ISSN: 2831 - 6789

Research Article

Journal of Robotics and Automation Research

Optimized Machine Learning Algorithms to Predict Wear Behavior of Tribo-Informatics

Poornima Hulipalled[1*], V Lokesha[2]

[1]*Department of Studies in Computer science, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India*

[2]*Department of Studies in Mathematics, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India*

***Corresponding Author**
Poornima Hulipalled, Department of Studies in Computer science, Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India.

Submitted:17 Nov 2022; **Accepted:** 25 Nov 2022; **Published:** 29 Nov 2022

Citation: *Hulipalled, P., Lokesha, V. (2022). Optimized Machine Learning Algorithms to predict wear behavior of Tribo-Informatics. J Robot Auto Res, 3(3), 316-321.*

Abstract
Wear rate prediction is most important in industrial applications. Machine learning (ML) has made an admirable contribution to the field of tribology. Standard ML models are extremely dependent on the parameter values; hence, tuning plays a crucial role in enhancing predictive performance. ML models largely work empirically, based on the data availability and application domain, the parameter tuning process effectively attains the desired accuracy of the models. The main aim of this study is to develop optimized ML models which render better accuracy than the previous study by using a grid search hyperparameter optimization technique. Five ML models namely Random Forest (RF), Support Vector Machine (SVM), K- Nearest Neighbor (KNN), Gaussian Process Regression (GPR), and Linear Regression (LR) are designed by tuning the parameters which lead to the optimization of models concerning the prediction accuracy.

Keywords: Grid search, Hypereutectoid steel, Hyperparameter optimization, Machine learning, Wear rate, Pin-on-disc, Random Forest, Support Vector Machine, K-Nearest Neighbor, Gaussian Process Regression, Linear Regression.

Introduction

Peter H. Jost coined the name "tribology" in 1966, and since then, the field has grown to encompass many other scientific disciplines. The field of tribology, which studies the interactions between surfaces, has recently grown into various new subfields, such as biomimetic tribology, nanotribology, and bio-tribology [1]. Despite the many attempts to develop principles on friction and wear, tribology is still an empirical study because of the complicated nature of friction and wear. Despite this, tribological studies had collected a wealth of information about the wear, friction, and surface property of a wide range of materials, opening the door to the data-driven study. As the processing power of computers has increased, new data-driven analytical methods have emerged that can provide novel insights.

Analysts are turning to artificial intelligence (AI) and ML methods of "Big Data" analytics to discover new connections in data-driven sectors. Triboinformatics is a relatively recent area of tribology made possible by the use of data-driven methods [2].

In most cases, the constituent known as carbon is responsible for modifying steel's properties [3-5]. Increases in the percentage of carbon in steel lead to improvements in both the material's tensile strength and its hardness [6-8]. Hypereutectoid steels have attracted much interest from different areas of engineering, especially the railway sector, because of how well they perform in terms of wear resistance and mechanical properties [9-10]. Using a grid search, determined the best settings for ML models used in tribo-informatics [11]. Applied the Bayesian hyperparameter optimization technique for the enhancement of tool wear rate prediction [12]. Hyperparameters for serval control were optimized by Hao et al. using Random Search [13]. Numerous research has attempted to predict wear performance using various machine learning (ML) methods however, their performance is dependent on hyperparameters. In our previous work we used various ML algorithms with default argument values to predict the wear behavior of hypereutectoid steel [14-21]. In this article, we have focused on the improvement of those ML models' performances using the concept of hyperparameter tuning. Grid search is used as an optimization technique for hyperparameter tuning to find the efficient combination of parameter values for the ML models.

Materials and Methods

Dataset

Collecting relevant data is an important first step in creating a powerful data-driven ML algorithm. Training the ML algorithms on a sizable, relevant dataset yields better predicting accuracy. Tribological experimentation is a time-consuming and resource-intensive endeavor necessary for ML analysis. Considering this issue, we have collected an effective dataset of the wear behavior of hypereutectoid steel by conducting experiments using the pin-on-disc wear testing machine. This dataset is composed of input features, namely sliding speed, normal pressure, and sliding distance followed by wear rate as an output parameter.

Data Pre-Processing

Here, we have made use of Spyder, an open-source integrated development environment for programming in Python (version 3.7) including its software packages and frameworks for analyzing ML algorithm performance. Several crucial data processing stages were completed before ML model installation, including data cleaning, data separation as the training and testing sets, the transformation of data, and data standardization. The dataset is split into ten-fold cross-validation to evaluate the efficiency of ML models.

