Machine learning models to predict components decay in a naval propulsion system

Quatrini E.*, Colabianchi S.*, Costantino F.*, Tronci M.*

**Department of Mechanical and Aerospace Engineering, University of Rome "La Sapienza", Via Eudossiana, 18, 00184 Rome, Italy*

(elena.quatrini@uniroma1.it, silvia.colabianchi@uniroma1.it, francesco.costantino@uniroma1.it, massimo.tronci@uniroma1.it)

Abstract: The decay of a single component in a naval propulsion system can affect the performance of the entire system, involving expensive maintenance costs for restoring efficient conditions. Therefore, a regular control of the decay of key components of these systems is essential for properly handle maintenance actions. Moreover, in naval propulsion systems it is necessary to consider the difficulty in implementing an onboard maintenance action or returning a vessel. Two relevant components in naval propulsion systems are the turbine and the compressor. This study develops two machine learning models to predict turbine and compressor decay, i.e. based on classification and regression approaches. The former classifies whether the components are decayed or not, thus highlighting a state of criticality, the latter predicts a specific value of each decay coefficient. For each approach, different algorithms are compared, e.g. boosted trees, linear regression or support vector machines. A case study considering sixteen inputs has been used to test the effectiveness of the proposed solution, starting from a dataset of about twelve thousand instances referred to a naval vessel. A sensitivity analysis of relevant parameters has been developed to verify the robustness of the approach.

Keywords: decay coefficients; naval propulsion system; condition-based maintenance; turbine; compressor; maritime maintenance

1.Introduction

Vessels are nowadays the most important means of transport in global trade, with more than ten billion tons of goods shipped (Brooks & Faust, 2018). Maintaining naval propulsion systems can be very expensive and can strongly affect the total amount of costs to sustain (Coraddu et al., 2016). Specifically, naval propulsion systems must deal with both the high costs of individual components and the high costs of maintenance actions. One can affirm that maintenance costs in maritime settings may be 20% of total operating costs and may be even higher for an offshore plant (Sebastiani, Pescetto, & Ambrosio, 2013). For all these reasons, maintenance of the several components of a naval propulsion system is an onerous activity, that must be efficiently programmed. The distinctive characteristic of naval propulsion systems is that they can require a dry dock to be maintained and the difficulties of bringing a boat back to port are not negligible (Cipollini, Oneto, Coraddu, Murphy, & Anguita, 2018a). Sometimes it is impracticable to perform maintenance actions on board and offshore. Modelling the behavior and the interactions of the components of a naval propulsion system can be very complicated, affecting or even making the implementation of proper maintenance impossible. There are many variables at stake and their a priori physical modelling is difficult. Data-driven models can help to solve this problem, achieving considerable results (Cipollini, Oneto, Coraddu, Murphy, & Anguita, 2018b). Based on all these considerations, this paper proposes two machine learning models, to predict turbine and compressor decay in a naval propulsion system characterized by a gas turbine propulsion plant. The paper suggests two different approaches: the former is a regression analysis, that predicts a specific value of each decay coefficient; the latter is a classification approach, that classifies whether the components are decayed or not, thus highlighting a state of criticality. The aim is to compare the approaches and evaluate their respective performances, to determine whether they are both performing, or you need to prefer one over the other. Besides, the case study checks for any significant differences between the results of the compressor and turbine analyses, both for the regression and the classification approach. A sensitivity analysis has been developed to assess how much the variations of individual inputs as well as combinations of inputs affect the output, to evaluate the robustness of the approach. Some considerations about the correlation between the inputs themselves have been carried out, with different tests for the dimensionality reduction, to decrease the variables redundancy. For this purpose, the technique applied is the Principal Component Analysis (PCA), comparing PCA with different levels of variability covered. The case study presented in this paper relies on a dataset available online on the UC Irvine Machine Learning Repository (Coraddu et al., 2016, 2015). The remainder of the paper is organized as follows. Section 2 presents the related literature. Section 3 presents the machine learning models applied to a case study and Section 4 explains the analyses carried out in the case study. The discussion of the paper results, the conclusions, and future research are presented in Section 5.

