

Intelligent Prediction of Wear Loss Based on Ensemble Learning in Dual-Phase Steel

Poornima Hulipalled

Department of Computer science,
Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India
Email: poornima.algur@gmail.com

V Lokesha

Department of Mathematics,
Vijayanagara Sri Krishnadevaraya University, Ballari, Karnataka, India
Email: v.lokesha@gmail.com

(Received January 23, 2023, Revised March 16, 2023, Accepted May 15, 2023)

Abstract: Dual-phase (DP) steels are widely used in the automotive and aerospace sectors for their excellent tensile strength, ductility, and wear resistance. Experimental determination of wear loss is expensive and time-consuming. These limitations can be eliminated by using data-driven machine-learning approaches. Wear test experiments were conducted to evaluate the volumetric wear loss of DP steels under varying conditions of normal pressure, sliding speed, and sliding distance; this data was then used as input for machine learning (ML) models to forecast wear loss. The primary purpose of this work is to develop efficient ensemble machine learning (EML) models to make predictions regarding the amount of wear loss experienced by dual-phase steel. Random forest (RF), Ada boost (AdaB), and gradient boost (GB) were used to predict wear-loss quantities. The dataset collected from the wear loss experiment is preprocessed and then split into training and testing sets in a 70:30 ratio. The EML models were implemented along with adopted optimization techniques like grid search and 10-fold cross-validation to get the best numerical values for the parameters of ensemble algorithms. EML models have achieved their intended accuracy through a process of empirical tuning of their parameters using k-fold validation and grid search optimization. GB has outperformed others with the highest model accuracy and the fewest statistically significant errors. As a result, the ELM technique shows great promise in aiding the manufacture of wear-resistant components for a wide range of applications.

Keywords: dual-phase steel, ensemble machine learning, hyperparameter optimization, wear loss.

1. Introduction

Tribology focuses on the study of the frictional and sliding interactions that occur between moving and sliding surfaces with the goal of understanding and managing these interactions¹⁻². Significant economic losses are incurred in the automotive sector owing to wear and friction on mechanically powered machine parts³. The increased demand for light weighting in manufacturing motivated more research into improved high-strength steels. Dual-phase (DP) steels with high strength and great formability provide the distinct potential of combining weight reduction with better passive safety⁴. DP steel is commonly used in automobile parts such as wheel discs, rims, bumpers, bodywork, and bumpers⁵. The most typical failure mode for such materials is wear⁶. Consequently, it is essential to comprehend the wear rate across various operating parameters.

Numerous studies have been published on the wear rate of ultrahigh-carbon steels under various operating parameters. In contrast, surveyed the wear behavior of hypereutectoid steels at various working conditions⁷. Gao⁸ looked at how full pearlite wears on bimetallic hypereutectoid steels at ambient temperature with and without the application of thermomechanical treatments.

Several experimental studies were carried out to examine the wear performance of hypereutectoid steels under different operating circumstances. However, this analysis is time-consuming. As a result, machine learning algorithms have been adopted to estimate the wear rate of materials in order to minimize the expenses associated with experimentation. Methods of machine learning (ML) have been the subject of a significant amount of research in recent years, and their influence can be observed in a wide number of fields. For instance, the amount of wear for coated ferroalloy at different operating conditions was predicted based on ML algorithms⁹. Siqui¹⁰ looked at how the ML approach can be used in the field of materials engineering as well as the functional sector. The linear regression (LR) method was utilized so that an estimation of the amount of wear loss could be provided¹¹. Furthermore, we uncover numerous common themes related to the application of ML techniques, such as material characteristics¹², crystal structures¹³ and phase diagram predictions¹⁴, which greatly speed the discovery of new materials via a data-driven material research approach. Taking into account the available data, ML methods typically display outstanding performance when tackling the challenging multidimensional exponential relationship between input and output variables.

