Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2013) 21: 1440 – 1454

doi:10.3906/elk-1108-19

Anticipating the friction coefficient of friction materials used in automobiles by means of machine learning without using a test instrument

Mustafa TİMUR,[1] **Fatih AYDIN**[2,*]

[1]Machinery Program, Vocational School of Technical Sciences, Kırklareli University, Kirklareli, Turkey
[2]Computer Programming Program, Vocational School of Techni50cal Sciences, Kırklareli University, Kırklareli, Turkey

Received: 06.08.2011 • **Accepted:** 14.05.2012 • **Published Online:** 12.08.2013 • **Printed:** 06.09.2013

Abstract: The most important factor for designs in which friction materials are used is the coefficient of friction. The coefficient of friction has been determined taking such variants as velocity, temperature, and pressure into account, which arise from various factors in friction materials, and by analyzing the effects of these variants on friction materials. Many test instruments have been produced in order to determine the coefficient of friction. In this article, a study about the use of machine learning algorithms instead of test instruments in order to determine the coefficient of friction is presented. Isotonic regression was selected as the machine learning method in determining the coefficient of friction. The correlation coefficient between the results of isotonic regression algorithms and the results taken from the test instruments was measured as 0.9999 and the root mean squared error was 0.0014 in the experiments conducted. Selection of the number of optimum samples was enabled by taking bias–variance tradeoff into account, and this increased the performance of the classifier in use.

The target of this study was to prevent the practice of time-consuming test activities by using machine learning methods instead of test instruments in determining the friction coefficient. This presents a solution for decreasing the factors of time and cost.

Key words: Machine learning, isotonic regression, bias–variance tradeoff, friction

1. Introduction

Although friction absorbs a great deal of useful work energy, it is often unnoticed in our daily life. For instance, friction is present in many moving surfaces, ranging from walking or moving a car to handling a pen or blinking our eyes [1]. In a general sense, friction is defined as the resistance shown by the surfaces that are in contact and the objects moving on each other or towards each other [2,3]. There are a wide number of theories that try to explain friction. Among these, the one that is closest to reality is the that of Bowden and Tabor. According to their theory, friction results from 2 basic factors. The first of these emerges in the real contact part in the adhesion. In this part, source links come into existence and the power that makes it possible to dismantle these links is Fs. The other is the influence of (S) power that the coign on the hard surface makes on the soft surface at the time of friction [4].

The systems that enable transformation of the motion energy of vehicles into heat via friction to get the motion under control are called braking systems. Lining at the braking process is used in order to slow or stop the skidding of the mating face between the lining itself and the object that it tries to brake in a controlled way

*Correspondence: fatih.aydin@kirklareli.edu.tr

[5]. Today, instead of the asbestos-based linings, which are sensitive to high temperatures and threaten human health, the goal is to produce linings that are manufactured through powder metallurgy, are resistant to high temperatures, and do not threaten human health [6].

Coefficient friction and wearing rates according to the pressure, velocity, and temperature values have been obtained by doing experiments of sample linings prepared numerously at braking conditions specified in valid standards and at the test instrument designed with the aim of determining the coefficient friction of braking linings [7]. That the object does not accelerate while determining the friction coefficient, or that it moves at a stable speed, is a result of this (F = ma = 0) [8]. The friction coefficient (u) is divided into 2 according to TS 555 as hot and cold coefficient numbers. The cold coefficient number is the arithmetical average of friction coefficients measured at 100, 150, and 200 °C under 1050 kPa pressure. The hot friction coefficient is, however, the arithmetic average of friction coefficients measured at 350 and 400 °C under 300, 350, 1050, and 3000 kPa pressures [9].

The temperature occurring at the time of friction as a result of the unevenness and absorption of the surfaces highly affects the friction coefficient, while the increase of the temperature causes the decrease of the friction coefficient in the material and the system does not work [10]. For this reason, after the necessary grinding over the surface in the experimental machine, rubbing processes are done; the surface is rubbed with a rubber whose grain size is 320 and powders are cleaned with dry air or equivalent material that does not leave behind fiber. For the friction materials to be able to offer a good friction surface, in the experimental machines where the tests are made, generally campanes and disks made of perlitic cast iron are used. Their Brinell hardness values are desired to be between 170 and 255 HB (kg/mm^2). Friction must be between related surfaces and certain values. As an example, because of the lack of friction in car brakes, the car cannot stop at the desired distance and time or blockage of the car's brakes occurs because of the over-friction [11].

Obtaining the data from the test instrument produced requires much time. Because of such negative factors, the activities aimed to be done with the test instrument can be done instead by using machine learning algorithms.

We can now save vast amounts of data thanks to the improvements in information technologies. Every moment, millions of data reach centers where data are kept from supermarket checkout counters, ATMs, POS machines, and e-trade applications. It is not possible to analyze and process such vast data by hand. In order to find a solution for these problems, the machine learning method was developed and still continues to be used. Machine learning methods try to find the most suitable model for the data using the data from the past. They analyze the new data according to this method. Different applications have different expectations from the analysis. It is possible to classify machine learning methods according to those expectations [12].

2. Materials and methods

2.1. Tools used in experiments

Tools used in the study presented in this paper include a test instrument [7] and WEKA Software Developer, version 3.7.1 [13].