2.Literature review

Naval propulsion systems are widely discussed in relation to condition-based maintenance. This is due to the fact that marine diesel engines are obliged to have high levels of reliability, and availability, to meet stringent in-service, and operating requirements (Giorgio, Guida, & Pulcini, 2007). Considering the naval marine sector, nowadays several ship programs currently use "Combined Diesel Electric and Gas Turbine" (CODELAG) propulsion (Hendry & Bellamy, 2019), as in the presented case study. Particularly, the gas turbines have been the subject of study for many years (Anon, 1986; Kerpestein, 1987; Pierce & Shu, 1984). As presented in the introduction, this is due to the high maintenance costs that can account for most of the costs of ship machinery maintenance (Coraddu et al., 2016). Considering the naval propulsion systems, they are made up of very expensive equipment (Coraddu et al., 2016), such as gas turbines or cooler, compressor and fuel injection system (Altosole et al., 2014; Basurko & Uriondo, 2015). Modelling a naval propulsion system can be very complicated and in such complex systems the application of data-driven models facilitates the maintenance implementation (Cipollini et al., 2018b; Györfi, Kohler, Krzyzak, & Walk, 2002). It is, therefore, necessary to consider the difficulty of having a labelled dataset (Altosole et al., 2014), since it can be very difficult to retrieve data from fault situations (Tan, Niu, Tian, Hou, & Zhang, 2019). Two of the main components of a naval propulsion system, specifically in a gas turbine (GT) propulsion plant, are the GT compressor and the GT turbine (Cipollini et al., 2018b). Their appropriate maintenance becomes essential not only for the proper functioning of the system but also for the containment of emissions (Lorencin, Anđelić, Mrzljak, & Car, 2019). In fact, the fouling of the GT compressor increases the specific fuel consumption and the temperature of the exhaust gas (Tarabrin, Schurovsky, Bodrov, & Stalder, 1998). Consequently, the effects of fouling also affect the efficiency of the GT, which can be represented with a reduction of the GT flow rate, explained by *kMt* (Altosole et al., 2014; Coraddu et al., 2016). The analysis of the pollution emissions can even help to implement a structured prognosis for the gas turbines (Kacprzynski et al., 2001). The dataset used in this paper has been used for testing the applicability of some well-known regression algorithms such as regularized kernel least squares and support vector regression (Coraddu et al., 2016, 2015) as well as with neural networks (Cisotto & Herzallah, 2019; Lorencin et al., 2019). Based on the analyzed literature, the authors present two machine learning models for predicting turbine and compressor decay in a GT propulsion plant. The paper also suggests considerations on the dimensionality reduction and the influence of individual variables as well as their variability on the model.

3.Machine learning models application to a case study

The machine learning models developed to study the turbine and compressor decay are shown by a case study application. The dataset used in this case study (Coraddu et al., 2016, 2015) has been generated with a numerical simulator of a naval vessel characterized by a GT propulsion system. The complete simulator is constituted by 5 blocks, i.e. propeller, hull, GT, gearbox, and controller. The considered outputs of this dataset are 2:

- *kMc* = GT Compressor decay state coefficient
- *kMt* =GT Turbine decay state coefficient

Specifically, *kMc* describes the reduction of the values of the airflow rate *Mc*. This value is indicative of the compressor status. Similarly, *kMt* describes the gas flow rate reduction factor on duty, providing the turbine state. The *kMc* has been investigated in the domain [1; 0.95], and the *kMt* in the domain [1; 0.975]. Ship speed has been investigated sampling the range of feasible speed from 3 knots to 27 knots with a sampling distance of 3 knots. As asserted in the abstract and the introduction, the authors have decided to analyse and predict the decay state of the turbine and the compressor with two approaches, i.e. a regression approach and a classification approach. The former predicts the punctual values of *kMc* and *kMt*. The latter only classifies whether the components are decayed or not. For the classification approach, the authors have followed what suggested in (Cipollini et al., 2018a):

- *kMc:* [0.95−0.98) decayed; [0.98−1] not decayed
- *kMt:* [0.975−0.99) decayed; [0.99−1] not decayed

Those thresholds are founded on the assumption that an effective time service of 2000 h per year is a reasonable operating time for these vessel types. Based on these thresholds, the authors have created a different dataset where the punctual values of *kMc* and *kMt* are replaced with "1" if they represent a decay state and "0" if they do not represent a decay state. So, at the end of these steps, there are two different datasets: the one with the specific value of *kMc* and *kMt* that is used for the regression analysis and the one with the sole representation of a state of decay or otherwise of the compressor and turbine, used for the classification approach. The inputs dataset is composed by 16 variables:

- x_1 = Lever position (lp) [ ]
- x_2 = Ship speed (v) [knots]
- x_3 = Gas Turbine (GT) shaft torque ($\mathbf{GT_T}$) [kN m]
- x_4 = GT rate of revolutions (GTn) [rpm]
- x_5 = Gas Generator rate of revolutions (GGn) [rpm]
- x_6 = Starboard Propeller Torque (Ts) [kN]
- x_7 = Port Propeller Torque (Tp) [kN]
- x_8 = Hight Pressure (HP) Turbine exit temperature (T48) [C]
- x_9 = GT Compressor inlet air temperature (T1) [C]
- x_{10} = GT Compressor outlet air temperature (T2) [C]
- x_{11} = HP Turbine exit pressure (P48) [bar]
- x_{12} = GT Compressor inlet air pressure (P1) [bar]
- x_{14} = GT exhaust gas pressure (Pexh) [bar]
- x_{15} = Turbine Injecton Control (TIC) [%]
- x_{16} = Fuel flow (mf) [kg/s]

4.Sensitivity analysis and correlation analysis

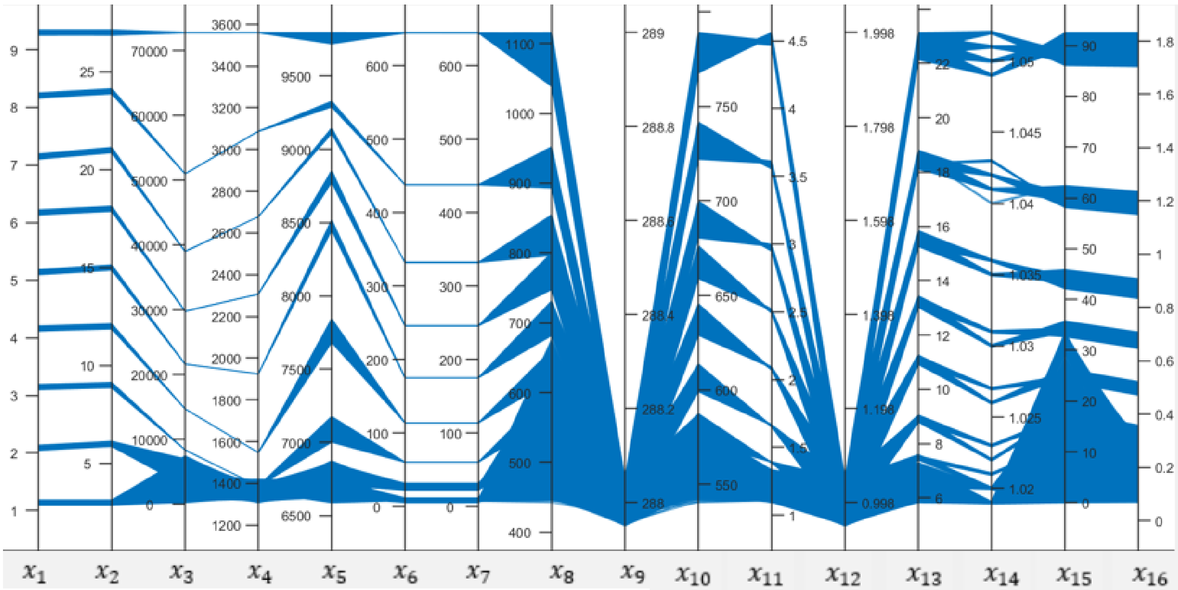

Figure 1: Data overall distribution

In this section, the use of the expression "step" has the function of representing the path taken by the authors to structure the models presented in the paper. It is, therefore, a way of representing and highlighting the sequence of actions carried out and their order of execution. The first step to apply is to verify the presence of variables that do not vary over time, to eliminate them from the analysis. Figure 1 presents the overall distribution of the data. In Figure 1, each line in the plot represents a complete instance, and each coordinate variable in the plot corresponds to a variable in the dataset. Two variables always assume the same value, which are x_9 and x_{12}. Neglecting these variables, the result is a dataset with 14 variables. The second step is the application of a sensitivity analysis. The latter explains how the variations of an input data affect the outputs. A model can be sensitive to a parameter in two main modes: the variability, or uncertainty, associated with one parameter propagates throughout the whole of the model, making a strong contribution to the variability of model outputs; model results can be highly correlated with a parameter of input so that small changes in input cause significant changes in the output. Specifically, feature values have been randomly shuffled, one column at a time and then testing the different possible combinations of the inputs. The performance of the model is measured before and after, and the output of this analysis is an index, $s_{i,j}$, representing the sensitivity coefficient of the variable *i* on the output *j*. The authors have implemented 4 different sensitivity analyses, one for every possible output. For the compressor, x_2 and x_7 have an $s_{i,j} = 0$, thus allowing them to be removed from the ranking of the variables. x_2 and x_7 have an $s_{i,j} = 0$ even for the turbine, enabling to neglect them for the following analyses. Moreover, x_2 is a linear function of x_1. Considering both variables, regardless of the $s_{i,j}$ obtained, would have provided redundant information. The third step for the development of the machine learning models was the analysis of the correlations between the selected inputs. This step permits to verify the presence of redundant variables that can weigh down the model. This analysis shows a strong correlation between the variables, always and for all of them higher than 0.89. An analysis of the significance of the p-values obtained confirms the rejection of the null hypothesis. Based on these observations, the authors decided to apply the PCA for reducing both the redundancy and the dimensionality of the variables. With PCA it is possible to apply an orthogonal transformation to convert a set of correlated variables into a set of values of linearly uncorrelated variables.