In this research, we used the machine learning algorithms RF, AdaB, and GB to forecast the wear loss of DP steel in dry sliding conditions.

2. Materials and Methods

This section includes the methodology of the work carried out. The following procedures form the methodology of this study: data collection, dataset pre-processing, ensemble model implementation, evaluation, optimization, and model prediction. Figure 1 shows a graphic representation of the entire workflow, which is followed by more detailed explanations.

2.1 Data collection and pre-processing

The DP steel material sample was prepared with 0.92 percent carbon, and the steps taken to achieve this concentration are described in detail¹⁵. Gathering appropriate data is a critical initial step in developing an effective data-driven ML system. Tribological experimentation requires a significant investment of both time and finances, which makes it difficult to carry out a large number of tests. In order to address this problem, we used the pin-on-disc wear testing equipment to conduct tests and gathered a useful dataset of the wear behavior of DP steel. Figure 2 shows the wear test experimental setup. The dataset contains various input features such as sliding speed, normal pressure, and sliding distance, with wear rate serving as the output parameter.

Data preprocessing is a major step to be taken before implementing the EML models, hence the data was cleaned, separated into training and testing sets (70:30), transformed, and standardized. We used Spyder, an open-source IDE for Python (version 3.7), to analyze EML algorithm performance.

2.2 Ensemble Machine learning models

In supervised ML, regression analysis quantifies the input-output relationship¹⁶. In this work, the below-mentioned EML algorithms were used to forecast DP steel's tribological behaviour based on input and output parameters.

2.2.1 Random Forest

The supervised learning method known as RF employs a plethora of decision trees. The foundational idea behind RF is to combine several decision trees into a single one to determine the final output, rather than relying on the results of individual trees¹⁷.

2.2.2 AdaBoost

The adaptive boosting algorithm is an eminent machine learning algorithm introduced by Freund and Schapire which works on the concept of boosting. It is also identified as an ensemble learning algorithm¹⁸ which reduces variance and bias under two classifications namely bagging-based and boosting-based, respectively. This method is called iteratively until all the features are classified correctly.

2.2.3 Gradient Boost

Gradient Boosting Machines was developed by Friedman which trains many models in a gradual, additive, and sequential manner. Its goal is to reduce the model's loss by employing a gradient descent-like approach to add weak learners. To forecast the outcome, it employs several additive functions¹⁹.

2.3 Evaluation metrics

Three evaluation criteria were used, including R-squared (R^2), the mean absolute error (MAE), and the root mean square error (RMSE), to estimate the performance of EML models which can be calculated as:

$$(2.3.1) \quad \text{MAE} = \frac{1}{n} \sum_i^n |y_i - x_i|$$

where n is number of data points, y_i is prediction and x_i is true value.

$$(2.3.2) \quad \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - Y_i}{Y_i} \right)^2}$$

Where n is the sample points, y_i is the predictive value and Y_i is the experimental value.

$$(2.3.3) \quad R^2 = 1 - \frac{\sum_i (s_i - \hat{s}_i)^2}{\sum_i (s_i - \bar{s}_i)^2}$$

where s_i is variable to be predicted, $\hat{s}_i$ is the predicted value and $\bar{s}_i$ is total variability in s .

2.4 Parameter Optimization

In practice, the ML models achieve their desired accuracy through a process of parameter optimization based on the availability of data and the application area. Hyperparameters customize the model training. They influence a model's learning and reduce test dataset error²⁰.

2.4.1 Grid search (GS)

In this work, we utilize GS optimization technique to tune hyperparameters. We start by making a grid of searchable parameters, and then it generates the parameters that are most suitable for our data.

2.4.2 k-Fold Cross-Validation

An interpolation technique known as cross-validation can be utilized to increase the accuracy of the evaluation of EML models based on a limited amount of available data. The first fold is a validation set, and indeed the remaining $k-1$ folds are fit. In this work, we performed an optimization on the value of k equal to 10.