Machine Learning Models

The purpose of regression analysis in supervised machine learning is to quantify the relationship that exists between the input parameters and the output parameters. In the course of this study, the below-mentioned ML algorithms were put into action for the prediction of tribological behavior based on input and output factors by utilizing the tribological data of hyper eutectoid steel.

Linear Regression

The linear relationship between the variables is one of the most significant and widespread regression approaches used to predict the result of a dependent variable based on the independent variables [22].

Random Forest

The RF method, which is essentially a collection of random decision trees, is far less dependent on the initial set of inputs than the decision tree. This is achieved because of randomly selected subset features used for training the model, reducing the covariance between the trees. In context to the regression, the average of predictions is considered and hence the generalized predictive model is produced which reduces the notorious overfitting tendency [23].

Support Vector Machine

SVM regression uses trial and error to determine the hyperplane that has the biggest margin and is the region where the majority of data points fall [24].

K-Nearest Neighbor

Non-parametric KNN regression provides an estimate of the relationship between independent variables and the continuous outcome based on an intuitive average of data from neighboring locations [25].

Gaussian Process Regression

In the field of machine learning, the nonparametric, Bayesian regression method known as Gaussian process regression (GPR) is ruffling feathers. GPR has several advantages, including the ability to cope with smaller datasets and can provide confidence estimates for predictions [26]. We have already provided in-depth discussions for the above ML models in our previous work [15].

Parameter Optimization of ML Models

Through optimization, the developed machine learning regression models' potential for accurate prediction may be raised to its highest level. We examined many ML models and used grid search in combination with ten-fold cross-validation to ascertain the parameters that would lead to the highest level of performance. These optimization strategies may be added to the already established ML models. These methods include repeatedly running ML models with varying parameter settings, to determine which combinations of settings produce the best results and then recommending those parameter values. The performance of the learning process as well as the accuracy of predictions can be directly influenced by hyperparameters, which are necessary parameters of the machine learning algorithm [27-28]. There is no "magic" combination of hyperparameters that, when applied to every dataset, will produce the best possible results. Before applying ML models, the records of the dataset were shuffled [29-30]. This method helped in getting good prediction accuracy because the ML models select (split) the test set randomly. Hence the shuffled dataset made the ML models learn better with a variety of data available.

At the stage of developing ML models, an optimization technique called grid search is used to select the effective combination of hyperparameters that gives the best performance metrics of wear rate prediction [31-32]. Table 1 maintains a record of the parameters that have been determined to be effective for a variety of ML regression models.

Table 1: Optimized parameters

Model name	Selected parameters
SVM	C=5, gamma=0.22, kernel='rbf', epsilon=0.05
GPR	random_state=29
RF	n_estimators=100, max_features=3, max_depth=5
LR	fit_intercept=True, normalize=True
KNN	n_neighbors=3

Results and Discussion

The In this section, to analyze the regression performance, MAE, RMSE, and R2 evaluative matrices were used. In regressor, MAE is a loss function that takes into account both positive and negative deviations from the real and predicted values when calculating a loss. The MSE metric evaluates how well a fitted line represents the data points. It is the difference that is squared between the values that are predicted by a model and the values that are observed.

In our earlier work we trained three ML models, namely SVM, GPR, and RF, to predict the wear behavior of hypereutectoid steel using the default parameter values or without passing any explicit values to the arguments [21]. From Table 2, it is seen that random forest showed the best results in training and test datasets with 95.4% and 94% accuracy respectively. Even the MAE and RMSE are also lower for the random forests as well.

Table 2: The efficiency of the algorithms without tuning the hyperparameters

	R^2- training	R^2-test	MAE	RMSE
SVM	0.895	0.832	0.307	0.454
GPR	0.907	0.845	0.362	0.466
RF	0.954	0.940	0.237	0.294

In this work, we have added two more ML models namely LR and KNN and the total of five comparative performance outcomes of different models with hyperparameter tuning are shown in Table 3.

To enhance the effectiveness of ML models, we have been working on optimization strategies including grid search and ten-fold cross-validation. Table 3 illustrates the level of accuracy achieved by the improved machine learning regression models when used to predict wear rate. The values of MSE, RMSE, and MAE were much lower than expected for each of the five standalone ML models, contributing to an R2 value that ranged between 0.80 and 0.98. When it came to the prediction of wear rate, the ML models performed "satisfactorily" according to the statistics. Among the models, the RF outperformed others with R^2 = 0.98, MAE=0.16, and RMSE=0.22 with a remarkable degree of concordance between the observed and anticipated rates of wear. With an R2 value of 0.97, an MAE value of 0.11, and an RMSE value of 0.19, SVM has been the second-best prediction model execution for wear rate.