4.1 PCA and predictive models

Analyzing the dataset with the PCA, it is possible to do some interesting considerations on the data about the variations of the PCA according to the covered variability of the dataset. Specifically, with just one PCA component it is possible to cover 97.5% of the total variability of the dataset. In the same way, with 2 PCA components, it is possible to cover 99.4% of the total variability and with 6 PCA components, all the dataset variability can be covered. To summarize, the different variabilities covered by PCA are:

- 1 PCA component → 97.5% of covered variability
- 2 PCA components → 99.4% of covered variability
- 6 PCA components → 100% of covered variability

These results enable assessments to be made on the next steps of the prediction model. The three different combinations of PCA are tested both for the regression approach and then for the classification approach. The aim is to find the best options for the machine learning models and in the same way, see the distinctions with the different settings of the PCA. For the classification approach, the tested algorithms are two ensemble methods, i.e. the boosted trees and the bagged trees, the logistic regression, and the support vector machines (SVM). For the regression approach, the tested algorithms are the bagged trees, the linear regression, the robust regression, the interactions regression, and the stepwise linear regression (SL regression). The results of the predictive models are presented from Table 1 to Table 4. Specifically, Table 1 and Table 2 show for each combination of the PCA components and the tested algorithms the R^2, i.e. the coefficient of determination, which represents a proportion between the variability of the data and the correctness of the statistical model used. Table 3 and Table 4, on the other hand, show the accuracy of the classification models, i.e. the proportion between corrected predictions and total predictions. Table 1 and Table 2 show that both for *kMt* and *kMc*, the best results for the regression approach are obtained with the bagged trees, using the PCA with 6 components. Table 3 and Table 4 show the classification of the state of decay of the turbine and the compressor respectively. Both obtain the best results with the bagged trees, with PCA with 6 components.

Table 1: Results of the regression model for *kMt*

	97.5% of variability	*99.4% of variability*	*100% of variability*
Bagged trees	R^2: 0.15	R^2: 0.75	R^2: **0.97**
Linear regression	R^2: 0	R^2: 0	R^2: 0.28
Robust regression	R^2: 0	R^2: 0	R^2: 0.28
Interactions regression	R^2: 0	R^2: 0	R^2: 0.79
SL Regression	R^2: 0	R^2: 0	R^2: 0.79

Table 2: Results of the regression model for *kMc*

	97.5% of variability	*99.4% of variability*	*100% of variability*
Bagged trees	R^2: 0.58	R^2: 0.85	R^2: **0.98**
Linear regression	R^2: 0	R^2: 0	R^2: 0.51
Robust regression	R^2: 0	R^2: 0	R^2: 0.51
Interactions regression	R^2: 0	R^2: 0.01	R^2: 0.87
SL Regression	R^2: 0	R^2: 0.01	R^2: 0.87

Table 3: Accuracy of the classification model for the turbine decay

	97.5% of variability	*99.4% of variability*	*100% of variability*
Boosted decision trees	63.8%	66.4%	86.7%
Bagged trees	60.7%	89.9%	**97.1%**
Logistic regression	57.7%	57.7%	69.1%
SVM	53%	47.5%	69.2%

