Evolved model for early fault detection and health tracking in marine diesel engine by means of machine learning techniques

Tolga Şahin[*1], C. Erdem Imrak[1], Altan Cakir[2], Adem Candaş[1]

[1] Department of Mechanical Engineering, Faculty of Mechanical Engineering, Istanbul Technical University, Inönü St. No:65 Gümüşsuyu/Beyoğlu, Istanbul 34437, Turkey, e-mail: sahintolga@itu.edu.tr
[2] Physics Engineering, Faculty of Science and Letters, Istanbul Technical University, Katar St. B4-207 Reşitpaşa /Maslak/Sarıyer, Istanbul 34469, Turkey
[*] Corresponding author

ABSTRACT

The Coast Guard Command, which has a wide range of duties as saving human lives, protecting natural resources, preventing marine pollution and battle against smuggling, uses diesel main engines in its ships, as in other military and commercial ships. It is critical that the main engines operate smoothly at all times so that they can respond quickly while performing their duties, thus enabling fast and early detection of faults and preventing failures that are costly or take longer to repair. The aim of this study is to create and to develop a model based on current data, to select machine learning algorithms and ensemble methods, to develop and explain the most appropriate model for fast and accurate detection of malfunctions that may occur in 4-stroke high-speed diesel engines. Thus, it is aimed to be an exemplary study for a data-based decision support mechanism.

ARTICLE INFO

1 Introduction

Control, safety and reliability of the engines are very important in addition to high power generation, low fuel consumption and low maintenance/repair costs. MTU 16V 4000 M90 model main diesel engines are in the inventory of MRTP-33 type ships of the Turkish Coast Guard Command especially for sudden and rapid interventions with the advantages of low volume and high power. As in many ships, faults of different severity occur also at the main diesel engines operated in MRTP-33 type ships due to failure to detect fatal changes exceeding the limits in operational values accurate and timely. Innovative and developing information technology applications such as statistical learning techniques based on machine/deep learning to facilitate the detection and diagnosis of such malfunctions of engines.

In recent years, different measurement methods have been used to monitor the status of marine diesel engines and to detect faults. Failure modeling studies that deal with acoustic emission, cylinder pressure, vibration, angular velocity and multiple measurements of internal combustion engines are available in the literature. Chandroth et al. [1–2] investigated the cylinder pressures and vibration of the internal combustion engine. In order to increase the efficiency of the engines and reduce the maintenance costs, the engine vibration and acoustic emission data were examined and a machine learning model was created by Yasir [3]. In another study, Vladimir [4] investigated monitoring of the state of rotary motion machines and fault detection by using vibration data. Zhixiong et al. [5–6] investigated the fault detection of ship diesel engines using instantaneous angular velocity data. Tsaganos et al. [7] investigated the fault detection and diagnosis of a two-stroke low-speed ship diesel engine by using machine learning method with pressure and temperature data. Mohd Noor et al. [8] created and analyzed the performance model of a ship diesel engine with artificial neural networks.

These methods are very effective to evaluate the working condition of the diesel engine. Multiple measurements, such as examining the pressure and temperature values obtained from different sensors together, provide results in the evaluation of the working condition of the diesel engine. The use of an information technology-based system

Table 1 Fault-reason relations of marine diesel engine common system faults

Problem	Primary Source
Wear/Fatigue/Corrosion	Life Cycle
Contamination/Clogging	User/Maintainer
Stress induced crack/Fracture	User/Maintainer
Changed dimensions after wear or maintenance	Life Cycle/User/Maintainer
Permanent deformations in system components	Life Cycle/User/Maintainer
Not using suitable fuel, water, oil or losing its properties	User/Maintainer

Source: Authors

together with multiple measurements in the diagnosis of diesel engines contributes to the elimination of external effects such as human error and to increase safety and reliability. In addition, timely detection of faults prevents the occurrence of larger faults with major damage and high cost and ensures operational continuity. Today, the use of electronic systems and information technologies in fast and accurate fault detection of diesel engines is increasingly desired, especially in terms of safety, reliability and ease of control.

Since various abrasions and malfunctions may occur in the system components of an engine depending on the operating time, it is significant to detect malfunctions especially in the lubrication, cooling, fuel/exhaust systems of the engine. In addition to the fact that the malfunctions of these systems are caused by wear or the loss of properties of liquids (diesel fuel, cooling water, oil, etc.), due to different reasons may also cause malfunctions that will affect the normal operation of the engine. Such unexpected malfunctions of engines may cause undesirable effects that may harm the environment and/or people, as well as cause the engine to operate with low efficiency or to prevent operational activities.

An engine failure can be defined as being out of limit of the acceptable operating values of certain functional parameters. In this context, damage due to malfunction can also be called as an irregularity or symptom that occurs during the working process of the engine and affects the present and the future negatively. The detection of these irregularities or symptoms can be achieved by comparing the parameters determined in the design of the system. System failures in the main diesel engine used on a ship due to the reasons shown in Table 1 have a direct impact on the life of the diesel engine itself and the system elements. It is possible to detect the malfunctions indicated in Table 1 by examining the obvious deviations in some basic data of the engine, especially in temperature and pressure data.

System failures that occur in ship diesel engines can be caused by the elements of that system or sometimes by an element belonging to a different system. This situation makes it difficult to find the root cause of the malfunction of the engine. For example, the reason for the high exhaust temperature of the engine may be the failure of an element in the combustion/exhaust system, or the failure of an element in the cooling system which is not fulfill its duty. In this paper, the root causes of ship diesel engine system failures were systematically investigated.