Table 3: The efficiency of the algorithms with tuning the hyperparameters

	R^2- training	R^2-test	MAE	RMSE
SVM	0.988	0.966	0.111	0.193
GPR	0.947	0.923	0.277	0.331
RF	0.980	0.979	0.163	0.220
LR	0.826	0.808	0.528	0.647
KNN	0.923	0.918	0.335	0.421

From looking at Tables 2 and 3, it is clear that there is a significant difference in the levels of performance achieved by the models before and after hyperparameter tuning. In this comparison, the accuracy of all models has increased with hyperparameter tuning.

Fig. 1 depicts the regression analysis graph for LR, KNN, GPR, SVM, and RF with the adopted optimization technique. In the graph, blue points represent the test data, and red colored dots represent the data used to train the corresponding ML model. The regression line of training prediction is shown by the red dashed line, and its slope indicates how successfully the model was trained by making use of the training dataset. The ability of the model to make accurate predictions on the test dataset is represented by the dashed blue line.

Based on the findings above, we can conclude that every model did a better job predicting the rate at which hypereutectoid steel would wear out. The method of normalizing and training the models using hyperparameter tuning allows us to obtain more accurate performance measurements for the models. SVM and RF models achieved the highest percentage accuracy by comparing the results without hyperparameter tuning.

