflicting interests

The authors declare no competing interests.

Funding

The author received no financial support for the research and/or authorship of this article.

Abbreviations

Abbreviation	Description
NOx	Nitrogen Oxides
NOxER	NOx emission rates
CI	Compression Ignition
RCCI	Reactivity-Controlled Compression Ignition
PM	Particulate Matter
ML	Machine Learning
NG	Natural Gas
CFD	Computational Fluid Dynamics
SOI	Start of Injection
kNN	k-Nearest Neighbor
Fuzzy kNN	Fuzzy kNN
WkNN	Weighted kNN
PIFWkNN	Parameter Independent Fuzzy Weighted kNN
ADAkNN	Adaptive kNN
IFROWANN	Imbalanced Fuzzy-Rough Ordered Weighted Average Nearest Neighbor
LCkNN	Locality Constrained Representation-Based kNN
GMKNN	Generalized Mean Distance-Based kNN
LMRKNN	Local Mean Representation-Based kNN
BM-Fuzzy KNN	Fuzzy kNN Based on the Bonferroni Mean
***fpfs*-matrices**	Fuzzy Parameterized Fuzzy Soft Matrices
***pfs*-matrices**	Picture Fuzzy Soft Matrices
***ifpifs*-matrices**	Intuitionistic Fuzzy Parameterized Intuitionistic Fuzzy Soft Matrices
PFS-kNN	Picture Fuzzy Soft kNN
FPFS-EC	Euclidean Distance-Based Fuzzy Parameterized Fuzzy Soft Classifier
FPFS-CMC	Compare Matrix-Based Fuzzy Parameterized Fuzzy Soft Classifier
FPFS-AC	Aggregation Operator-Based Fuzzy Parameterized Fuzzy Soft Classifier
IFPIFS-EC	Euclidean Distance-Based Intuitionistic Fuzzy Parameterized Intuitionistic Fuzzy Soft Classifier
IFPIFS-C	Intuitionistic Fuzzy Parameterized Intuitionistic Fuzzy Soft Classifier
FPFS-kNN	Fuzzy Parameterized Fuzzy Soft kNN

References

[1] Sener, R., Yangaz, M. U., and Gül, M. Z., 2020. *Effects of injection strategy and combustion chamber modification on a single-cylinder diesel engine*, Fuel, 266, 117122,

[2] Ni, P., Wang, X., and Li, H., 2020. *A review on regulations, current status, effects and reduction strategies of emissions for marine diesel engines*, Fuel, 279(118477).

[3] Fayyazbakhsh, A., Bell, M. L., Zhu, X., Mei, X., Koutný, M., Hajinajaf, N., et al., 2022. *Engine emissions with air pollutants and greenhouse gases and their control technologies*, J. Clean. Prod., 376(134260).

[4] Şener, R., and Gül, M. Z., 2021.*Optimization of the combustion chamber geometry and injection parameters on a light-duty diesel engine for emission minimization using multi-objective genetic algorithm*, Fuel, 304(121379).

[5] Çelik, T., and Şener, R., 2023.*Optimization of a biogas fueled reactivity-controlled compression ignition engine using NSGA-II algorithm*., J. Eng. Mach.,64(713), pp.670–695.

[6] Gössling, S., Meyer-Habighorst, C., and Humpe, A., 2021. *A global review of marine air pollution policies, their scope and effectiveness*, Ocean Coast. Manag.,212, pp.105824.

[7] Yüksel, O., Bayraktar, M., and Sokukcu, M., 2023. *Comparative study of machine learning techniques to predict fuel consumption of a marine diesel engine*, Ocean Eng., 286, pp.115505.

[8] Xin, M., Gan, H., Cong, Y., and Wang, H., 2024. *Numerical simulation of methane slip from marine dual-fuel engine based on hydrogen-blended natural gas strategy*, Fuel, 358, pp.130132.

[9] Tang, Z., Wang, S., Chai, X., Cao, S., Ouyang, T., and Li, Y., 2022. *Auto-encoder-extreme learning machine model for boiler NOx emission concentration prediction*, Energy, 256, pp. 124552.

[10] Hoque, K. E., Hossain, T., Haque, A. M., Miah, M. A. K., and Haque, M. A., 2024. *NOx emission predictions in gas turbines through integrated data-driven machine learning approaches*, J. Energy Resour. Technol.,146(7), pp. 071201.