In this study, data were collected through various sensors of MTU brand engine in order to systematically determine the status of the diesel engine running on the MRTP-33 class ship. These are the processes in which cooling system failure, lubrication system failure, combustion/exhaust system failure, pressure and temperature values are measured within the scope of 16V 4000 M90 model main engine sensors. The collected data was transferred to more than one machine learning algorithm and prediction results were obtained on the basis of accuracy, which is one of the machine learning performance criteria. By developing and improving the obtained results with the ensemble methods, an optimum machine learning model was developed for the detection of engine system failures, including the F-measure (F1) and the accuracy performance metrics and the models' build time.

2 Materials and methods

2.1 Machine learning and decision-support system

The general name given to computer algorithms that models a situation using historical data can be defined as machine learning. The aim here is to create the model that will give the highest performance by using the available data sets and computer algorithms.

Pitts and McCulloch's creation of the first mathematical model of a neural network in 1943 forms the basis of machine learning [9]. In 1950, Turing [10] questioned whether machines could use decision making and problem-solving skills by using existing knowledge in addition to logic, and the "Turing Test" emerged. The term Machine Learning was coined to describe the pattern recognition tasks that provide the "learning" component in artificial intelligence systems [11]. In 1997, Mitchell [12] defined machine learning as "Training a computer software with a measure of performance P to perform a desired task G using experience D.".

Statistical learning plays a key role in solving problems in many fields, especially in science, finance and industry [13]. Machine learning algorithms are widely applied in the studies of classification of neuromuscular diseases, brain tumors, colon cancer patients and healthy people in the field of health, in the studies of estimating the stock

Table 2 Alarm limits according to the design parameters of MTU 16V 4000 M90 marine diesel engine

Model Features	Engine Parameters	Alarm Limits
T_Coolant	Coolant Temperature (^{0}C)	95
P_Coolant	Coolant Pressure (Bar)	Figure 1
P_Raw_Water	Raw Water Pressure (Bar)	Figure 1
T_Oil	Lubricating Oil Temperature (^{0}C)	92
P_Oil	Lubricating Oil Pressure (Bar)	Figure 1
P_Fuel_After_Filter	Fuel Pressure After Filter (Bar)	Figure 1
T_Exhaust_A/B_Side	A/B Bank Exhaust Temperature (Mean) ^{0}C	750
Turbo_Charger_1/2 _Speed	Turbocharger Speed (rpm/1000)	55.8
T_Cylinder_Exhaust	Cylinder Exhaust Temperature 1-16 (^{0}C)	Figure 2

Source: Authors

price in the field of economics/finance and in the studies of measuring the innovations and development of companies, and also in the use of control systems and robotics in engineering problems [14–20]. In addition, there are machine learning applications in the creation and development of decision support systems in many areas [21, 22].

There are three main approaches to machine learning: Supervised Learning, Unsupervised Learning, and Reinforcement Learning. On the other hand, supervised learning is separated into two types as classification and regression.

In this study, the data collected for solving the multiclass classification problem with the supervised machine learning approach to be used in the modeling of a diesel engine will be evaluated in four classes as normal operating condition, cooling system failure, combustion/exhaust system failure, lubrication system failure. The multiclass classification was made by considering the fault logs of the MRTP-33 class ship in the inventory of the Turkish Coast Guard Command and the significant changes in the data collected from the MTU 16V 4000 M90 model main diesel engine operating on the ship, and the design and operating limits of the main diesel engine.

2.2 Fault diagnosis of four-stroke marine diesel engine

In order to meet low volume and high-speed needs at MRTP-33 class ships, direct injection, four-stroke with four valves in each cylinder, 2720 KW nominal power, common-rail injection electronically controlled, without full load and low load usage limits MTU 16V 4000 M90 diesel main engines are used [23, 24].

2.2.1 Model process based on engine design parameters

Ship diesel engines are equipped with electronic control and display systems, and according to this, with various sensors that can monitor engine operating values. By use of these sensors, problems that may occur in the future, such as costly breakdowns and prolonged downtime of ships can be prevented. Basically, temperature and pres-

sure parameters are the most studied and paid attention in the design and operation of ship machinery. It is absolutely necessary to observe and control the sudden and significant changes that may occur in these parameters during the operation of the engines.

Fault diagnosis can be basically divided into three different stages, which consist of measuring the structural and operating parameters, collecting data for comparison with previous data from the same engine, identifying the problems causing these faults and their causes [25]. The sequential completion of these stages allows for successful fault diagnosis.

The design/operation parameter data of the examined MTU 16V 4000 M90 model engine, as well as the alarm limits according to the design parameters for the cooling water temperature, lubricating oil temperature, A and B bank exhaust temperatures, and number 1 and 2 turbochargers are shown in Table 2.
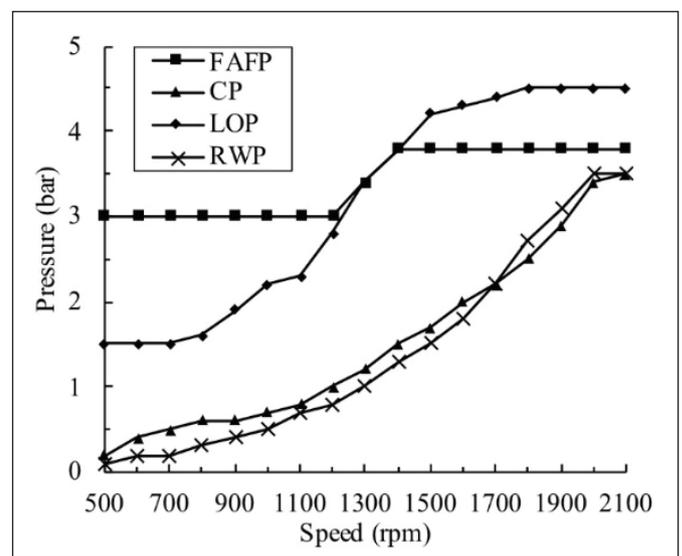