Table 4: Accuracy of the classification model for the compressor decay

	97.5% of variability	*99.4% of variability*	*100% of variability*
Boosted decision trees	81.2%	86.1%	95.6%
Bagged trees	77,7%	92.8%	**98.5%**
Logistic regression	58.8%	58.8%	78.2%
SVM	49.5%	47.8%	68%

Figure 2 and Figure 3 show the prediction trend with the regression approach both for *kMt* and *kMc,* obtained with the best settings of PCA and with the most performant algorithm, i.e. with the bagged trees and 6 PCA components. Going into detail, the black line represents "the perfect prediction", i.e. the situation in which predicted response and actual response coincide. Each point on the graph then highlights what is the prediction of the model and what is the real value of the decay coefficient of the component, allowing you to have a better insight into the performance of the models.

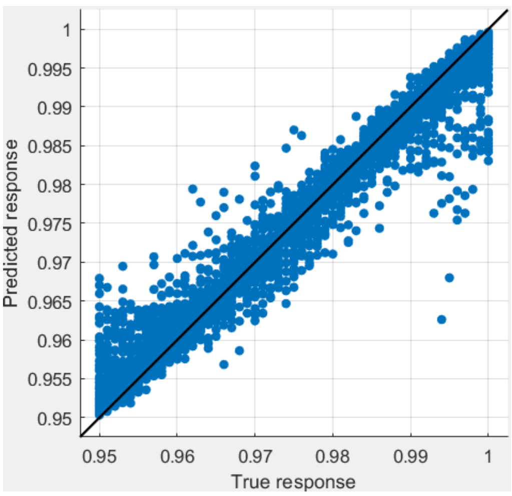

Figure 2: Prediction trend for *kMc* with the regression approach

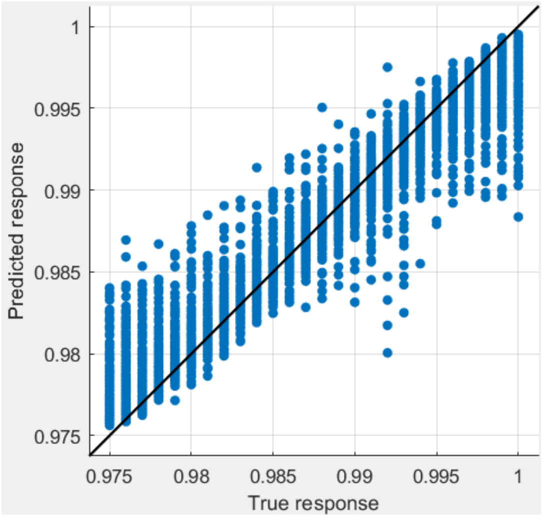

Figure 3: Prediction trend for *kMt* with the regression approach

As can be noticed, both models manage to obtain good predictions for both coefficients, being able to predict very well the dependent variable, i.e. the decay coefficients, based on the independent variables, i.e. the input parameters. Figure 4 and Figure 5 introduce respectively the ROC curve for the classification of the compressor decay and the turbine decay. From Figure 4 to Figure 7 the presented results refer to the classification machine learning models using 6 PCA components and the bagged trees as algorithm. The ROC curve is created by plotting the true positive rate and the false positive rate at various threshold settings. The closer the curve is to the upper left corner, the better the performance of the classifier. Curves that are close to the diagonal result from classifiers that tend to make estimates based on randomness. The results obtained by both models are very positive, as can be seen from the proximity of both ROC curves to the upper left corner.