[11] Zadeh, L. A., 1965. *Fuzzy sets*, Inf. Control, 8, pp. 338–353.

[12] Molodtsov, D., 1999. *Soft set theory-first results*, Comput. Math. Appl., 37(4–5), pp. 19–31.

[13] Atanassov, K. T.,1986. *Intuitionistic fuzzy sets*, Fuzzy Sets Syst., 20(2), pp. 87–96.

[14] Cuong, B. C., 2014. *Picture fuzzy sets*, J. Comput. Sci. Cybern. 30(4), pp. 409–420.

[15] Memiş, S., 2021. *A study on picture fuzzy sets*, 7th IFS Contemporary Mathematics and Engineering Conference (IFSCOM'2021), Mersin University Press, Turkey, pp. 125–132.

[16] Mahmood, T., 2020. *A novel approach towards bipolar soft sets and their applications*, J. Math., 2020(2), pp. 4690808.

[17] Yang, Y., Liang, C., Ji, S., Liu, T., 2015. *Adjustable soft discernibility matrix based on picture fuzzy soft sets and its applications in decision making*, J. Intell. Fuzzy Syst., 29(4), pp. 1711–1722.

[18] Memiş, S.,2022. *Another view on picture fuzzy soft sets and their product operations with soft decision-making*, J. New Theory, (38), pp. 1-13.

[19] Memiş, S., 2023. *Picture fuzzy soft matrices and application of their distance measures to supervised learning: Picture fuzzy soft k-nearest neighbor (PFS-kNN)*, Electronics, 12(19), pp. 4129.

[20] Enginoğlu, S. and Çağman, N.,2020. *Fuzzy parameterized fuzzy soft matrices and their application in decision-making*, TWMS J. Appl. Eng. Math.,10(4), pp. 1105–1115.

[21] Enginoğlu, S. and Arslan, B.,2020. *Intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices and their application in decision making*, Comput. Appl. Math., 39(325).

[22] Cover, T. M. and Hart, P. E., 1967. *Nearest neighbor pattern classification*, IEEE Trans. Inf. Theory,13, pp. 21–27.

[23] Memiş, S., Enginoğlu, S., and Erkan, U., 2022. *Fuzzy parameterized fuzzy soft k-nearest neighbor classifier*, Neurocomputing, 500, pp. 351–378.

[24] Richards, K. J., Senecal, P. K., and Pomraning, E., 2019. CONVERGE 2.4 Manual, 1078.

[25] Keller, J. M., Gray, M. R. and Givens, J. A., 1985. *A fuzzy k-nearest neighbor algorithm*, IEEE Trans. Syst. Man Cybern., 15, pp. 580–585.

[26] Dubey, H., and Pudi, V., 2013. *Class based weighted k-nearest neighbor over imbalance dataset*, 17th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, Gold Coast, Australia, pp. 305–316.

[27] Biswas, N., Chakraborty, S., Mullick, S. S., and Das, S., 2018. *A parameter independent fuzzy weighted k-nearest neighbor classifier*, Pattern Recognit. Lett., 101, pp. 80–87.

[28] Mullick, S. S., Datta, S., and Das, S., 2018. *Adaptive learning-based k-nearest neighbor classifiers with resilience to class imbalance*, IEEE Trans. Neural Netw. Learn. Syst., 29, pp. 5713–5724.

[29] Ramentol, E., Vluymans, S., Verbiest, N., Caballero, Y., Bello, R., Cornelis, C., and Herrera, F., 2015. *IFROWANN: Imbalanced fuzzy-rough ordered weighted average nearest neighbor classification*, IEEE Trans. Fuzzy Syst., 23, pp. 1622–1636.

[30] Gou, J., Qiu, W., Yi, Z., Shen, X., Zhan, Y., and Ou, W., 2019. *Locality constrained representation-based K-nearest neighbor classification*. Knowl.-Based Syst., 167,pp. 38–52.

[31] Gou, J., Ma, H., Ou, W., Zheng, S., Rao, Y., and Yang, H., 2019. *A generalized mean distancebased k-nearest neighbor classifier*, Expert Syst. Appl., 115, pp. 356–372.