Figure 1 Change of pressure alarm limits depending on engine speed (FAFP: Fuel pressure after filter, CP: Cooling pressure, LOP: Lube-oil pressure, RWP: Raw water pressure)
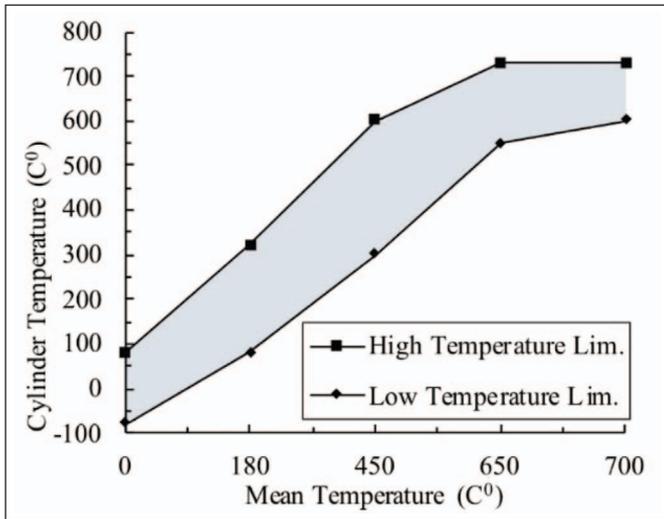
Source: Authors

Figure 2 Change of cylinder temperature alarm limits depending on average temperature for each cylinder

Source: Authors

The pressure curves of cooling water, sea water, lubricating oil and fuel, that alarm limits change according to the engine speed, are shown in Figure 1. According to the design parameters of the engine, in normal operating condition, the pressure values are expected to be above the curve of the relevant parameter. The high and low temperature curves of each cylinder, that alarm limits change according to the average exhaust temperature of the 16 cylinders, are shown in Figure 2. In the normal operating condition of the engine, the exhaust temperature of each cylinder is expected to be within the shaded area. High exhaust temperature alarms occur if the shaded area is exceeded, and low exhaust temperature alarms occur if it goes below.

2.2.2 Experimental process

In the data acquisition process of MTU 16V 4000M90 diesel main engine, PT 1000 type with measuring range of -40 to 150 °C and 1000-1385 ohms for cooling water temperature and lubricating oil temperature, PT 100 type with measuring range -40 to 900 °C and 100-408 ohms for A and B bank exhaust temperature, K type thermocouples with measuring range 0-850 °C for the exhaust temperature of each of the 16 cylinders, inductive type with a measuring range of 50-100000 rpm for the speed of turbochargers 1 and 2, measuring range 0 to 6 bar and 0.5-4.5 volts for cooling water pressure and seawater pressure, measuring range 0 to 10 bar and 0.5-4.5 volts for lubricating oil pressure, piezoelectric type with measuring range 0 to 15 bar and 0.5-4.5 volts for low fuel pressure, sensors are used. Coolant temperature sensor is located at the fresh water cooler inlet, lubricating oil temperature sensor is located at the oil cooler inlet, A and B bank exhaust temperature sensors are located at the exhaust manifold outlet of both sides, thermocouples used in the exhaust temperature measurement of the cylinders are located on the upper side of each cylinder, turbocharger speed sensors are located on the top of each turbocharg-
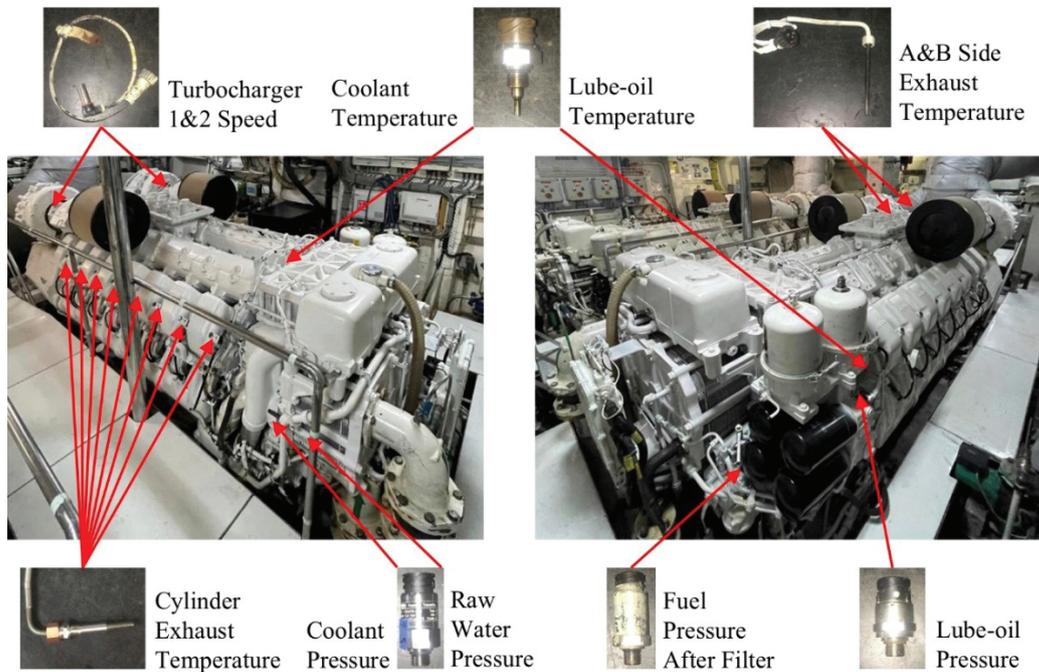