2.2. Feature selection

Feature selection is one of the most crucial steps of many pattern recognition and artificial intelligence problems [14]. One of the factors affecting the decision processes of the machine learning algorithm is whether the qualities are suitable or not. When selecting a good attribute subset, there are 2 fundamentally different

approaches. One is to make an independent assessment based on general characteristics of the data; the other is to evaluate the subset using the machine learning algorithm that will ultimately be employed for learning [15]. In our research, we used both approaches in the selection of quality. Before the learning stage is started, using the first approach, the qualities that could help compose good training data so that the learning-based system can make influential decisions were assessed. With the second approach, the results of the machine learning algorithms that were used were evaluated. Making use of the second approach, at the same time, the utility theory approach in the selection of the qualities was applied. Utility theory is concerned with making rational decisions when we are uncertain about the state [12]. We can express this as follows: let the quality that we observed beforehand be x and let S position be composed of the following detailed situation: S_k, k = 1, ..., n. According to this, the probability of an S_k that is known to belong to x quality is counted as $P(S_k|x)$. Let us define the decision motion that designates the x quality to the S_k position as α_i, and our utility function as U_{ik}. Thus, the expected utility is:

$$EU(\alpha_i|x) = \sum_k U_{ik} P(S_k|x). \tag{1}$$

If we suppose that the motion that maximizes the expected utility is α_i, the expected utility of α_i for x quality is shown as in Eq. (2):

$$EU(x) = \max_i \sum_k U_{ik} P(S_k|x). \tag{2}$$

Let our new quality added into the quality cluster be y. The expected utility is then demonstrated as in Eq. (3):

$$EU(x, y) = \max_i \sum_k U_{ik} P(S_k|x, y). \tag{3}$$

According to this, if EU(x, y) > EU(x), y quality can be said to be a remarkable quality. In other words, if the expected utility of the new quality added into the quality cluster is more than the expected utility of the former quality cluster, then the new quality can be said to be useful.

Maximizing the joint dependency with a minimum size of variables is generally the main task of feature selection. For obtaining a minimal subset while trying to maximize the joint dependency with the target variable, the redundancy among selected variables must be reduced to a minimum [16]. If any newly added quality increases EU with the selection of the quality carried out according to EU, the maximum relation between qualities (maximum relevance) is provided. If EU does not increase, there is no need to choose a quality. Thus, the minimum surplus (minimum redundancy) principle is provided and unnecessary qualities are left out.

2.3. Preparation of training data by using a testing instrument

The testing instrument is designed in such a way that reliable results can be obtained through certain standards. Every detail is taken into consideration in order that experimental stages can be done practically, either on designation or after the start of production. The coefficient friction of brake linings under different factors such as speed, heat, and pressure is investigated. A total of 1050 training data points were acquired from the testing instrument as a result of the experiments.

There is electrohydraulic system in the engine that can provide a brake pressure system and the friction coefficient experimental pressures stated in TS 555. With the control of the computer program and electronic

recharge equipment in the system, a closed cycle was composed. While running the system for the first time, the desired system pressure was detected by the electronic control card, and it gave way to electrohydraulic on/off and proportional valves. Via the closed cycle constituted by the operation of the system, it was tried to keep the feedback data and system pressure stable. In the determination of the friction coefficient, calculation is done by considering the system pressure at that time. The friction speed is determined by the number of revolutions per minute entered into the system [17]. The transfer of circular motion coming out of the electric motor to the disk through the shaft is realized thanks to an invertor. These cycles are between 0 and 1400 1/d.

In order to enable the 1050 kPa and 3000 kPa balata surface pressure described in TS 9076 and spin brake disk in the experimental mechanism, a 1400 d/d 3-phase electric motor of 5.5 kW power was chosen. The circular motion of the electric motor was arranged to provide 6 m/s speed in the reference caliber where the erosion experiment would be carried out using a belt and pulley and spinning the brake disk at 680 d/d in the unloaded situation [18].

In all of the experiments, the experiment initiation temperature was taken as 50 °C and the experiments were made under 10.5 MPa pressure. Temperature showed changes faster or slower according to the friction coefficients of these brake linings. Again, the friction surface temperature of the linings rose to 300 °C and oscillation began at this temperature change. When the temperature change enters this process, by becoming a part of the activity, the manual control and electrical heater provide a heat increase and enable the possibility of watching the changes seen at up to 400 °C of friction coefficient stated in TS 555. The friction coefficient of linings under the influence of such factors as different cycles, temperatures, and pressures was examined.

2.4. Machine learning

Inductive machine learning algorithms can learn patterns from labeled data, i.e. cases that have a known outcome [19]. In knowledge discovery, machine learning is most commonly used to mean the application of induction algorithms, which is one step in the knowledge discovery process [20]. There are very different machine learning approaches for pattern learning on data [21]. One of these approaches is regression. Regression is named for the problems in which output is a numerical value inside training data.

2.4.1. Classification algorithm used in the experiment: isotonic regression

The isotonic regression classifier learns an isotonic regression model. It picks the attribute to the results in the lowest squared error. Missing values are not allowed. It can only deal with numeric attributes. It considers the monotonically increasing case as well as the monotonically decreasing case [22]. Isotonic regression is also sometimes referred to as monotonic regression. Isotonic regression is used when the direction of the trend is strictly increasing, while monotonic regression could imply a trend that is either strictly increasing or strictly decreasing [23].

