Review

Recent Progress of Machine Learning Algorithms for the Oil and Lubricant Industry

Md Hafizur Rahman 1, Sadat Shahriar 2 and Pradeep L. Menezes 1,*

1 Department of Mechanical Engineering, University of Nevada, Reno, NV 89557, USA; mdhafizurr@unr.edu
2 Department of Computer Science, University of Houston, Houston, TX 77204, USA
* Correspondence: pmenezes@unr.edu

Abstract: Machine learning (ML) algorithms have brought about a revolution in many industries where otherwise operation time, cost, and safety would have been compromised. Likewise, in lubrication research, ML has been utilized on many occasions. This review provides an in-depth understanding of seven ML algorithms from a tribological perspective. More specifically, it presents a comprehensive overview of recent advancements in ML applied to lubrication research, organized into four distinct categories. The first category, experimental parameter prediction, highlights the significant contributions of artificial neural networks (ANNs) in accurately forecasting operating conditions related to friction and wear. These predictions offer valuable insights that aid in forensic preparation. Discriminant analysis, Bayesian modeling, and transfer learning approaches have also been used to predict experimental parameters. Second, to predict the lubrication film thickness and identify the lubrication regime, algorithms such as logistic regression and ANN were useful. Such predictions provide up to 99.25% accuracy. Third, to predict the friction and wear for a given experimental condition, support vector machine (SVM), polynomial regression, and ANN offered an accuracy above 93%. Finally, for condition monitoring for bearings, gearboxes, gear trains, and similar critical situations where regular in-person inspection is difficult, Naïve Bayes, SVM, decision trees, and ANN were utilized to predict the safe life of lubricants. This review highlighted these four aspects with state-of-the-art examples and discussed the current situation and projected future possibilities of lubricant design facilitated by ML techniques.

Keywords: algorithm; friction; machine learning; neural network; wear

1. Introduction

The modern world is dependent on versatile machinery that is susceptible to continuous friction and wear. As a result, numerous lubricants and additives have been developed to reduce the cost associated with friction and wear [1]. As per the literature, effective use of lubricants could contribute from 1% to 1.55% of the yearly gross domestic product (GDP) of the United States [2–4]. Therefore, effective lubrication is extremely important to strengthen the national economy. Lubricants can be categorically divided into three types: liquid, solid, and gaseous [5,6]. Each type has its own application, but overall the purpose is the same: optimizing friction and wear.

Among all lubricants, liquid lubricants have unique functional features that have helped them expand their applicability over many years [7,8]. Solubility, viscosity, heat capacity, wettability, dampening ability, resistance to corrosion, wear, and friction are typical characteristics of liquid lubricants [9–11]. Engineers have added solid particles to liquid lubricants and utilized their synergy to reduce friction and wear [12]. Tribologists have also developed superior lubricants that can provide high thermal stability, exhibit low volatility, and have a broad liquid range [13,14]. Therefore, the list of lubricants and their properties has become innumerable nowadays. As a result, it became difficult even for an expert to choose the best lubricant or simply predict the optimum condition for a chosen
lubricant without going through a series of trials and errors. In addition, oil analysis is important to maintain the reliability of the equipment and enhance its lifespan. However, using conventional techniques, a data analyst can observe the tribological failure and take steps. But if an earlier prediction could be made, that could minimize the breakdown time. Therefore, tribologists have left their footprints in the computational field and found a wide potential to utilize ML algorithms to predict the optimum lubricant for tribological applications [15].

Machine learning is the study of algorithms that can train and improve themselves through observation. It can be carried out in two ways: (1) supervised learning and (2) unsupervised learning. Both of these types have their pros and cons. As a result, they need to be chosen wisely. Due to the complexity of parameters and methods, computer scientists and tribologists are working together to develop unique tools to predict optimum lubricants and optimum parameters to reach superior performance. Therefore, this approach’s success opens up wide potential to minimize the time and money associated with trial and error, which could strengthen the economy to a great extent [15]. At the current status quo, many tribological studies are devoted to the estimation of the marginal life of lubricants [16–19]. In practice, the prediction of the limit state for a lubricant is dependent on the manufacturer’s recommendation. When a pre-estimated service life is finished, an oil change is needed. However, the actual health of oil might be better to go on or could be already much worse even before the pre-estimated timeline. Therefore, to optimize the usage of lubricants, continuous monitoring of oil is a strong tool to save machine life and maintenance costs. By monitoring the performance of a lubrication system in real time, it is possible to detect potential issues before they become more serious problems. This can improve the reliability of the lubricants and the performance of the equipment. Therefore, with the development of accurate and efficient models, ML is poised to revolutionize the industry in the coming years [15]. Companies that adopt these technologies will be better positioned to deliver high-performance lubrication solutions that are optimized for specific applications while also reducing downtime and maintenance costs.