Figure 3 Locations of sensors on four-stroke marine diesel engine

Source: Authors

ers. The cooling water pressure sensor is located at the fresh water pump outlet, the seawater pressure sensor is located at the seawater pump outlet, the lubricating oil pressure sensor is located at the lubricating oil pump outlet, and the low fuel pressure sensor is located at the fuel filter outlet. The positions of the sensors used in the measurements on the engine are shown in Figure 3.

It is possible to monitor online the data recorded from the sensors on the 16V 4000 M90 model diesel main engine operating on the ship TCSG-312, which is in the inventory of the Turkish Coast Guard Command, by transferring it to the computer via the MTU DIASYS interface over the MTU MCS-5 (control and display system). The machine fault tracking system used in the developed data collection process is shown in Figure 4.

Connecting to the MTU MCS-5 system is achieved by joining the ECU 4 unit on the machine and the laptop via RS232 cable. The data transmission cycle time between individual data blocks is 70 ms [26] and the data sampling time is 440 ms.

After data collection and manipulation, an open source PYTHON program and PYCARET library which allows to use different PYTHON machine learning libraries together, were used for analysis. The flow chart of the system developed for the creation of a model by processing the data collected from the sensors coupled to the diesel engine is presented in Figure 5.

The data set consists of 3237 samples containing all data classes. 10% of the data set (324 samples) is reserved as unseen data in order to make predictions based on the analysis results. The remaining 90% (2913 samples) was used in modeling. The data used in the modeling were divided into training (75%) and test (25%) data. 'Stratified
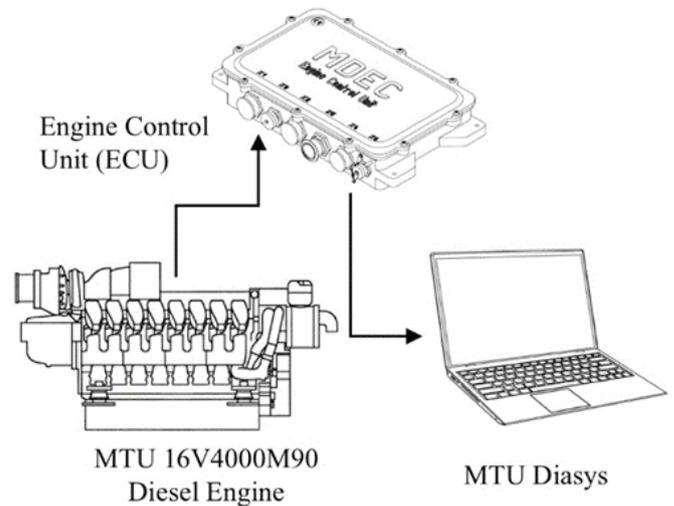

Figure 4 Engine fault tracking system

Source: Authors

KFold' was used as a cross validator. The number of layers of the cross validator was chosen as 10. In the model setup, the 'SMOTE Fix Imbalance' method was used because the sample numbers of the classes were not equal and 'Z-Score Normalize' method was used to standardize numerical data. After the model setup was completed, 13 most commonly used algorithms for classification problems (Light Gradient Boosting Machine, Random Forest Classifier, Gradient Boosting Classifier, Extra Trees Classifier, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Decision Tree Classifier, K Neighbors Classifier,
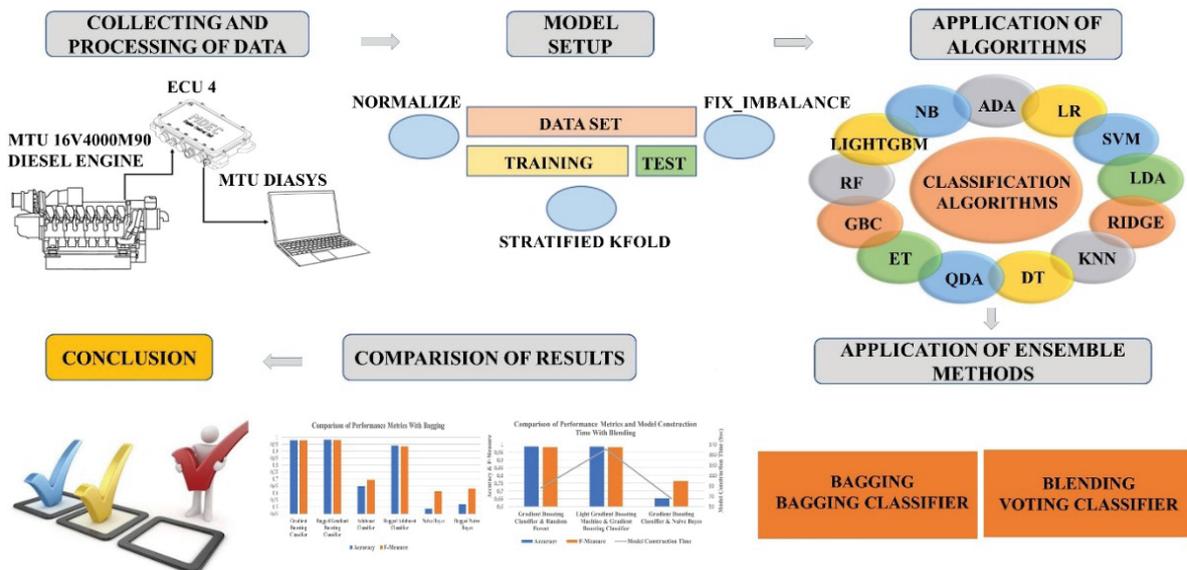

Figure 5 Flow chart of system architecture

Source: Authors

Ridge Classifier, SVM-Linear Kernel, Logistic Regression, Ada Boost Classifier, Naive Bayes) was applied to the developed model and the results were obtained.