In simple regression we assume a linear relationship between a predictor $\mathrm{p} = (\mathrm{p}_1, ..., \mathrm{p}_i, ..., \mathrm{p}_n)$ and a response $\mathrm{y} = (\mathrm{y}_1, ..., \mathrm{y}_i, ..., \mathrm{y}_n)$. Note that for the predictors we use p instead of the common notation x since later on we embed this algorithm into a convex programming problem where the target variables are typically denoted by x. However, the loss function to be minimized is a least squares problem of the form:

$$f(\alpha, \beta) = \sum_{i=1}^{n} w_i (y_i - \alpha - \beta p_i)^2 \rightarrow \min, \tag{4}$$

where α and β are the regression parameters and w_i represents some optional observation weights. Extensions can be formulated in terms of polynomials or other nonlinear parametric regression specifications. In many situations the researcher has no information regarding the mathematical specification of the true regression function. Rather, it can assume a particular shape that can be characterized by certain order restrictions. Typically, this involves that the y_i values increase with the ordered p_i values. Such a situation is called isotonic regression [24].

Suppose that P is the finite set $\{p_1, p_2, ..., p_n\}$ of the ordered predictors with no ties, i.e. $p_1 < p_2 < \cdots < p_n$. Let y be again the observed response vector and $x = (x_1, ..., x_i, ..., x_n)$ the unknown response values to be fitted. The least squares problem in monotonic regression can be stated as:

$$f(x) = \sum_{i=1}^{n} w_i (y_i - x_i)^2 \rightarrow \min, \tag{5}$$

which has to be minimized over x under the inequality restrictions $x_1 \leq x_2 \leq \cdots \leq x_n$ for isotonic regression [24].

Isotonic regression implements the method for learning an isotonic regression function based on the pair-adjacent violators (PAV) approach. Estimating a piecewise constant monotonically increasing function is an instance of isotonic regression, for which there is a fast algorithm based on the PAV approach. The basic PAV algorithm iteratively merges pairs of neighboring data points that violate the monotonicity constraint by computing their weighted mean – initially, this will be the mean of 0/1 values – and using it to replace the original data points. This is repeated until all conflicts have been resolved. It can be shown that the order in which data points are merged does not affect the outcome of the process. The result is a function that increases monotonically in a stepwise fashion [25].

Isotonic regression is simpler than commonly used statistical methods and it is often used among statisticians [26]. Isotonic regression is a single-sided test used to compare group averages when there are sequential expectations of in-line or categorical data [27]. Isotonic regression is in fact the simplest ordering situation. Simple line hypothesis ($(\mu_1 \leq \mu_2 \leq \mu_3 \leq \mu_4)$ can be put in order not leaving space for decrease. Isotonic regression happens with the comparison and mixture of the data in different groups until the group averages left are paired off at H_2 [26].

2.4.2. Other classification algorithms used in the experiment

The results obtained through using such classifiers as linear regression, least median squared linear regression (LeastMedSq), Gaussian processes, pace regression, simple linear regression, and SMO regression in experiments were compared with isotonic regression. The reasons for using isotonic regression are explained in the experimental results. In addition, the selections of all of the classifiers used in the experiments were determined as assumed.

Linear regression: Uses the Akaike criterion [28] for model selection and is able to deal with weighted instances [29].

LeastMedSq: Implements a least median squared linear regression utilizing the existing WEKA LinearRegression class to form predictions. Least squared regression functions are generated from random subsamples of the data. The least squared regression with the lowest median squared error is chosen as the final model [30]. Classical least squares regression consists of minimizing the sum of the squared residuals. In LeastMedSq, a different approach is introduced, in which the sum is replaced by the median of the squared residuals. The

resulting estimator can resist the effect of nearly 50% of contamination in the data. In the special case of simple regression, it corresponds to finding the narrowest strip covering half of the observations. Generalizations are possible for multivariate location, orthogonal regression, and hypothesis testing in linear models [31].

Gaussian processes: Implements Gaussian processes for regression without hyperparameter tuning. To make choosing an appropriate noise level easier, this implementation applies normalization/standardization to the target attribute as well as the other attributes (if normalization/standardization is turned on). Missing values are replaced by the global mean/mode. Nominal attributes are converted to binary ones. Note that kernel caching is turned off if the kernel used implements Cache Kernel [32].

Pace regression: Under regularity conditions, pace regression is provably optimal when the number of coefficients tends to infinity. It consists of a group of estimators that are either overall optimal or optimal under certain conditions [33].

Simple linear regression: Learns a simple linear regression model. Picks the attribute that results in the lowest squared error. Missing values are not allowed [34].

SMO regression: SMOreg [35] implements the support vector machine for regression [36]. The parameters can be learned using various algorithms. The algorithm is selected by setting the RegOptimizer [37].

2.4.3. K-fold cross-validation

K-fold cross-validation is a method for estimating the accuracy (or error) of an inducer by dividing the data into k mutually exclusive subsets (the "folds") of approximately equal size. The inducer is trained and tested k times. Each time it is trained on the data set minus a fold and tested on that fold. The accuracy estimate is the average accuracy for the k folds [20]. Limitations can be placed on the amount of the data used for training data by using cross-validation [15]. Thus, sorting the unnecessary amount of data out of the training data can be achieved. This prevents the learning algorithm from memorizing instead of learning. We choose the k figure as 10 in our experiments.

2.5. Evaluation criteria

Different criteria are used to learn the performance of algorithms. Root mean squared error (RMSE) and correlation coefficient values are used for the evaluation of the classifiers used in the experiment. These evaluation criteria are explained below.