Although there is a lot of hype about ML, the available research on this ML perspective for lubrication is scattered. There is a necessity for a review article in this field focusing on different algorithms for researchers to understand and pursue a suitable approach. In this review, nearly 100 scientific journals have been studied to infer current progression and project future path lines for this growing area. Also, this article will be helpful for a tribologist to understand the basics of ML in order to probe its advantages for lubricant selection. This review will provide state-of-the-art insight into ML approaches for lubrication research.

2. Machine Learning

Machine learning could be simply addressed as a computer program that can learn from experience (“E”), with respect to some tasks (“T”), and performance measures (“P”). As per Tom Mitchell, the performance of ML at tasks in “T”, as measured by “P” improves with experience “E” [20]. ML is a subset of artificial intelligence (AI), which helps computers act without explicit programming and instead utilizes the concept of statistics. ML has been widely utilized to find data patterns to make predictions. Nowadays, ML has advanced from pattern recognition studies and is capable of focusing on the recognition of patterns and regularities in data [21].

2.1. Machine Learning Types and Processes

Machine learning techniques can be categorically divided into supervised learning and unsupervised learning [22,23]. Each of these types consists of different processes and algorithms, as shown in Figure 1.
Figure 1. Categorical classification of machine learning applications as per Mathworks.
In the case of supervised learning, the model has access to both the training data and their labels, and the model can be trained on the labeled data to perform prediction tasks on any unseen data. The supervised learning task can be formulated as a classification or a regression problem [24]. Classification provides discrete-valued output. It is helpful to predict a characterization for a particular case; for a naive example, machine learning in tribology can predict if wear is coming from a specific type of coated surface or an uncoated surface, which is a discrete-valued output [25].

Regression, on the other hand, provides a continuous numerical value. It is helpful to estimate any relationship between variables. Some of the common regression algorithms are linear regression and logistic regression, to name a few. One example of regression could be the investigation of the relationship between the coefficient of friction (COF) and surface roughness in sliding contact [26,27]. For instance, let us consider a study where different sample surfaces with varying roughness values are subjected to sliding against a counter material. The COF is measured under controlled conditions, and the corresponding surface roughness parameters (Ra, Rq, Rsk, and Rku) are quantified for each sample. A regression analysis could then be performed to determine the correlation between the COF and surface roughness, allowing for the formulation of a predictive model. Such a model could help estimate the friction behavior of similar materials with known surface roughness, aiding in the design and optimization of tribological systems.

For unsupervised learning, no labeled data is provided to the model; rather, an unlabeled data set is fed, and the capabilities of the model are utilized to find the hidden structures in the dataset. Under the unsupervised learning category, cluster analysis is helpful to group a similar set of objects in one group while grouping another set of similar objects in another group. The wear particle data obtained from the lubricated system could be analyzed using this technique. Clustering can form distinct clusters of data that share similarities [28]. Such clustering can provide insights into the wear mechanisms, lubricant effectiveness, or the presence of any irregular wear conditions.

2.2. Machine Learning Algorithms

There are multiple algorithms available to utilize the characterization of a dataset. Each such algorithm needs a unique approach to learning. The choice of one algorithm for favoring a particular benefit could cause a tradeoff of other benefits, including accuracy, complexity, and speed [24]. Therefore, the success of an algorithm depends on the specific nature of the problem, its complexity, cost, and often trial and error.

Currently, there are different platforms to utilize ML algorithms, such as Matlab, Python, C, and C++, to name a few. In Figure 1, different algorithms are presented that are compatible with the Matlab platform [15].

2.3. Explored Machine Learning Algorithms

Machine learning is a process that involves computer programs using software such as MATLAB and Python and needs to be structured following some algorithms. Scientists have explored many fundamental algorithms for lubrication, and a few of them have repeatedly appeared in tribological studies, such as linear regression, logistic regression, support vector machines, discriminant analysis, Naïve Bayes, decision trees, and artificial neural networks. The mathematical modeling of these seven algorithms with relevant tribological examples is summarized here.

2.3.1. Linear Regression

Regression analysis is a widely used technique to analyze multifactor data [29]. The basic idea behind linear regression is to get a line that best describes the relationship between the input and output variables. For example, if $x$ represents an experimental time and $y$ represents a wear volume, then as shown in Figure 2, the straight-line equation relating these two variables could be as follows: where $\beta_0$ is the intercept and $\beta_1$ is the slope.
relating these two variables could be as follows: where $\beta_0$ is the intercept and $\beta_1$ is the slope. 

$$h(x) = \beta_0 + \beta_1 \times x$$ (1)

Figure 2. Linear regression model.