As a result of calculating the model with the determined algorithms, performance metrics (Accuracy, F-Measure (F1)) and processing time were obtained, and they were listed on the basis of 'Accuracy'. Then, the improvement of the results had been achieved by ensemble methods.

For the quality and accuracy of the classification, the performance metrics 'Accuracy' and 'F-Measure' are important criteria. Although model evaluation does not consider accuracy alone, it is the prime metric when comparing data models.

As the prime metric of determining the performance of the model, accuracy is the percentage of predictions each model got right but especially in multi-class classification problems, sometimes 'accuracy' can't measure the performance successfully by itself and it might be deceptive.

In this manner, as the weighted average of Precision and Recall which are also performance metrics, F-Measure (F1) was also used to evaluate the performance of the model correctly.

3   Results and discussion

In this analysis, thirteen (13) different algorithms are used for multi-class classification, which are Light Gradient Boosting Machine, Random Forest Classifier, Gradient Boosting Classifier, Extra Trees Classifier, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Decision Tree Classifier, K Neighbors Classifier, Ridge Classifier, SVM-Linear Kernel, Logistic Regression, Ada Boost Classifier and Naive Bayes. Additionally, as ensemble methods, Bagging (Bagging Classifier) and Blending (Voting Classifier) were included in the study to improve the performance

of the algorithms. Then, to evaluate the efficiency of each algorithm, Accuracy and F-Measure as the performance metrics and the time required to build the model were used. According to the performance metrics, the algorithms were evaluated and compared. In addition, a comparison was made between ensemble methods and the algorithms to examine the usefulness of ensemble methods to improve the performance of the algorithms in engine system fault diagnosis. The results of performance metrics and model construction times are presented in Table 3 by calculating the created multiclass classification model with 13 different algorithms via PYTHON programming language.

As seen in Table 3, It was determined that Light Gradient Boosting Machine had the most efficient and accurate performance with relatively little time to build the model. Although it required little time to build the model, Naive Bayes had the worst performance of the algorithms selected in the survey. Ada Boost Classifier did not have satisfactory performance neither. Besides, Logistic Regression, SVM, Linear Discriminant Analysis and Ridge Classifier had close performances and accuracy to each other but less than 90% accuracy. In addition, with the variable construction times, K Neighbors Classifier, Decision Tree Classifier, Quadratic Discriminant Analysis, Extra Trees Classifier, Gradient Boosting Classifier and Random Forest Classifier had over than 90% accuracy performance.

After the evaluation of the main algorithms that used to predict the engine system faults, the ensemble methods were applied to improve the analysis.

By applying the Bagging Classifier as an ensemble method Bagging to the survey, it was discovered that Gradient Boosting Classifier had the best performance and accuracy with the increase of around 0.4%, however, needed much time to build the model. With a little improvement of performance metrics, Light Gradient Boosting Machine,

Table 3 Results of research algorithms (MCT: Model construction time including CPU time)

	Model	Accuracy	F1	MCT (sec)
lightgbm	Light Gradient Boosting Machine	*0.9803*	*0.9773*	18.6
rf	Random Forest Classifier	0.9794	0.9759	14.3
gbc	Gradient Boosting Classifier	0.9794	0.9760	161.0
et	Extra Trees Classifier	0.9766	0.9722	9.65
qda	Quadratic Discriminant Analysis	0.9734	0.9679	0.934
dt	Decision Tree Classifier	0.9716	0.9714	2.68
knn	K Neighbors Classifier	0.9592	0.9574	1.93
ridge	Ridge Classifier	0.8947	0.9136	0.989
Ida	Linear Discriminant Analysis	0.8883	0.9105	1.76
svm	SVM – Linear Kernel	0.8563	0.8880	2.5
Ir	Logistic Regression	0.8494	0.8873	8.52
ada	Ada Boost Classifier	0.6474	0.6947	9.99
nb	Naive Bayes	0.4863	0.6133	0.981

Source: Authors

Table 4 Results of research algorithms with the ensemble method Bagging (MCT: Model construction time including CPU time)

	Model	Accuracy	F1	MCT (sec)
lightgbm	Light Gradient Boosting Machine	0.9812	0.9783	123.0
rf	Random Forest Classifier	0.9776	0.9733	75.0
gbc	Gradient Boosting Classifier	*0.9831*	*0.9808*	446.0
et	Extra Trees Classifier	0.9776	0.9739	51.6
qda	Quadratic Discriminant Analysis	0.9725	0.9669	2.86
dt	Decision Tree Classifier	0.9776	0.9761	9.36
knn	K Neighbors Classifier	0.9648	0.9618	10.5
ridge	Ridge Classifier	0.8942	0.9139	1.5
Ida	Linear Discriminant Analysis	0.9405	0.9358	3.46
svm	SVM – Linear Kernel	0.9240	0.9333	14.0
Ir	Logistic Regression	0.8796	0.9066	24.5
ada	Ada Boost Classifier	0.9396	0.9346	236.0
nb	Naive Bayes	0.5174	0.6322	9.93