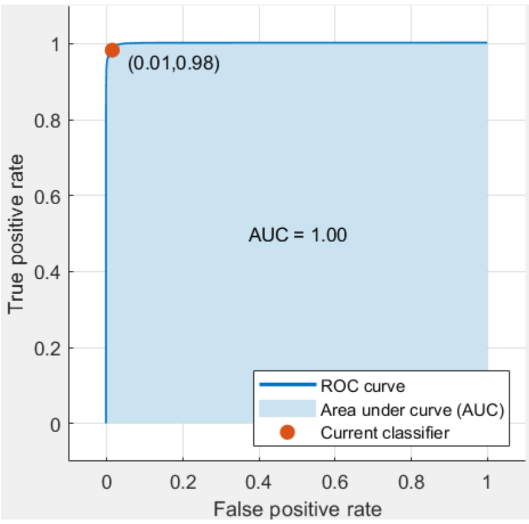

Figure 4: Best result for the compressor decay with the classification approach

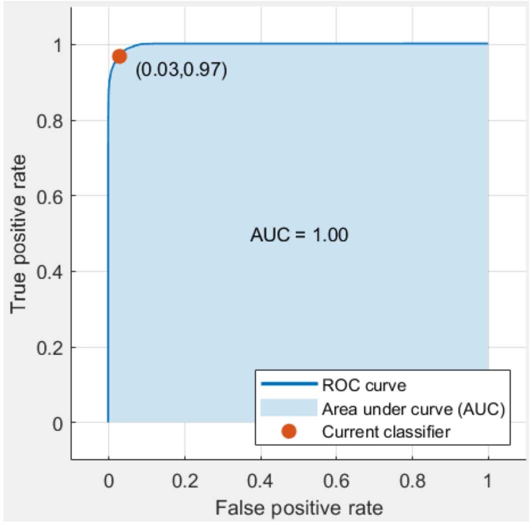

Figure 5: Best result for the turbine decay with the classification approach

Figure 6 and Figure 7 present respectively the confusion matrices for the compressor decay and the turbine decay, with the rate of correct predictions as well as of true and false positives and true and false negatives. The confusion matrices make it possible to understand the trend of classification errors, giving evidence of the statistical classification accuracy. Each column of the matrix represents the predicted values, while each row represents the real values. Both the compressor and turbine classification model have a very low false positive and false negative rate, but the latter are slightly better as far as the compressor is concerned.

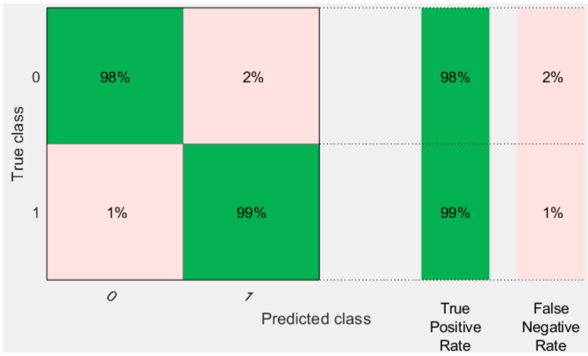

Figure 6: Confusion matrix for compressor decay

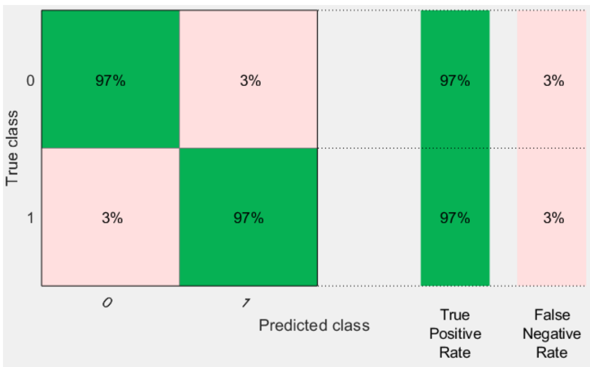

Figure 7: Confusion matrix for turbine decay

All machine learning models have been validated with cross-validation. Specifically, 10 folds were used for each model. So, the cross-validation process starts partitioning the data into 10 disjoint folds. The following step is that for every fold, the model is trained using the out-of-fold observations, and then it is assessed with the in-fold data, calculating the selected metrics over all folds. Cross-validation avoids overfitting and makes efficient use of all the data. Moreover, it can help to solve the problem of the asymmetric sampling of the training dataset, which may involve bias. The results obtained however open up the scenario for future research and reflections. In fact, PCA components 3, 4, 5, and 6 provide 0,6% of the variance, but the machine learning models improved a lot their performance. The results obtained with the application of cross-validation do not show the presence of overfitting. This leads to suspect the presence of physical characteristics of the system that determine these results. Based on this, in future research developments, it will be interesting to investigate the results obtained, to determine in detail the characteristics that define them.