Here, $h(x)$ is called the hypothesis function that is used to approximate the target variable (i.e., the wear volume here). By defining an objective function or cost function, computer scientists try to minimize the error between the obtained line and the actual data points. If $y$ is the actual output variable and there are $m$ total training samples, we can write,

$$cost(\beta_0, \beta_1) = \frac{1}{2} \sum_{i=1}^{m} (h(x^{(i)}) - y^{(i)})^2$$ (2)

To estimate the parameter values ($\beta_0$ and $\beta_1$) so that the line best fits the data, the cost function needs to be minimized. Researchers use a gradient descent algorithm to minimize the cost function and find the best set of parameters. The multiple linear regression model is used in the case of more than one input variable [30]. There are different variations of linear regression, such as Lasso and Ridge regression, to ensure the generalizability of the model [31].

2.3.2. Logistic Regression

Logistic Regression is a classification algorithm that estimates the probability that a certain data point belongs to a particular class. An investigation to classify the lubrication regime of hydrodynamic journal bearings was carried out using logistic regression [32]. Logistic regression follows a similar learning principle to linear regression. However, unlike linear regression, the output variable is binary, and thus, the input to the output mapping goes through a non-linear transformation called the sigmoid function to make a binary prediction. To illustrate, let us assume we want to fit the data with our hypothesis function $h(x)$, where

$$h(x) = \text{sigmoid}(\beta_0 + \beta_1 \times x)$$ (3)

The sigmoid function can transform any real-numbered value into a range between 0 and 1, which makes it a powerful choice of transformation in classification algorithms. For example, in Figure 3, if $x \to \infty$, then $\text{sigmoid}(x) \to 1$, and if $x \to -\infty$, then $\text{sigmoid}(x) \to 0$. The sigmoid function is defined as

$$\text{sigmoid}(z) = \frac{1}{1 + e^{-z}}$$ (4)
The sigmoid function can transform any real-numbered value into a range between 0 and 1, which makes it a powerful choice of transformation in classification algorithms. For example, in Figure 3, if $x \to \infty$, then $\text{sigmoid}(x) \to 1$, and if $x \to -\infty$, then $\text{sigmoid}(x) \to 0$. The sigmoid function is defined as

$$\text{sigmoid}(z) = \frac{1}{1 + e^{-z}}$$

Therefore, we obtain

$$h(x) = \frac{1}{1 + e^{-\beta_0 - \beta_1 x}}$$

(5)

Here, the cost function is derived using the maximum log-likelihood method.

$$\text{cost}(\beta_0, \beta_1) = \frac{1}{2} \sum_{i=1}^{m} y^{(i)} \log h\left(x^{(i)}\right) + \left(1 - y^{(i)}\right) \log \left(1 - h\left(x^{(i)}\right)\right)$$

(6)

In a similar way to linear regression, the gradient descent method is used to minimize the cost function and achieve the right set of parameters [33,34].

2.3.3. Support Vector Machines (SVM)

Support vector machines (SVM) are one type of ML algorithm that analyzes data for supervised learning. SVM provides a sorted data map with two groups, with some margins to the farthest possibility. This algorithm has been widely utilized in text recognition, image classification, and many other scientific applications [35,36]. This algorithm was observed as useful for the oil and gas exploration phase of the hydrocarbon industry [37]. It was reported that for hydrocarbon reservoir prediction, the SVM classification method was more suitable than expert judgment or filter approaches [37]. Especially for relatively complex problems, SVM was proven to be more suitable. A similar technique could be useful in tribology for predicting any cracks or irregularities on a material surface. Such a prediction would help to estimate the necessity of additional lubricants.

SVM was first invented in 1992, and over the last thirty years, it has become one of the most powerful ML algorithms [38]. The basic intuition behind the SVM is to build a separating hyperplane in an optimized way to create the maximum margin between two classes.

Let’s assume that in Figure 4, our separating hyperplane forms the following equation:

$$h(x) = g(\beta_0 + \beta_1 * x)$$

(7)
useful in tribology for predicting any cracks or irregularities on a material surface. Such a prediction would help to estimate the necessity of additional lubricants.

SVM was first invented in 1992, and over the last thirty years, it has become one of the most powerful ML algorithms [38]. The basic intuition behind the SVM is to build a separating hyperplane in an optimized way to create the maximum margin between two classes.

Let’s assume that in Figure 4, our separating hyperplane forms the following equation:

\[ h(\mathbf{x}) = \mathbf{g}(\beta_0 + \beta_1 \times \mathbf{x}) \]  

(7)

Figure 4. Support vector machine output with separating hyperplanes.