Source: Authors

Table 5 Main results of the ensemble method Blending (MCT: Model construction time including CPU time)

Model	Accuracy	F1	MCT (sec)
Lightgbm & Rf	0.9821	0.9796	124.0
Lightgbm & Gbc	0.9853	0.9832	161.0
Rf & Gbc	*0.9863*	*0.9843*	86.0
Lightgbm & Rf & Gbc	0.9812	0.9785	171.0
Qda & Rf & Gbc	0.9831	0.9806	82.0
Ada & Qda & Rf & Gbc	0.9799	0.9764	321.0
Qda & Lightgbm & Rf & Gbc	0.9840	0.9817	184.0
Qda & Knn & Rf & Gbc	0.9831	0.9806	83.0
Qda & Lightgbm & Et & Gbc	0.9826	0.9800	217.0
Qda & Lightgbm & Dt & Gbc	0.9849	0.9829	173.0
Nb & Gbc	0.6529	0.7658	66.0

Source: Authors

Random Forest, Extra Trees Classifier and Decision Trees Classifier had close results. With 45% improvement of accuracy, Ada Boost Classifier had the most remarkable result when the ensemble method of Bagging Classifier added the analysis, although it needed long model construction time. All the results of applied Bagging ensemble method including model construction time are shown in Table 4.

On the other hand, Blending (Voting Classifier) analysis has been added as another ensemble method. Random Forest and Gradient Boosting Classifier are achieved with the help of Blending ensemble method, with 98.63% accuracy and 98.43% F-Measure performance metrics, with the best results in all analyzes and also a relatively shorter model build time. In the context of results, if the analysis included Gradient Boosting Classifier in the Blending ensemble method, the performances of the combinations are mostly observed around 98% accuracy and F-Measure with variable model building time. However, Naive Bayes shows the worst predictions of damage even the fact that it is attempted to improve the performance by including Gradient Boosting Classifier and adding Voting Classifier as Blending ensemble method with 65.29% accuracy and 76.58% F-Measure. The main results of applied Blending ensemble method are seen in Table 5.

As a result, the comparison of significant performance results of the research algorithms with the combination of Bagging ensemble method, which are Gradient Boosting Classifier, Adaboost Classifier and Naive Bayes, are shown in Figure 6. In addition, the comparison of significant performance results and model construction time of the research algorithms with the combination of Blending ensemble method, which are Gradient Boosting Classifier & Random Forest, Light Gradient Boosting Machine & Gradient Boosting Classifier and Gradient Boosting Classifier & Naive Bayes are shown in Figure 7.
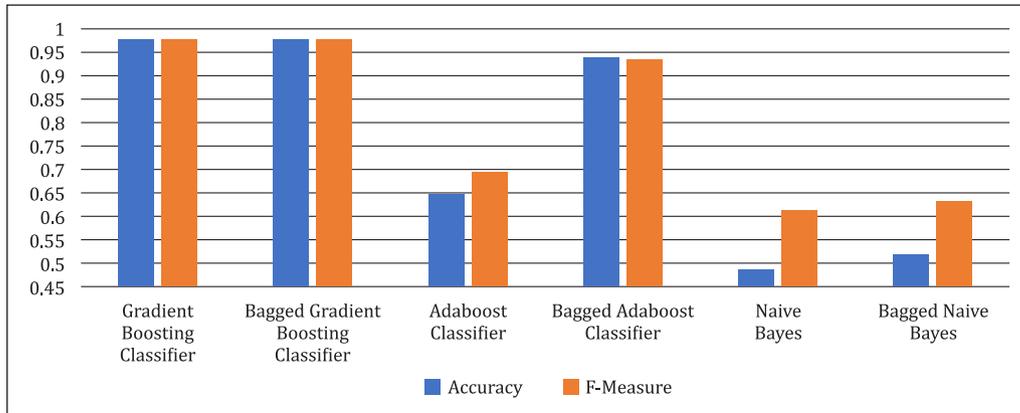
Figure 6 Comparison of performance metrics of Gradient Boosting Classifier, Adaboost Classifier and Naive Bayes algorithms with the Bagging ensemble method.
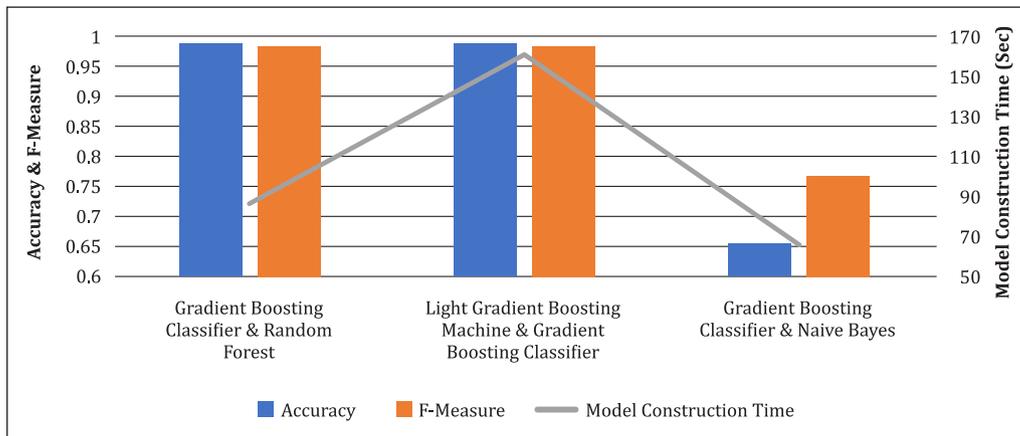
Source: Authors