5.Discussion and conclusions

It is interesting to note that both the regression approach and the classification approach achieve the best results with the same algorithm and under the same conditions of PCA. Specifically, the previous expression refers to the fact that bagged trees are the algorithm that guarantees the best performance for the prediction of *kMt* and *kMc* and for the classification of the presence or absence of a decay state in the turbine or in the compressor and the PCA that covers the 100% of the variability of the data reaches the best results both for the regression and the classification approach. These results can be found in Table 1, Table 2, Table 3, and Table 4. Considering the same tables, it can still be argued that already the PCA that covers 99.4% of variability achieved acceptable results in both approaches, but the PCA that covers the 97.5% of the data variability is not able to guarantee appropriate outcomes. According to the metrics of the different machine learning models, it can be noted that those relating to the compressor are significantly better than those relating to the turbine with the PCA with 1 and 2 components, as can be noted in Table 1, Table 2, Table 3 and Table 4. As shown in the same tables, this difference is practically zero when 6 components are considered in the PCA. For the evaluation of the regression algorithms, given the small values to be predicted, it was considered that the only metric relevant for their evaluation was R^2. Examining in detail the results obtained for *kMt* and *kMc* shown in Table 1 and Table 2, it is interesting to note that with the 1 and 2 components PCA no algorithm can guarantee a R^2>0, except for bagged trees that however return acceptable values only considering at least 2 components. Based on the same tables, with the 6-components PCA, bagged trees can obtain R^2 values close to the optimal value. As regards the classification approach, again the 6-components PCA applied with bagged trees guarantees optimal results, with accuracy in both cases greater than 97%, as shown in Table 3 and Table 4. For the regression approach, the results for *kMt* and *kMc* are substantially identical, but for the classification approach, one can note that for the analysis of the compressor the percentages of false positives and false negatives are lower, although the results can be considered basically comparable (Figure 6 and Figure 7). In conclusion, both the classification and regression approach can be applied with excellent results, also performing an important reduction in the considered variables. At the end of the analysis carried out in this paper, some considerations can be deduced from the results obtained. Both the proposed approaches, i.e. regression and classification, can be successfully applied. Neither performative with better results than the other. As can be seen from the results shown in Table 1, Table 2, Table 3 and Table 4, the classification is perfectly in line with how successful the fit is in explaining the variation of the data for the regression approach. It is interesting to note that despite a small variability gap covered between models with 6 PCA components and with 1 and 2 PCA components, the results between them are not comparable. The application of the PCA makes it possible to successfully reduce redundancy in the model and to lighten its development and implementation. Many advantages can be achieved by applying models of this type. Firstly, letting a machine work until a complete breakdown can lead to high stranded costs. This problem is exacerbated in contexts where machinery costs are extremely high, as in the context considered in this contribution. Secondly, through these types of approaches, it is possible to manage the execution of maintenance interventions to reduce downtime. Going into detail, the use of both classification and regression algorithms allows a dual viewpoint. Through regression algorithms, it is possible to have a more specific view of the health state of the machinery. Since regression can show not only a situation of decay or not, but also its level of intensity, through it is possible to monitor the development of degradation and consequently to plan maintenance interventions based on the working activity of the machine. Classification algorithms, on the other hand, since they only show the presence or absence of decay, can serve as a control for more detailed regression results. In fact, errors with regression algorithms could lead to erroneous considerations about the health of the machinery, which can instead be identified and corrected through the classification approach. The main objective therefore that the two proposed approaches want to pursue is based on their combined use. Regression allows a picture of the current state of the machine, but with a focus on its evolution over time, allowing for long-term maintenance management. Prediction errors in this context can, however, lead to erroneous considerations, which can be identified using classification algorithms. The latter, however, provide only a current view of the state of health of the machinery. In conclusion, these two approaches have a complementary role, but each one adds more value to the implemented condition monitoring. Future research will focus on the comparison on the same case study of PCA and other dimensionality reduction techniques, to find similarities and differences between them and related results. It would be interesting to apply the models developed on a real naval propulsion system dataset, to see how they perform in real-life situations and further validate their applicability. Moreover, as previously mentioned, future research will be directed towards the analysis of the characteristics of the system that lead to the results obtained, i.e. why 4 PCA components together provide 0,6% of the variance, but the machine learning models improved a lot. Finally, future interesting research could be carried out on the future forecasting of decay coefficient trends. This would permit the application of prognosis analyses on the naval propulsion system.

References

Altosole, M., Campora, U., Martelli, M., & Figari, M. (2014). Performance decay analysis of a marine gas turbine propulsion system. *Journal of Ship Research*, *58*(3), 117–129. https://doi.org/10.5957/JOSR.58.3.130037

Anon. (1986). Colloquium on gas turbines: control theory and practice. *IEE Colloquium (Digest)*, (1986 /116).