2.5.1. Root mean squared error

c is the unknown value of a parameter of a distribution, and X is an estimator of this parameter. In this context, it is usual to denote the estimator by $\theta*$ (instead of X) and the value of the parameter by θ_0 (instead of c). ε is called the "estimation error" of $\theta*$. A good estimator is close to θ_0 on average. Just how close is usually measured by the mean of the squared estimation error ε. This quantity is called the mean square error (MSE) of the estimator $\theta*$ [38]. The RMSE is the square root of the MSE.

$$MSE = E\left[(\theta^* - \theta_0)^2\right] \tag{6}$$

$$MSE = Var(\theta^*) + Bias(\theta^*)^2 \tag{7}$$

As can be observed from Eq. (7), a high degree of bias and variance increase the value of the MSE. From this point of view, the RMSE presented by WEKA is very important in the process of the preparation of data.

2.5.2. Correlation coefficient

The correlation coefficient (CC) is a coefficient that indicates relation between independent variables of force and direction. It changes between –1 and +1. The positive values of CC specify a direct directional relation, and the negative values specify a reverse directional relation. If the CC is 0, it indicates there is no relationship between variables. Detailed explanation of CC values is shown in Table 1 as given by Cohen [39].

Table 1. Cohens correlation table.

Correlation	Negative	Positive
Low	–0.29 to –0.10	0.10 to 0.29
Middle	–0.49 to –0.30	0.30 to 0.49
High	–0.50 to –1.00	0.50 to 1.00

3. Classification results and discussion

Let us examine the test results of 7 classifiers used on 1050 training data points. Our criteria at the stage of testing were determined as accuracy rate, RMSE, and model forming. The parameters that the classifiers used in the experiments are taken as the assumed parameters. According to this, when the results in Table 2 are examined, the classifier whose correlation coefficient is the highest (CC = 0.9999), with the lowest error rate (RMSE = 0.0014), is seen to be isotonic regression. Moreover, isotonic regression comes in second place in terms of composing the model within 0.02 s. As the k value at the time of the CV process was chosen as 10, 10 models were formed. Because 0.02 s passes for a single model, the duration for 10 models to be formed will be only 0.2 s because of all of these positive performance values; isotonic regression was thus chosen to compose a model.

Table 2. The comparison of performance criteria of the classifiers.

Classifiers	CC	RMSE	Time taken to build model (s)
Linear regression	0.9997	0.0029	0.1
LeastMedSq	0.9997	0.0029	6.3
Gaussian processes	0.9907	0.0166	14.48
Pace regression	0.9997	0.0029	0.08
Simple linear regression	0.9997	0.0029	0.01
SMO regression	0.9997	0.0029	1.99
Isotonic regression	0.9999	0.0014	0.02

Training data prepared separately for each quality were tried to be learnt using isotonic regression. The statistical results gained after the learning process are shown in Table 3. According to Table 3, learning results on the training data prepared only with power data give better results compared with training data prepared according to other qualities. For this, it can be said that power quality is more dominant compared with pressure and heat qualities.

The number of training data used in the process of learning a classifier is very important. An insufficient amount of training data might cause the classifier to make inexact predictions. The reason for this is that the complexity of the function obtained by using training data is less than the complexity of the real function. In this case, the function created with training data is more complex than the real function. Thus, the prediction of a new test datum might be different from what it is meant to be. That is, a generalization mistake is made. Experiments are done for bias–variance tradeoff by using 1050 examples that are acquired by the testing

instrument in order to prevent this situation. In these experiments, our first criterion for bias–variance tradeoff is the RMSE value. We do the experiments by increasing the number of our training data from 10 to 1050 in order to observe the change in this value. At the end of the experiments, examples used for training are compared with the RMSE value. This comparison is shown in Figure 1. Furthermore, the reason why the number of our training data starts from 10 is that the k value was taken as 10 during cross-validation.

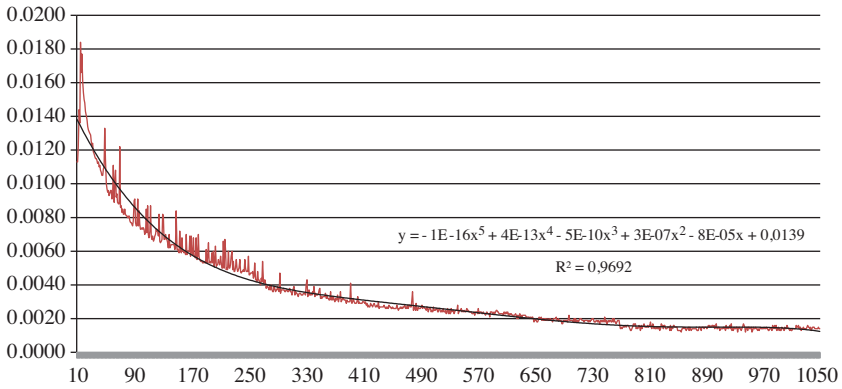

Figure 1. Red line: correlation between RMSE and sample number. Black line: tendency line of the correlation between RMSE and sample number.

Table 3. Statistical information gained as a result of the training data prepared separately according to each quality using isotonic regression.