Figure 7 Comparison of performance metrics and model construction time of Gradient Boosting Classifier & Random Forest, Light Gradient Boosting Machine & Gradient Boosting Classifier and Gradient Boosting Classifier & Naïve Bayes combined algorithms with the Blending ensemble method.
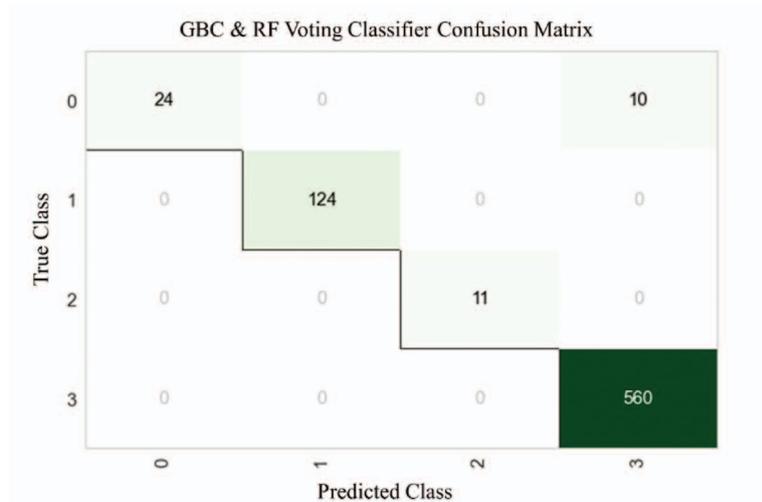
Source: Authors

Figure 8 Confusion matrix for Gradient Boosting Classifier & Random Forest Blended by Voting Classifier

Source: Authors

In the context of the classification system performance table, Confusion Matrix were obtained for Gradient Boosting Classifier & Random Forest, which are blended by Voting Classifier and had the best performance metric results with successful model construction time. In Figure 8, '0' represents Combustion/Exhaust System Failure class, '1' Cooling System Failure class, '2' Lubrication System Failure class and '3' Normal Operating Condition class. As seen in Figure 8, the predictions for class 1, 2 and 3 are rather successful and the prediction success for class 0 is relatively less than the other classes.

Unlike similar studies in the literature, in this study, cooling system, lubrication system and combustion/exhaust system failures, which are the three most important sub-systems of a ship diesel engine, are investigated together. The multi-class classification model, built on these 3 sub-system faults, is basically based on determining which element of the system a fault originates from. For example; the occurrence of a high exhaust temperature alarm in cylinder number 12 according to the design parameters may be caused by an injector failure, which is an element of the combustion/exhaust system, as well as the failure of the oil jet, which is an element of the lubrication system, or insufficient cooling liquid to reach that cylinder, which is an indication of a cooling system element failure. Our machine learning model aimed to find out which system element caused the source of the fault (root cause). When the obtained results are examined, it has been shown that although the detection of combustion/exhaust system failures is partially lower success than the detection of other failures, the overall success of the model has successful results when blending ensemble method is used.

4 Conclusions

Machine learning algorithms are expected to produce successful and reliable solutions for the rapid detection of engine failures, which is an important problem in the machine management of ships used for special purposes. In this study, thirteen different machine learning algorithms are studied on the model developed for fault diagnosis of four-stroke high-speed marine vehicles.

As a result, among the results shown, the F-measures of seven algorithms (Light Gradient Boosting Machine, Random Forest Classifier, Gradient Boosting Classifier, Extra Trees Classifier, Quadratic Discriminant Analysis, Decision Trees Classifier and K Neighbors Classifier) were in the range of 95.92% and 98.03%. Besides, the accuracy values varying in the range of 95.74 and 97.73%. However, Naive Bayes (48.63% accuracy, 61.33% F-Measure) and Ada Boost Classifier (64.74% accuracy, 69.47% F-Measure) outputs had the worst performance among these calculations.

Although it required a long model construction time, Gradient Boosting Classifier had the most successful results with 98.31% accuracy and 98.08% F-Measure values when as the Bagging ensemble method, Bagging Classifier added to the evaluations. On the other hand, Ada Boost Classifier had an effective improvement (45% accuracy, 35% F-Measure) with 93.96% accuracy and 93.46% F-Measure values. Naïve Bayes had the worst performance although Bagging Classifier added to the evaluations.

When the blending ensemble method, Voting Classifier applied to the evaluations, most successful results obtained by the combination of Gradient Boosting Classifier and Random Forest. The combination of Gradient Boosting Classifier and Random Forest had 98.63% accuracy and 98.43% F-Measure values with rather less model construction time (86 sec). By this result, we obtained the best solution of this multi-class classification problem which focused on fault diagnosis of four-stroke marine diesel engine.

As the result of this research, although ship diesel engines have complex systems, the general fault condition of the engine can be easily diagnosed by modeling the sub-systems with a certain independence with the help of sensor data. With this study, *root causes of engine system failures* are detected via machine learning algorithms and ensemble methods by using real-time data collected from an MTU brand four-stroke high-speed diesel engine belonging to a special-purpose Coast Guard ship. Although the study includes a small number of samples, it is basically a proof of concept. Increasing the number of samples and balancing the sample size of the classes and/or working with large-scale statistics will yield much more successful results. With the additional features, malfunctions of specific components (sub-elements of the systems) on the engine can be detected or precautions can be taken before a malfunction occurs and also studies can be carried out to detect engine start system, turbocharger system and power generation system (crank-piston system) malfunctions. In addition, it will be possible to develop and use this model in a study to be carried out to determine the necessary maintenance in order to prevent the occurrence of malfunctions.