Basurko, O. C., & Uriondo, Z. (2015). Condition-based maintenance for medium speed diesel engines used in vessels in operation. *Applied Thermal Engineering*, *80*, 404–412. https://doi.org/10.1016/j.applthermaleng.2015.01.075

Breiman, L. (2001). Random Forest. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brooks, M. R., & Faust, P. (2018). 50 years of review of maritime transport, 1968–2018: reflecting on the past, exploring the future. *Technical Report*.

Cipollini, F., Oneto, L., Coraddu, A., Murphy, A. J., & Anguita, D. (2018a). Condition-based maintenance of naval propulsion systems: Data analysis with minimal feedback. *Reliability Engineering & System Safety*, *177*, 12–23. https://doi.org/10.1016/J.RESS.2018.04.015

Cipollini, F., Oneto, L., Coraddu, A., Murphy, A. J., & Anguita, D. (2018b). Condition-Based Maintenance of Naval Propulsion Systems with supervised Data Analysis. *Ocean Engineering*, *149*, 268–278. https://doi.org/10.1016/J.OCEANENG.2017.12.002

Cisotto, S., & Herzallah, R. (2019). Performance Prediction using Neural Network and Confidence Intervals: A Gas Turbine application. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 2151–2159. https://doi.org/10.1109/BigData.2018.8621919

Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., & Figari, M. (2016). Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment*, *230*(1), 136–153. https://doi.org/10.1177/1475090214540874

Coraddu, A., Oneto, L., Ghio, A., Savio, S., Figari, M., & Anguita, D. (2015). Machine learning for wear forecasting of naval assets for condition-based maintenance applications. *Electrical Systems for Aircraft, Railway and Ship Propulsion, ESARS*, *2015-May*. https://doi.org/10.1109/ESARS.2015.7101499

Giorgio, M., Guida, M., & Pulcini, G. (2007). A wear model for assessing the reliability of cylinder liners in marine diesel engines. *IEEE Transactions on Reliability*, *56*(1), 158–166. https://doi.org/10.1109/TR.2006.890901

Györfi, L., Kohler, M., Krzyzak, A., & Walk, H. (2002). A distribution-free theory of nonparametric regression. *A Distribution-Free Theory of Nonparametric Regression*.

Hendry, M. L., & Bellamy, N. (2019). Advantages and experience of using SSS (synchro-self-shifting) clutches in hybrid propulsion such as CodeLog or CodeLAg naval marine systems. *Proceedings of the ASME Turbo Expo*, *1*. https://doi.org/10.1115/GT2019-91873

Kacprzynski, G. J., Gumina, M., Roemer, M. J., Caguiat, D. E., Galie, T. R., & McGroarty, J. J. (2001). A prognostic modeling approach for predicting recurring maintenance for shipboard propulsion systems. *Proceedings of the ASME Turbo Expo*, *1*. https://doi.org/10.1115/2001-GT-0218

Kerpestein, J. B. (1987). High performance gearing for modern naval gas turbine propulsion systems. *American Society of Mechanical Engineers (Paper)*.

Lorencin, I., Anđelić, N., Mrzljak, V., & Car, Z. (2019). Multilayer perceptron approach to condition-based maintenance of marine CODLAG propulsion system components. *Pomorstvo*, *33*(2), 181–190. https://doi.org/10.31217/p.33.2.8

Pierce, R. M., & Shu, H.-T. (1984). System evaluation of advanced technology gas turbines for naval ship propulsion. *Proceedings of the Intersociety Energy Conversion Engineering Conference*, 1409–1414.

Sebastiani, L., Pescetto, A., & Ambrosio, L. (2013). The Condition Monitoring System for Optimal Maintenance – Possible Application on Offshore Vessels. *Offshore Mediterranean Conference and Exhibition*, p. 14. Retrieved from https://doi.org/

Tan, Y., Niu, C., Tian, H., Hou, L., & Zhang, J. (2019). A one-class SVM based approach for condition-based maintenance of a naval propulsion plant with limited labeled data. *Ocean Engineering*, *193*, 106592. https://doi.org/10.1016/J.OCEANENG.2019.106592

Tarabrin, A. P., Schurovsky, V. A., Bodrov, A. I., & Stalder, J.-P. (1998). An analysis of axial compressor fouling and a blade cleaning method. *Journal of Turbomachinery*, *120*(2), 256–261. https://doi.org/10.1115/1.2841400