The detailed explanation of GS and k -fold cross-validation is discussed in previous work²¹.

We investigated EML models with GS in conjunction through ten-fold cross-validation to determine the individual parameter values that can result in the best performance. Table I records the significant parameter values of EML models.

3. Results and Discussion

The efficient EML models were developed through the use of grid search optimization and 10-fold cross-validation. The performance of these models was evaluated through the utilization of R^2 , MAE, and RMSE. In this part, we will present the outcomes of an EML analysis that was conducted in order to anticipate the amount of wear loss in dual-phase steel. Table II presents the performance of EML models in wear loss prediction. A high level of satisfaction is indicated by an R^2 value between 0.92 and 0.99. The statistically significant minimal error values for both MAE and RMSE range from 0.065 all the way up to 0.49 is observed.

EML regression models are shown in Fig.3, together with their accuracy throughout training and testing, for both the experimental and predicted wear rates. The training and testing data are represented by red and blue scatter points, the predictions for training and testing are represented by red and blue dashed lines, and the solid black line indicates the model that provides the greatest fit to the training and testing data. Fig.4 compares the two-evaluation metrics for three EML models on the training (cyan) and test (blue) dataset. The performance of ML models is ranked in ascending order as AdaB, RF, and GB.

In contrast to the three EML models, the GB model has the ability to make the most precise forecasts. On the other hand, AdaB provides the least accurate predictions, despite the fact that its error values are significantly lower. AdaB has a minimal set of hyperparameters that can be changed to improve model performance; this could explain its poor accuracy.

4. Conclusion

1. Machine learning has made a significant contribution to predicting the wear loss of dual-phase steel.
2. Ensemble machine learning model has produced some encouraging results in terms of the accuracy of its wear loss prediction in dual-phase steel.
3. GB has performed significantly better than the rest of the group in both the training and the testing phases, achieving the greatest accuracy and a significantly lower error rate than the others.

Table I : Optimized parameters

Model name	Selected parameters
RF	n_estimators=150, max_features=3, max_depth=4
AdB	learning_rate=0.25, random_state=21, loss='linear'
GB	criterion='friedman_mse', n_estimators=100, max_depth=4

Table II: Performance of ensemble learning models

Models	Dataset	Evaluation criterion		
		R ²	MAE	RMSE
RF	Training	0.980	0.311	0.496
	Testing	0.97s9	0.245	0.439
AdB	Training	0.922	0.313	0.378
	Testing	0.916	0.253	0.320
GB	Training	0.994	0.077	0.108
	Testing	0.978	0.065	0.091