Attributes	CC	RMSE	Error sample count
Force	0.9999	0.0014	26
Pressure	0.7753	0.0763	1046
Heat	0.9118	0.0496	1043

Analyzing Figure 1, when sample amounts are between 770 and 1050, the RMSE value varies between 0.0012 and 0.0018 and consists of 281 data. The average of these RMSE values is 0.0014. The reason why we used a central tendency scale as average is that all values are close to each other and the difference between min–max values is very small. Our first sample giving amounts equal to or smaller than 0.0014 after the training data including 770 samples is the training data that have 771 samples. This example set's RMSE value is equivalent to 0.0014. That is why we delimited the amounts of sample being used in our training data at 771. We can say that our classifier's performance for the set including 771 samples is very good because, for that set of training data, the CC value is measured as 0.9999. That shows us that there is a perfect harmony between the real value and estimated value. We also indicated the correlation between the RMSE and sample amounts with a tendency line in Figure 1. This indicated that the function of this tendency line is a quadratic polynomial. The reliability of this polynomial's function is indicated with $R^2 = 0.969$. R^2 (the coefficient of determination) is the proportion of variability in a data set that is accounted for by the statistical model [40].

In other words, R^2 gives an idea about how good the function's predictions are. The R^2 value is expressed as the square of the CC and it can be said that the predictions are more successful however closer it is to 1. In this case, we can say that the fourth-degree polynomial function reveals well the relationship between the RMSE and the sample number.

In Figure 2, the correlation between the sample number and the CC has been indicated. According to this graphic, while the component number of the training data is 10, the CC value has been estimated as 0.9270,

and after the sample number has become 769, the CC value permanently stays at 0.9999. That is, as the sample number approaches 1050, it is a CC = 1.0 asymptote.

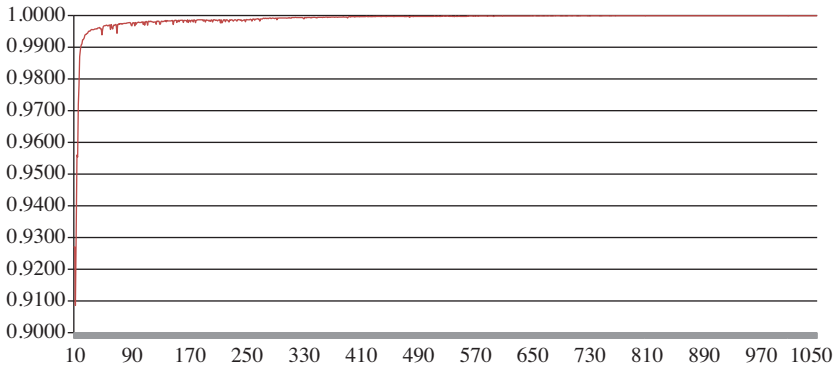

Figure 2. The correlation graphic between CC and sample number.

The difference between real values and estimated values was seen only in 16 samples in these experiments. For the rest of the samples, there is total relevance between real values and estimated values. This shows us that the classifier had a great hit on predictions. The 16 samples that show incongruity are shown in Table 4. When this is analyzed, it can be observed that the error value between real and estimated values is 0.01. This shows us that even among the 16 samples that show incongruity, the error rate is very low.

Table 4. The difference between actual and predicted values.

Instance no.	Actual value	Predicted value	Error
1	0.46	0.45	–0.01
2	0.32	0.31	–0.01
3	0.56	0.57	0.01
4	0.51	0.52	0.01
5	0.44	0.45	0.01
6	0.16	0.17	0.01
7	0.31	0.32	0.01
8	0.45	0.46	0.01
9	0.54	0.53	–0.01
10	0.15	0.16	0.01
11	0.42	0.41	–0.01
12	0.39	0.40	0.01
13	0.45	0.44	–0.01
14	0.09	0.1	0.01
15	0.36	0.37	0.01
16	0.20	0.21	0.01

In Figure 3, the range of the friction coefficient and the samples that have differences between real values and estimated values are shown.

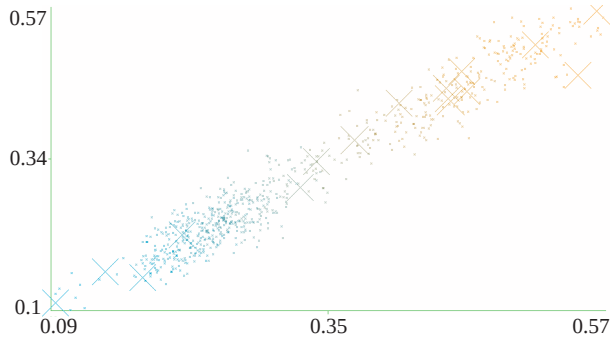

Figure 3. Range of friction coefficients between actual and predicted values. X axis: actual friction coefficient values, Y-axis: predicted friction coefficient values, large crosses: error between values.

Results of the isotonic regression classifier chosen as the learning algorithm are shown in Table 5. In consequence of using isotonic regression, the correlation coefficient between the real and estimated values was estimated as 0.9999. This shows us that the correlation is in a positive direction and at a high degree. At the same time, the facts that the mean absolute error and RMSE values are quite close to 0 and the relative absolute error and root relative squared error rates are quite low also show us that the classifier has excellent predictions with very few mistakes. An increase in the error number is seen in the results of increase of the sample number. This shows that the sample number of 771 is appropriate for the performance of the classifier. When the sample number is 1050, it is observed that the error number increases. In other words, the differences between the predictions that the classifier has made have appeared. This shows us that the variance has increased.