Author Contributions: Research, Tolga Şahin; Data Collection, Tolga Şahin; Writing, Tolga Şahin and Adem Candaş; Review and editing, Adem Candaş, C. Erdem Imrak and Altan Cakir; Supervision, C. Erdem Imrak and Altan Cakir.

References

[1] Chandroth GO, Sharkey AJC, Sharkey NE. Cylinder Pressures and Vibration in Internal Combustion Engine Condition Monitoring. Coxmoor Publishing Co: Chipping Norton, UK, 1999. pp. 141–151.

[2] Sharkey AJC, Chandroth GO, Sharkey NE. Acoustic emission, cylinder pressure and vibration: A multisensor approach to robust fault diagnosis. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium; 2000 July 24-27; Como, Italy. pp. 223–228.

[3] Yasir HA. Artificial Intelligence-Emerging Trends and Applications. Artificial Intelligence Application in Machine Condition Monitoring and Fault Diagnosis,2018. pp. 275–291.

[4] Vladimir D. Condition monitoring and fault diagnosis, ScienceDirect, Procedia Engineering, 2017;177:502–509.

[5] Zhixiong L, Xinping YCS. A novel remote condition monitoring and fault diagnosis system for marine diesel engines based on the compressive sensing technology, Journal of Vibroengineering, 2014;16(2):879–890.

[6] Zhixiong L, Xinping Y, Chengqing Y, Zhongxiao P. Intelligent fault diagnosis method for marine diesel engines using instantaneous angular speed, Journal of Mechanical Science and Technology, 2012;26(8):2413–2423.

[7] Tsaganos G, Papachristos D, Nikitakos N, Dalaklis D, Olcer AI. Fault Detection and Diagnosis of Two-Stroke Low-Speed Marine Engine with Machine Learning Algorithms, INT-NAM 2018 3rd International Naval Architecture and Maritime Symposium, Conference Paper, 2018 Apr 23-25; Besiktas, Istanbul. ResearchGate 2018; pp. 699–712.

[8] Mohd NCW, Mamat R, Najafi G, Mat YMH, Ihsan CK, Noor MM. Prediction of marine diesel engine performance by using artificial neural network model, Journal of Mechanical Engineering and Sciences (JMES). 2016;10(1): 1917–1930.

[9] McCulloch WS, Pitts W A. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematics and Biophysics, 1943(5):115–133.

[10] Arslan K. Artificial Intelligence and Applications in Education. Western Anatolia Journal of Educational Sciences. 2020;11(1):71–80.

[11] Gogas P, Papadimitriou T. Machine Learning in Economics and Finance. Computational Economics. 2021;57:1–4.

[12] Pulat M, Deveci I. Bibliometric Analysis of Theses Published on Machine Learning and Decision Trees in Turkey. Journal of Management and Economics. 2021;28(2):287–308.

[13] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning, Springer Series in Statistics, Springer-Verlag New York; 2009.

[14] Kucuk H, Eminoglu I, Balci K. Classification of neuromuscular diseases with artificial intelligence methods. Journal of the Faculty of Engineering and Architecture of Gazi University. 2019;34(4):1725–1742.

[15] Gurkahraman K, Karakıs R. Brain tumors classification with deep learning using data augmentation. Journal of the Faculty of Engineering and Architecture of Gazi University. 2021;36(2):997–1012.

[16] Toraman S, Turkoglu I. A new method for classifying colon cancer patients and healthy people from FTIR signals using wavelet transform and machine learning techniques. Journal of the Faculty of Engineering and Architecture of Gazi University. 2020;35(2):933–942.

[17] Zhang J, Li L, Chen W. Predicting Stock Price Using Two-Stage Machine Learning Techniques. Computational Economics. 2021;57:1237–1261.

[18] Guerzoni M, Nava RC, Nuccio M. Start-ups survival through a crisis. Combining machine learning with econometrics to measure innovation, Economics of Innovation and New Technology. 2021;30(5):468-493.

[19] Majumdar A, Hall G, Ahmadi AA. Recent Scalability Improvements for Semidefinite Programming with Applications in Machine Learning, Control, and Robotics. Annual Review of Control, Robotics, and Autonomous Systems. 2020;3(1):331–360.

[20] Shimizu R, Kobayashi S, Watanabe Y, Ando Y, Hitosugi T. Autonomous materials synthesis by machine learning and robotics. APL Materials. 2020;8(11): 111110.

[21] Inan T, Baba AF. Building a hybrid algorithm based decision support system to prevent ship collisions. Journal of the Faculty of Engineering and Architecture of Gazi University. 2020;35(3):1213–1230.

[22] Azuz E, Tatar T, Alabas C. A decision support system based on simulation for airline boarding problem. Journal of the Faculty of Engineering and Architecture of Gazi University. 2019;34(4):1777–1792.

[23] MTU_Value Service, Diesel Engine 16V 4000 M90, Functional Description:M013011/01E, 2010.

[24] MTU Diesel Engine 16V 4000 M90, Operation Instructions, Description: MW15413/06E, 2011.

[25] Margaronis, I. E., Function Diagnostic Systems in Marine Diesel Engines. Doctoral thesis, National Technical University of Athens. 1986.

[26] MTU_Value Service, System Documentation BlueLine with Multi-Function Display DIS 14 for Series 4000 FPP MDEC, Description: E532776/01E, 2009.