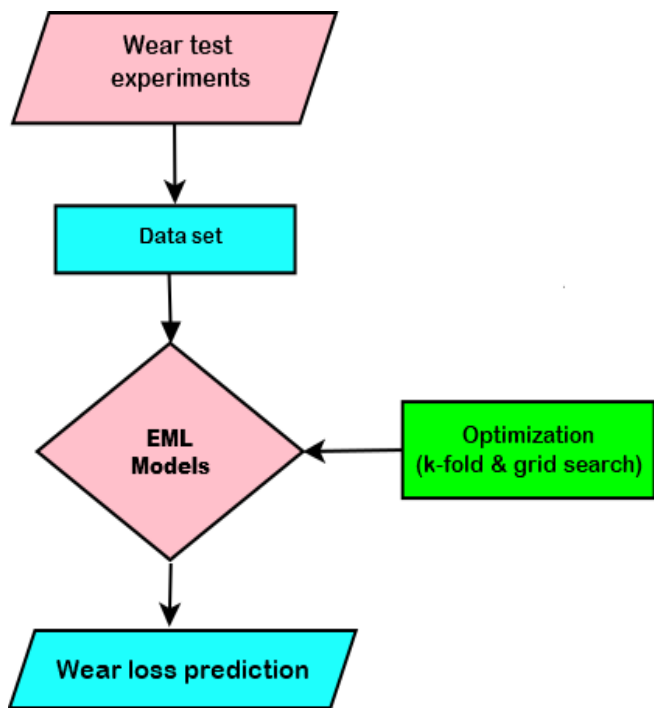


Figure 1: Methodology

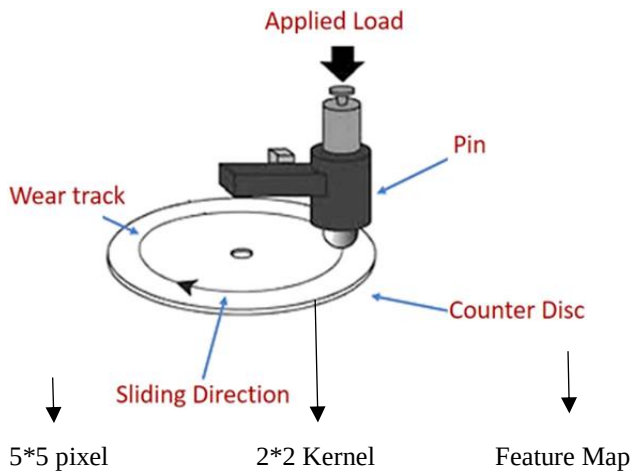


Figure 2: Wear test experiment

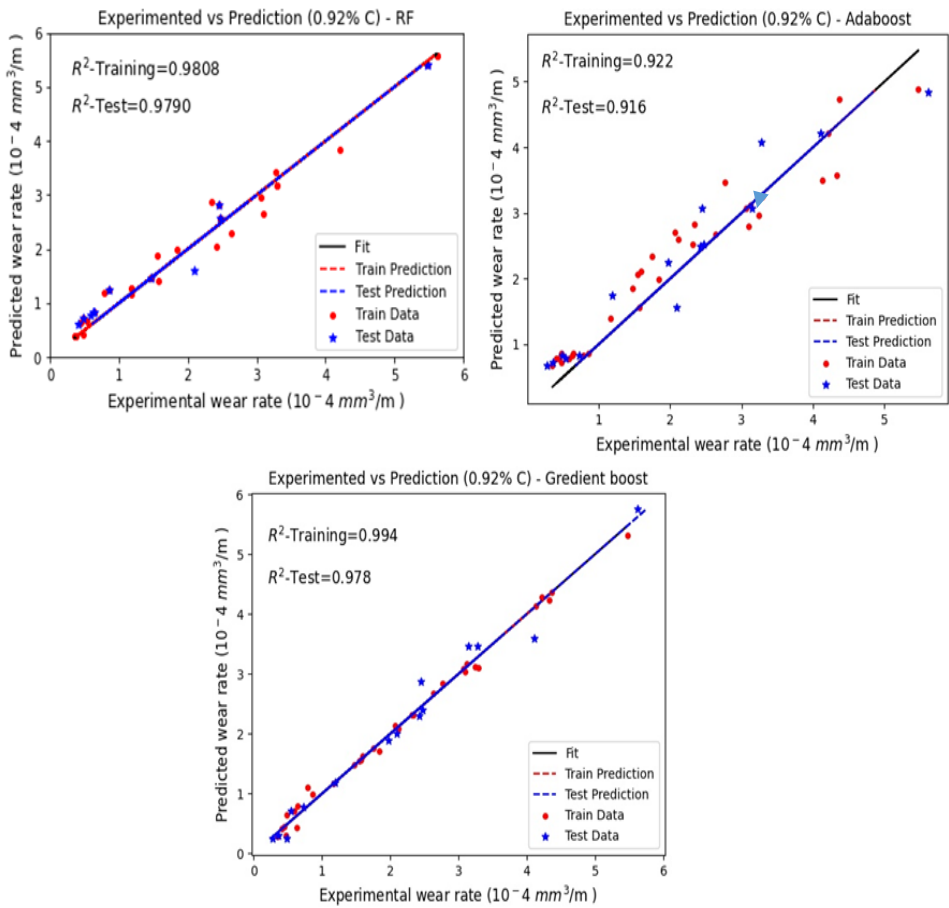


Figure 3: Predicted wear rate by EML models

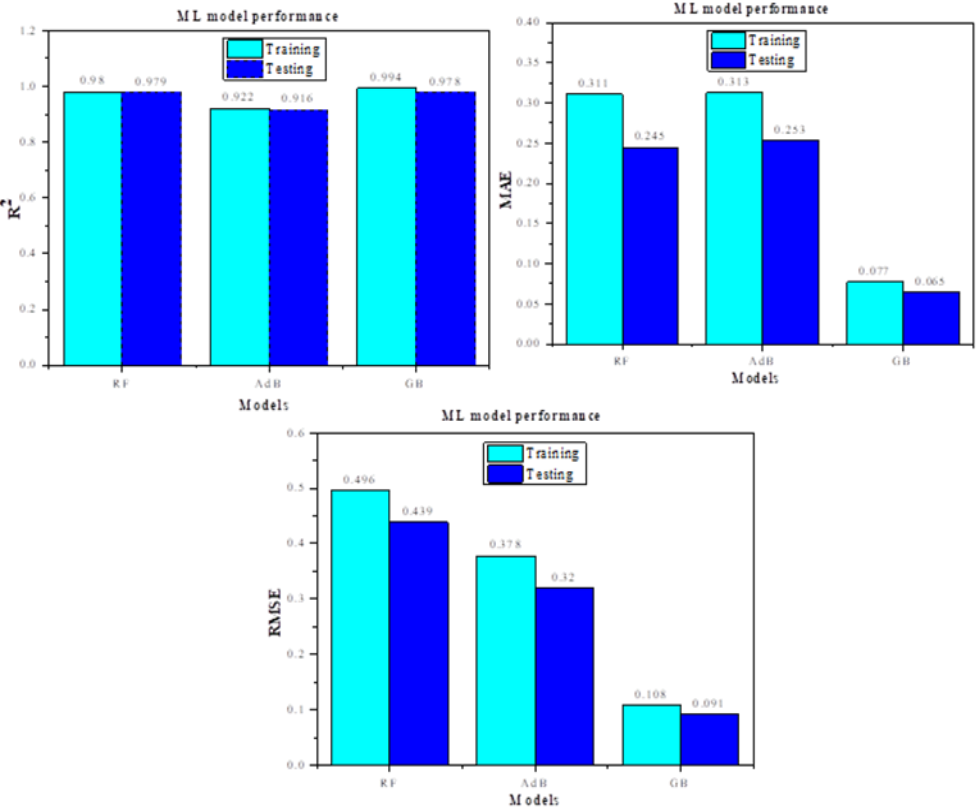


Figure 4: Evaluation metrics by EML models

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

References

1. G Britain; Education and research: A Report on the Present Position and Industry’s Needs, Lubrication (Tribology), H.M. Stationary Office, London, 1966.
2. D Dowson; History of Tribology, 2nd edition, Wiley, 1998.
3. M Ulas, O Altay, T Gurgenc, C Ozel; A new approach for prediction of the wear loss of PTA surface coatings using artificial network and basic, kernel-based and weighted extreme learning machine, *Friction*, 8 (2020), 1102-1116.
4. P I Menezes, M Nosonvsky, S P Ingle, S V Kailas, M R Lovell; *Tribology for scientists and engineers*, Springer, 2013.