Table 5. The comparison of the classifiers results as a consequence of the fact that the sample number is different.

Sample count	Error count	Mean (MAE) absolute error (MAE)	Root mean squared error (RMSE)	Relative absolute error (RAE)	Root relative squared error (RRSE)	Correlation coefficient (CC)
771	16	0.0002	0.0014	0.2217%	1.1952%	0.9999
1050	26	0.0002	0.0014	0.2217%	1.1952%	0.9999

Twenty-six errors in the predictions made for 1050 training data points are shown in Table 6. It can been seen that the errors that the classifier made ranged from –0.01 to 0.01.

Statistical information about the entrance features and the exit feature used in learning process are shown in Table 7.

The range of heat features is shown in Figure 4. When the range is analyzed, it is seen that heat feature has a range between 49 and 348.

The range of force features is shown in Figure 5. When the range is analyzed, it is seen that force feature has a range between 3.41 and 20.43.

Table 6. The actual and predicted values of the friction coefficient for 1050 training data points.

Instance no.	Actual value	Predicted value	Error
1	0.540	0.530	–0.01
2	0.560	0.570	0.01
3	0.420	0.410	–0.01
4	0.150	0.148	–0.003
5	0.200	0.210	0.01
6	0.390	0.400	0.01
7	0.190	0.200	0.01
8	0.190	0.200	0.01
9	0.090	0.100	0.01
10	0.110	0.120	0.01
11	0.450	0.460	0.01
12	0.150	0.147	–0.003
13	0.150	0.147	–0.003
14	0.310	0.320	0.01
15	0.510	0.520	0.01
16	0.130	0.120	–0.01
17	0.140	0.148	0.008
18	0.150	0.148	–0.003
19	0.440	0.450	0.01
20	0.130	0.140	0.01
21	0.140	0.150	0.01
22	0.460	0.450	–0.01
23	0.170	0.180	0.01
24	0.360	0.370	0.01
25	0.450	0.440	–0.01
26	0.320	0.310	–0.01

Table 7. Statistical information about features.

Features	Distinct value count	Unique value count	Minimum value	Maximum value	Mean	Std. dev.
Heat	207	58	49	348	236.817	77.035
Force	500	323	3.41	20.43	11.031	4.034
Pressure	82	19	2.15	11.04	10.539	0.904
Friction (outcome)	47	3	0.09	0.57	0.307	0.112

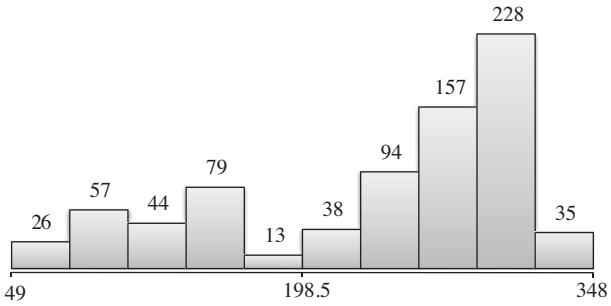

Figure 4. The range of heat features.

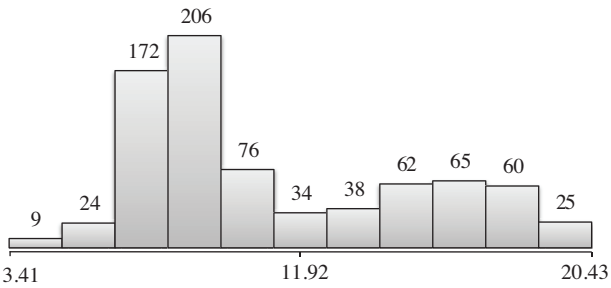

Figure 5. The range of force features.

The range of pressure features is shown in Figure 6. When the range is analyzed, it is seen that pressure feature has a range between 2.15 and 11.04. The central tendency of the range is 10.539. That is, the frequency of the values is between 10 and 11. In Figure 6, this situation is observed.

The range of friction coefficient features is shown in Figure 7. When the range is analyzed, it is seen that the friction coefficient feature has a range between 0.09 and 0.57.

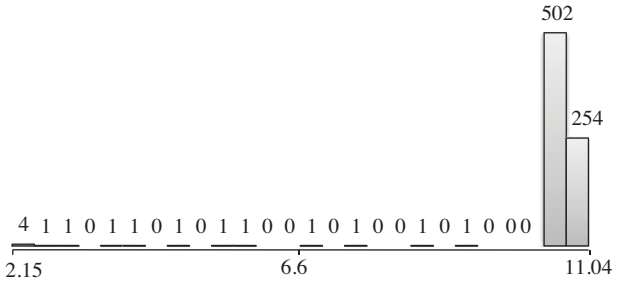

Figure 6. The range of pressure features.

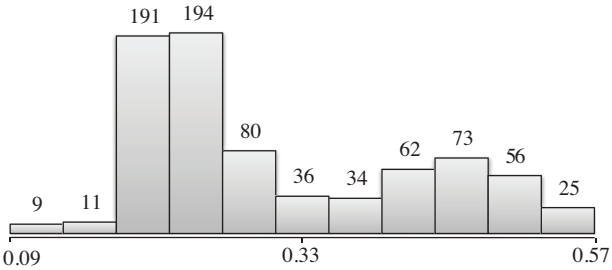

Figure 7. The range of friction coefficient features.