[32] Gou, J., Qiu, W., Yi, Z., Xu, Y., Mao, Q. and Zhan, Y., 2019. *A local mean representation-based k-nearest neighbor classifier*, ACM Trans. Intell. Syst. Technol.,10(3), pp. 29.

[33] Kumbure, M. M., Luukka, P., and Collan, M., 2020. *A new fuzzy k -nearest neighbor classifier based on the Bonferroni mean*, Pattern Recognit. Lett.,140, pp. 172–178.

[34] Memiş, S., Enginoğlu, S., and Erkan, U., 2021.*Numerical data classification via distance-based similarity measures of fuzzy parameterized fuzzy soft matrices*, IEEE Access, 9, pp. 88583–88601.

[35] Memiş, S., Enginoğlu, S., and Erkan, U., 2022. *A classification method in machine learning based on soft decision-making via fuzzy parameterized fuzzy soft matrices*, Soft Computing, 26(3), pp. 1165–1180.

[36] Memiş, S., Enginoğlu, S., and Erkan, U., 2022. *A new classification method using soft decision-making based on an aggregation operator of fuzzy parameterized fuzzy soft matrices*, Turk. J. Electr. Eng. Comput. Sci., 30(3), pp. 871–890.

[37] Memiş, S., Arslan, B., Aydın, T., Enginoğlu, S., and Camcı, Ç., 2021. *A classification method based on hamming pseudo-similarity of intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices*, J. New Results Sci., 10(2), pp.59–76.

[38] Memiş, S., Arslan, B., Aydın, T., Enginoğlu, S., and Camcı, Ç., 2023. *Distance and similarity measures of intuitionistic fuzzy parameterized intuitionistic fuzzy soft matrices and their applications to data classification in supervised learning*, Axioms, 12(5), pp. 463.

[39] Koç, M. A., and Şener, R.,2021. *Prediction of emission and performance characteristics of reactivity-controlled compression ignition engine with the intelligent software based on adaptive neural-fuzzy and neural-network*. J. Clean. Prod., 318, pp. 128642.

[40] Şener, R., Memiş, S., and Kılıç, A., 2024. *Investigation of Reactivity Controlled Compression Ignition Fueled with Diesel and Natural Gas in Marine Engine using Soft Decision-Making Methods*, Global Maritime Congress (GMC'24): ITU Press, pp. 125–130.

[41] Fawcett, T., 2006. *An introduction to ROC analysis*, Pattern Recognit. Lett., 27, pp. 861–874.

[42] Erkan, U., *A.*, 2021. *precise and stable machine learning algorithm: Eigenvalue classification (EigenClass)*, Neural Comput. Appl., 33, pp. 5381–5392.

[43] Stone, M., 1974. *Cross-validatory choice and assessment of statistical predictions*, J. R. Stat. Soc. Ser. B (Methodol.), 36, pp. 111–147.

[44] Tadayoni, B., Amiri, E., and Soorani, F.,2023. *Knowledge in medical prescriptions with the help of genetic fuzzy hybrid model*. Int. J. Pioneering Technol. Eng., 2(2), pp. 142–146.

[45] Ozer, O., 2022. *Hamacher prioritized aggregation operators based on complex picture fuzzy sets and their applications in decision-making problems*, J. Innov. Res. Math. Comput. Sci., 1(2), pp. 33–54.

[46] Jaleel, A.,2022. *WASPAS technique utilized for agricultural robotics system based on Dombi aggregation operators under bipolar complex fuzzy soft information*, J. Innov. Res. Math. Comput. Sci., 1(2), pp. 67–95.

[47] Aydın, T., and Enginoğlu, S., 2021. *Interval-valued intuitionistic fuzzy parameterized interval-valued intuitionistic fuzzy soft sets and their application in decision-making*. J. Ambient Intell. Humanized Comput., 12(2), pp. 1541–1558.

[48] Aydın, T., and Enginoğlu, S., 2022. *Interval-valued intuitionistic fuzzy parameterized interval-valued intuitionistic fuzzy soft matrices and their application to performance-based value assignment to noise-removal filters*. Comput. Appl. Math., 41(4), pp.192.

[49] Memiş, S.,2023. *Picture fuzzy parameterized picture fuzzy soft sets and their application in a performance-based value assignment problem to salt-and-pepper noise removal filters*. Int. J. Fuzzy Syst., 25(7), pp. 2860–2875.