5. Amit K Gupta, Deep Narayan Mishra; An experimental investigation of the effect of carbon content on the wear behavior of plain carbon steel, *IJSR*, 2(7) (2013), 222-224.
6. M Sharanabasappa, V R Kabadi, P B Shetty and Veerabhadrapppa Algur; Dry Sliding Wear Behaviour of Hypereutectoid Steel Under the Influence of Microstructures, Sliding speeds, and Normal pressures, *Int Journ of Mech Engg & Robot Reser*, 4(2) (2015), 1-12.
7. Etsuo Marui, Norihiko Hasegawa, Hiroki Endo, Katsuhito Tanaka, Taiji Hattori; Research on the wear characteristics of hypereutectoid steel, *Wear* 205 (1997), 186-199.
8. X Gao, Z Jiang, D Wei, B Kosasih, S Jiao, D Chen; Dry sliding wear behaviour of full pearlite obtained by cladding low carbon steel to hypoeutectoid steel, *Adv Mater Res* 1017 (2014), 147-153.
9. Osman Altay, Turan Gurgenc, Mustafa Ulas, Chian Ozel; Prediction of wear loss quantities of ferroalloy coating using different machine learning algorithms, *Friction*, 8(1) (2020), 107-114.
10. S Siqi, Jian Gao, Yue Liu, Yan Zhao, Qu Wu, Wangwei Ju, Chuying Ouyang, Ruijuan Xiao; Multi-scale computation methods: Their applications in lithium-ion battery research and development, *Chin, Phys. B* 25(1) (2016), 018212.
11. J.C.A Batista, C. Godoy, A. Matthews; Micro-scale abrasive wear testing of duplex and non-duplex (single-layered) PVD (Ti, Al) N, TiN and Cr-N coatings, *Tribology International*, 35(2002), 363-372.
12. Veerabhadrapppa Algur, Poornima Hulipalled, V Lokesha, Madeva Nagaral, V Auradi; Machine Learning Algorithms to Predict Wear Behavior Modified ZA-27 Alloy of Under Varying Operating Parameters, *Journal of Bio-and Tribo-Corr*, 8(1) (2022), 1-10.
13. K. Tsutsui, H. Terasaki, T. Maemura, K. Hayashi, K. Moriguchi, S. Morito; Microstructural diagram for steel based on crystallography with machine learning, *Computational Materials Science* 159 (2019), 403-411.
14. Ling Qiao, Zhonghong Lai, Yong Liu, Aorigele Bao, Jingchuan Zhu; Modelling and prediction of hardness in multi-component alloys: A combined machine learning, first principles and experimental study, *Journal of Alloys and Compounds*, 5 (2021), 853.
15. Veerabhadrapppa Algur, V. R. Kabadi, S. M. Ganeshari, V. R. Chavan; Effect of Mn content on tribological wear behaviour of ZA-27 alloy, *Materials today: Proceedings* 4 (2017), 10927-10934.
16. Poornima Hulipalled, Veerabhadrapppa Algur, V Lokesha; An Approach of Data Science for the Prediction of Wear Behaviour of Hypereutectoid Steel, *Journal of Bio-and Tribo-Corr*, 8(1) (2022), 1-10.

17. Md Syam Hasan, Amir Kordijazi, Pradeep K Rohatgi, Michael Nosonovsky; Tribo informatic modeling of dry friction and wear of aluminum base alloys using machine learning algorithms, *Trib Intern*, 161 (2021), 1-11.
18. Fatih Aydin, Rafet Durgut; Estimation of wear performance of AZ91 alloy under dry sliding conditions using machine learning methods, *Transc of Nonfer Met Soc of Ch*, 31(1) (2021), 125-137.
19. S Bermejo, J Cabestany; Adaptive soft k-nearest-neighbour classifiers, *Pattern Recogn*, 33 (2005), 1999.
20. Li Yang, Abdallah Shami; On hyperparameter optimization of machine learning algorithms: Theory and practice, *Neurocomputing*, 415 (2020), 295-316.
21. Poornima Hulipalled, V. Loksha; Optimized Machine Learning Algorithms to Predict Wear Behavior of Tribo-Informatics, *Journal of Robotics and Automation Research*, 3 (2022), 316-321.