The range between the values of heat, force, and pressure with friction coefficient features is respectively shown in Figure 8, 9, and 10. When Figure 8 is analyzed, it is seen that there is a nonlinear relation between friction coefficient and heat.

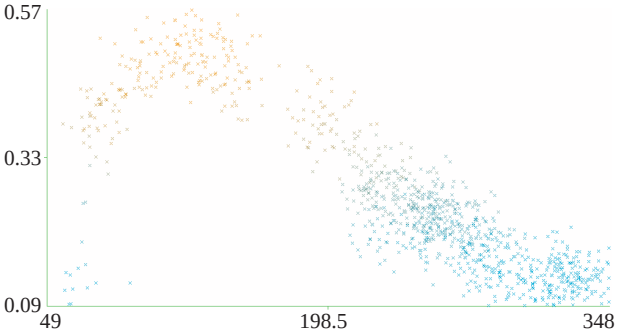

Figure 8. Relation between actual friction coefficient values and heat values. X-axis: heat, Y-axis: friction coefficient.

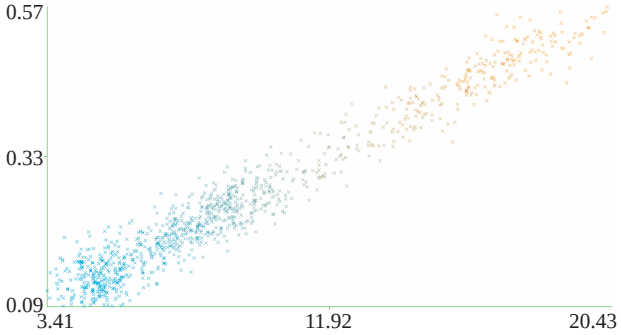

Figure 9. Relation between actual friction coefficient and force values. X-axis: force, Y-axis: friction coefficient.

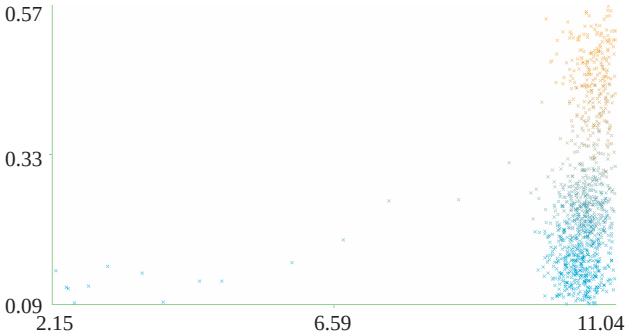

Figure 10. Relation between actual friction coefficient values and pressure values. X-axis: pressure, Y-axis: friction coefficient.

In Figure 9, the relation between force and friction coefficient values is shown. When the relation between the values is analyzed, it is seen that there is a relation close to a linear line between these 2 features.

The graph of the relation between pressure and friction coefficient values is shown in Figure 10. When it is analyzed, it is realized that there is a nonlinear relation between these 2 features.

Relations between other features can also be seen easily after these graphs are analyzed. According to this, it can also be said that there is a nonlinear relation between force–heat and force–pressure features. Furthermore, it can be said that there is a nonlinear relation between heat and pressure. It can be said that there is a relation close to a linear line between only the friction coefficient and force. According to this, it is seen that in problems in which there is a nonlinear relation between features, machine learning algorithms are very good at solving the relations between these features.

4. Conclusion

It was determined by our machine learning work that the desired results can be achieved easily without depending on a test instrument. The friction coefficient was predicted according to speed, pressure, and heat values by using isotonic regression. The correlation coefficient between the predicted friction coefficient and the friction coefficient acquired from the test instrument was 0.9999. This shows that there is a perfect relation between the predicted value and the real value. Meanwhile, the error rate among values is very close to 0 (RMSE = 0.0014), which is a sign that the difference between these 2 values is very low. The optimum sample number necessary for learning was selected from the 1050 data points acquired from the test machine by using RMSE criteria. Thus, bias–variance tradeoff was achieved. The number of samples in the training data prepared according to this was 771. The most important reason for error rates being close to 0 was that the friction coefficient of 755 training data out of 771 was the same as the real friction coefficient. As a conclusion, the cost was totally removed and the time factor was also decreased considerably with the method used.

Acknowledgment

The development of the test instrument used in detecting the friction coefficient was supported by TÜBİTAK under project number 106M006.

References

[1] D. Halliday, R. Resnick, Fiziğin Temelleri, Ayrım Yayınları, Ankara, 1986 (Turkish translation from English: Fundamentals of Physics, Wiley).

[2] F. Yılmaz, "Sürtünme ve aşınma", 9. UluslararasıMetalürji ve Malzeme Kongresi, İstanbul, pp. 229–247, 1997 (in Turkish).

[3] M. Ashby, R.H.D. Jones, Friction and Wear Engineering Materials, Cambridge University Press, Cambridge, 1996.

[4] F.P. Bowden, D. Tabor, The Friction and Lubrication of Solids, Oxford University Press, Oxford, 1964.

[5] C. Filikçioğlu, Taşıt Frenlerinde Balata Çeşidinin Ve Çalışma Şartlarının Frenleme Performansına Etkisi, MSc, Gazi University Institute of Science, Ankara, 1998 (in Turkish).

[6] N. Gemalmayan, "Hava kirliliğinde gizli tehlike asbest", TÜBİTAK Bilim ve Teknik Dergisi, Vol. 16, pp. 32–33, 1983 (in Turkish).