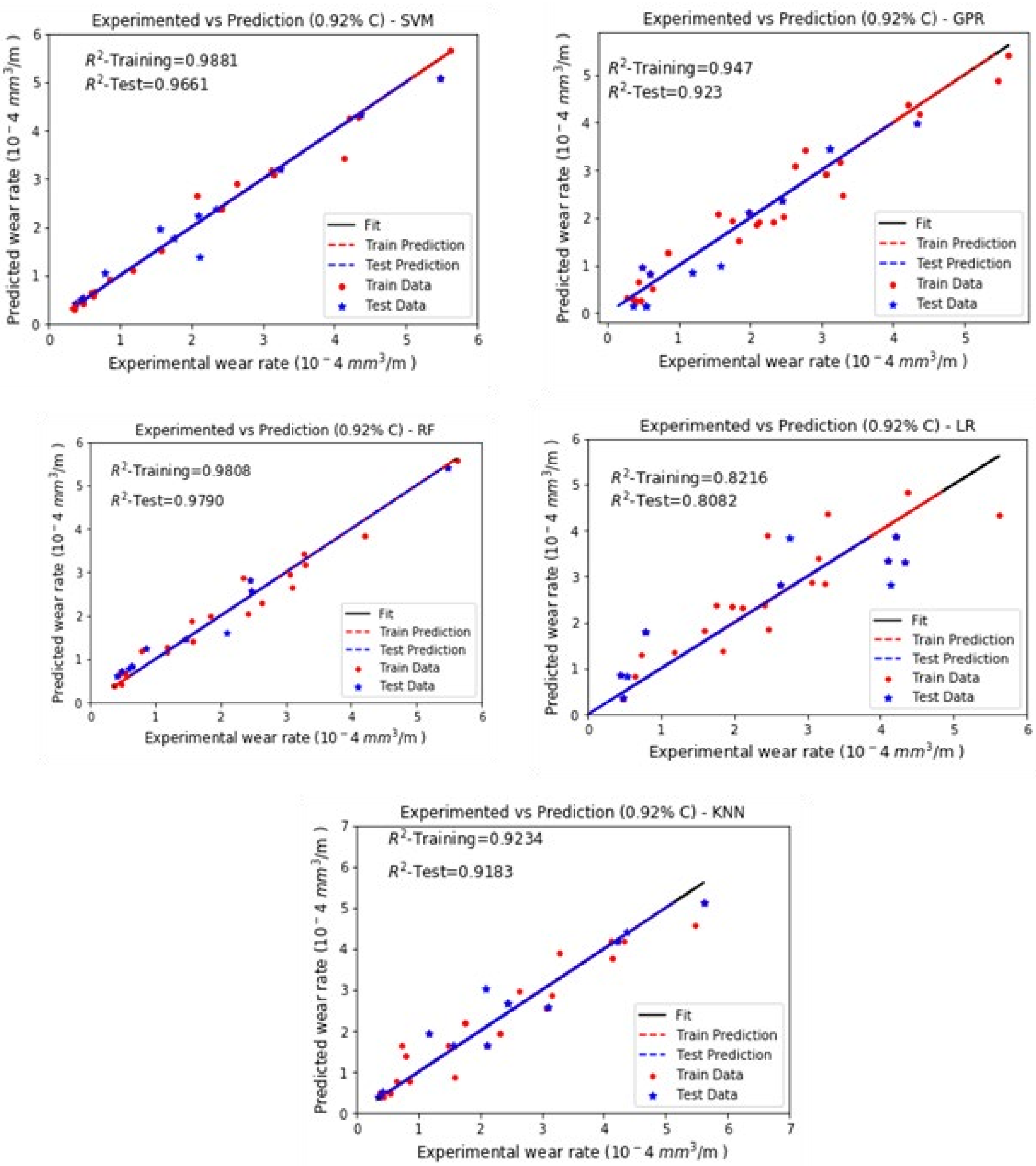

Figure 1: Regression analysis using SVM, GPR, RF, LR and KNN

Conclusions

Hyperparameter optimization aims to get maximum data performance within an acceptable time period. This is crucial to a machine learning algorithm's ability to make reliable predictions. In this study, we have provided the findings of an evaluation that was conducted on many different ML algorithms to predict the wear rate of hypereutectoid steel. The usefulness of the algorithms in predicting wear rate is evaluated, and they are contrasted with data obtained from earlier research. It would appear that almost all of the models reached the utmost feasible degree of accuracy once they were given access to a higher number of distinct machine learning matrices. The current ML models with hyperparameter tuning have at least 5% improved accuracy when compared to the results of our previously published ML models on the same dataset and all ML models show a discernible decline in error levels. Hence, optimization strategies like hyperparameter tuning and 10-fold cross-validation are applied to enhance the performance of the models.

Declarations

Conflict of Interest

The authors declare that there are no competing interests.

Author Contributions

Poornima Hulipalled: Data analysis and interpretation, Drafting the article, V Lokesha: Critical revision of the article. All authors reviewed the results and approved the final version of the manuscript.

Funding Declaration

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors would like to express their gratitude to Dr. Veerabhadrappa Algur, Associate Professor, Department of Mechanical Engineering, R.Y.M. Engineering College, Ballari, for providing the necessary dataset to conduct this research work.

Availability of Data and Materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request