[7] M. Timur, Otomotivde Kullanılan Sürtünme Malzemelerinin Sürtünme KatsayısınıTespit Eden Test CihazıTasarımıve İmalatı, MSc Thesis, Afyon Kocatepe University Institute of Science, Afyonkarahisar, 2007 (in Turkish).

[8] TÜBİTAK, http://www.tubitak.gov.tr, October 2007.

[9] M. Boz, Toz Metalurjisi ile Üretilmiş Bronz EsaslıFren Balata Malzemelerinin Sürtünme-Aşınma Davranışlarının İncelenmesi, MSc Thesis, Gazi University Institute of Science, Ankara, 1999 (in Turkish).

[10] M.E. Uygur, "Balata üretim teknolojisi", 1. Ulusal Toz Metalurjisi Kongresi, Bildiriler Kitapçığı, Gazi University, Ankara, 1996 (in Turkish).

[11] N. Gemalmayan, Sürtünme Malzemelerinin Özelliklerinin Deneysel İncelenmesi, PhD, Gazi University Engineering and Architecture Faculty, Ankara, 1984 (in Turkish).

[12] E. Alpaydın, Introduction to Machine Learning, MIT Press, Boston, 2004.

[13] WEKA, http://www.cs.waikato.ac.nz/ml/weka/, March 2011.

[14] J.G. Zhang, H.W. Deng, "Gene selection for classification of microarray data based on the Bayes error", BMC Bioinformatics, Vol. 8, p. 370, 2007.

[15] I.H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed., Morgan Kaufmann, San Francisco, 2005.

[16] O. Kurşun, C.O. Şakar, O. Favorov, N. Aydın, F. Gürgen, "Using covariates for improving the minimum redundancy maximum relevance feature selection method", Turkish Journal of Electrical Engineering & Computer Sciences, Vol. 18, pp. 975–989, 2010.

[17] TS 555, Karayolu Taşıtları-Fren Sistemleri-Balatalar-Sürtünmeli Frenler İçin, T.S.E., 1. Baskı, Ankara, 1992 (in Turkish).

[18] TS 9076, "Fren Balataları-Malzeme Sürtünme Özelliklerinin Küçük Deney Parçalarıİle Değerlendirilmesi", T.S.E., 1. Baskı, Ankara, 1991 (in Turkish).

[19] R.O. Duda, Hart PE, Stork DG, Pattern Classification, 2nd ed., Wiley-Interscience, New York, 2001.

[20] R. Kohavi, F. Provost, Applications of Machine Learning and the Knowledge Discovery Process, Vol. 30, pp. 271–274, 1998.

[21] C.M. Bishop, Pattern Recognition and Machine Learning, Springer Science+Business Media, Singapore, 2006.

[22] WEKA, http://weka.sourceforge.net/doc.packages/isotonicRegression/weka/classifiers/functions/IsotonicRegression.html, October 2010.

[23] J. de Leeuw, "Monotonic regression", in Encyclopedia of Statistics in Behavioral Science, Vol. 3, pp. 1260–1261, John Wiley & Sons, New York, 2005.

[24] J. de Leeuw, K. Hornik, P. Mair, "Isotone optimization in R: pool-adjacent-violators algorithm (PAVA) and active set methods", Journal of Statistical Software, Vol. 32, pp. 1–24, 2009.

[25] I.H. Witten, E. Frank, M.A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed., Morgan Kaufmann, San Francisco, 2011.

[26] H. Dağlıoğlu, SıralıSeçenek Hipotezlerinde Kullanılan Testler, MSc, Hacettepe University Institute of Science, 2008.

[27] L. Kohorn, "Shoot morphology and reproduction in Jojoba: advantages of sexual dimorphism", Ecology, Vol. 75, pp. 2384–2394, 1994.

[28] C.E. Lance, Akaike's Criterion, Encyclopedia of Statistics in Behavioral Science, John Wiley & Sons, New York, 2005.

[29] WEKA, http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/LinearRegression.html, April 2012.

[30] WEKA, http://weka.sourceforge.net/doc/weka/classifiers/functions/LeastMedSq.html, April 2012.

[31] P.J. Rousseeuw, "Least median of squares regression", Journal of the American Statistical Association, Vol. 79, pp. 871–880, 1984.

[32] WEKA, http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/GaussianProcesses.html, April 2012.

[33] WEKA, http://weka.sourceforge.net/doc/weka/classifiers/functions/PaceRegression.html, April 2012.

[34] WEKA, http://weka.sourceforge.net/doc/weka/classifiers/functions/SimpleLinearRegression.html, April 2012.

[35] S.K. Shevade, S.S. Keerthi, C. Bhattacharyya, K.R.K. Murthy, "Improvements to the SMO algorithm for SVM regression", IEEE Transactions on Neural Networks, Vol. 11, pp. 1188–1193, 1999.

[36] A.J. Smola, B. Schoelkopf, "A tutorial on support vector regression", Journal Statistics and Computing, Vol. 14, pp. 199–222, 1998.

[37] WEKA, http://weka.sourceforge.net/doc/weka/classifiers/functions/SMOreg.html, April 2012.

[38] AI ACCESS, http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_mean_square_error.htm, March 2011.

[39] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed., Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1998.

[40] R.G.D. Steel, J.H. Torrie, Principles and Procedures of Statistics, McGraw-Hill, New York, 1960.