References

1. Menezes, P. L., Nosonovsky, M., Ingole, S. P., Kailas, S. V., & Lovell, M. R. (Eds.). (2013). Tribology for scientists and engineers (pp. 9-28). New York: Springer.
2. Kordijazi, A., Roshan, H. M., Dhingra, A., Povolo, M., Rohatgi, P. K., & Nosonovsky, M. (2020). Machine-learning methods to predict the wetting properties of iron-based composites. Surface Innovations, 9(2–3), 111-119.
3. Gupta, A. K., & Mishra, D. A. (2013). An experimental investigation of the effect of carbon content on the wear behavior of plain carbon steel. Int J Sci Res, 2(7), 222-224.
4. Sharanabasappa, M., Kabadi, V. R., & Algur, V. (2014). Some investigation on dry sliding wear behaviour of ultra high carbon steel. Int J Mech Eng Res, 4(1), 75-82.
5. Qiao, L., Zhu, J., & Wang, Y. (2022). Machine Learning-Aided Process Design: Modeling and Prediction of Transformation Temperature for Pearlitic Steel. steel research international, 93(1), 2100267.
6. Sharanabasappa, M., Kabadi, V. R., & Algur, V. (2015). The effect of pearlite, cementite, and martensite phases on volumetric wear rate of hypereutectoid steel under dry sliding conditions. Int J Metall Mater Sci Eng, 5(1), 2278-2524.
7. Hosmani, S. D., Kurhatti, R. V., & Kabadi, V. K. (2017). Wear behavior of spherodized cementite in hyper eutectoid plain carbon steel. International Advanced Research Journal in Science, Engineering and Technology, 4(7), 257-262.
8. Wang, Y., Lei, T., & Liu, J. (1999). Tribo-metallographic behavior of high carbon steels in dry sliding: I. Wear mechanisms and their transition. Wear, 231(1), 1-11.
9. Liu, K. P., Dun, X. L., Lai, J. P., & Liu, H. S. (2011). Effects of modification on microstructure and properties of ultrahigh carbon (1.9 wt.% C) steel. Materials Science and Engineering: A, 528(28), 8263-8268.
10. Sharanabasappa, M., Kabadi, V. R., Shetty, P. B., & Algur, V. (2015). Dry sliding wear behaviour of hypereutectoid steel under the influence of microstructures, sliding speeds and normal pressures. International Journal of Mechanical Engineering and Robotics Research, 4(2), 60.
11. Hasan, M. S., Kordijazi, A., Rohatgi, P. K., & Nosonovsky, M. (2021). Triboinformatic modeling of dry friction and wear of aluminum base alloys using machine learning algorithms. Tribology International, 161, 107065.'
12. McParland, D., Baron, S., O'Rourke, S., Dowling, D., Ahearne, E., & Parnell, A. (2019). Prediction of tool-wear in turning of medical grade cobalt chromium molybdenum alloy (ASTM F75) using non-parametric Bayesian models. Journal of Intelligent Manufacturing, 30(3), 1259-1270.
13. Hao, M. R., Ahmad, M. A., Raja Ismail, R. M. T., & Nasir, A. N. K. (2018). Performance evaluation of random search based methods on model-free wind farm control. In Intelligent Manufacturing & Mechatronics (pp. 657-670). Springer, Singapore.
14. Qiao, L., Zhu, J., & Wang, Y. (2022). Machine Learning-Aided Process Design: Modeling and Prediction of Transformation Temperature for Pearlitic Steel. steel research international, 93(1), 2100267.
15. Algur, V., Hulipalled, P., Lokesha, V., Nagaral, M., & Auradi, V. (2022). Machine learning algorithms to predict wear behavior of modified ZA-27 alloy under varying operating parameters. Journal of Bio-and Tribo-Corrosion, 8(1), 1-10.
16. Altay, O., Gurgenc, T., Ulas, M., & Özel, C. (2020). Prediction of wear loss quantities of ferro-alloy coating using different machine learning algorithms. Friction, 8(1), 107-114.
17. Aydin, F., & Durgut, R. (2021). Estimation of wear performance of AZ91 alloy under dry sliding conditions using machine learning methods. Transactions of Nonferrous Metals Society of China, 31(1), 125-137.
18. Abdoos, H., Tayebi, A., & Bayat, M. (2018). Prediction of Wear Behavior in Porous Sintered Steels: Artificial Neural Network Approach. Powder Metallurgy Progress, 18(2), 111-115.
19. Hasan, M. S., Kordijazi, A., Rohatgi, P. K., & Nosonovsky, M. (2022). Machine learning models of the transition from solid to liquid lubricated friction and wear in aluminum-graphite composites. Tribology International, 165, 107326.
20. Ulas, M., Altay, O., Gurgenc, T., & Özel, C. (2020). A new approach for prediction of the wear loss of PTA surface coatings using artificial neural network and basic, kernel-based, and weighted extreme learning machine. Friction, 8(6), 1102-1116.
21. Hulipalled, P., Algur, V., & Lokesha, V. (2022). An Approach of Data Science for the Prediction of Wear Behaviour of Hypereutectoid Steel. Journal of Bio-and Tribo-Corrosion, 8(3), 1-7.
22. Al-Nasser, A. D., & Radaideh, A. (2008). Estimation of simple linear regression model using L ranked set sampling. Int. J. Open Problems Compt. Math, 1(1), 18-33.

23. Aydin, F., & Durgut, R. (2021). Estimation of wear performance of AZ91 alloy under dry sliding conditions using machine learning methods. Transactions of Nonferrous Metals Society of China, 31(1), 125-137.
24. Brereton, R.G., Lloyd, G.R. (2021). Support vector machines for classification and regression. Analyst. 135(2), 230–267.
25. Bermejo, S., & Cabestany, J. (2000). Adaptive soft k-nearest-neighbour classifiers. Pattern Recognition, 33(12), 1999-2005.
26. Hashemitaheri, M., Mekarthy, S. M. R., & Cherukuri, H. (2020). Prediction of specific cutting forces and maximum tool temperatures in orthogonal machining by support vector and Gaussian process regression methods. Procedia Manufacturing, 48, 1000-1008.
27. Hossain, M. R., & Timmer, D. (2021). Machine Learning Model Optimization with Hyper Parameter Tuning Approach. Global Journal of Computer Science and Technology.
28. Pannakkong, W., Thiwa-Anont, K., Singthong, K., Parthanadee, P., & Buddhakulsomsiri, J. (2022). Hyperparameter Tuning of Machine Learning Algorithms Using Response Surface Methodology: A Case Study of ANN, SVM, and DBN. Mathematical Problems in Engineering, 2022.
29. Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295-316.
30. Syarif, I., Prugel-Bennett, A., & Wills, G. (2016). SVM parameter optimization using grid search and genetic algorithm to improve classification performance. TELKOMNIKA (Telecommunication Computing Electronics and Control), 14(4), 1502-1509.
31. Montgomery, D. C. (2017). Design and analysis of experiments. John wiley & sons.
32. Kong, D., Chen, Y., & Li, N. (2018). Gaussian process regression for tool wear prediction. Mechanical systems and signal processing, 104, 556-574.

Copyright: *©2022: Poornima Hulipalled. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*