However, in most cases, we deal with more than one input variable. So, we can think of \( \mathbf{x} \) as an n-dimensional vector, where, 

\[ \mathbf{x} = [x_1, x_2, \ldots, x_n] \]

Here \( \mathbf{w} \) and \( b \) can be defined as model parameters, where, \( \mathbf{w} = [w_1, w_2, \ldots, w_n] \), and \( b \) is the intercept term. Therefore, the equation can be rewritten as:

\[ h_{\mathbf{w}, b} = \mathbf{g}(\mathbf{w}^T \mathbf{x} + b) \]  

(8)

Unlike the logistic regression, our target variable \( y \in \{-1, 1\} \). So, we define \( g(z) = 1 \), if \( z \geq 0 \), and \( g(z) = -1 \) otherwise. In SVM, the concept of maximum margin comes into play. The margin is defined as the closest distance from a data point to the hyperplane. It can be derived mathematically:

\[ \text{margin}, m = \frac{1}{||\mathbf{w}||} \]  

(9)

If there are \( m \) training examples, the constrained optimization problem of SVM can be written as:

\[ \text{minimize}_{\mathbf{w}, b} \frac{1}{2} ||\mathbf{w}||^2 \]  

(10)

subject to : \( y_i \left( \mathbf{w}^T \mathbf{x}_i + b \right) \geq 1, i = 1, \ldots, m \)  

(11)

There can be non-separable cases where positive and negative examples overlap. Therefore, the soft margin is used by introducing \( C \) and \( \gamma \) parameters. The parameter \( C \) determines the trade-off between the misclassification of training examples and the simplicity of the decision boundary, while the \( \gamma \) parameter determines the influence of individual training examples on the decision boundary. Different kernels can be used to learn the SVM in high-dimensional feature space, namely polynomial, gaussian, and radial basis function kernels [39].

2.3.4. Discriminant Analysis

Discriminant analysis is a statistical technique that can differentiate between two or more groups of objects while simultaneously taking several variables into account [40]. This technique was first developed by Fischer in 1936. Discriminant analysis has been used in anthropology, biology, criminology, and political science. Recently, some expeditions on tribology have been carried out with this analysis [41].
The main intuition behind a discriminant analysis (linear) is to find a hyperplane where the projection of the classes will have maximum separability, as shown in Figure 5. Let us assume that the projection of data point $X$ onto the line is $y$, where $X = [x_1, x_2, \ldots, x_m]$, and the projection vector $W = [w_1, w_2, \ldots, w_m]$. We can write

$$y = W^T X$$ (12)

Now, if the mean of the features ($X$) in classes 1 and 2 is $\mu_1$ and $\mu_2$, and the mean of the projections ($y$) in classes 1 and 2 are $\bar{\mu}_1$ and $\bar{\mu}_2$,

$$\mu_c = \frac{1}{N_c} \sum_{x \in c} X$$ (13)

$$\bar{\mu}_c = W^T \mu_c$$ (14)

From that, we define scatter as the sum of square differences between the projected samples and their class mean. Hence, the scatter for class $c$ can be written as:

$$S_{c}^2 = \sum_{x \in c} (y - \bar{\mu}_c)^2$$ (15)

According to Fisher’s linear discriminant, the objective function can be defined as:

$$J(W) = \frac{||\bar{\mu}_1 - \bar{\mu}_2||^2}{S_1^2 + S_2^2}$$ (16)

![Figure 5. The discriminant analysis.](image)

Maximizing $J$ ensures that we find a projection line where projections of the same class are in the closest proximity and their means are at the farthest distance.

2.3.5. Naïve Bayes

Naïve Bayes is a simple classification technique that utilizes Bayes’ rule along with a string assumption depicting that the attributes are conditionally independent given the class [42]. Naïve Bayes has been widely used due to its many desirable features. Sreenath et al. [43] implemented the Naïve Bayes algorithm to monitor the failure mode of a gearbox. According to Bayes theorem, we can write:

$$P(H|D) = \frac{P(H|D)P(h)}{P(D)}$$ (17)

where

- $P(h)$ is the probability that a hypothesis is true, irrespective of the data;
2.3.5. Naïve Bayes

Naive Bayes is a simple classification technique that utilizes Bayes’ rule along with a string assumption depicting that the attributes are conditionally independent given the posterior, which is also known as the Maximum Posteriori (MAP) decision rule. In Naive Bayes, we first calculate the prior probabilities for given classes. Then the conditional probability of the attributes given the class labels is computed. All these values are used to compute the posterior. We pick the class (the hypothesis) with the maximum posterior, which is also known as the Maximum Posteriori (MAP) decision rule.

2.3.6. Decision Trees

A Decision Tree is a non-parametric learning algorithm that works by splitting the dataset based on different conditions. This algorithm helps to model the consequences of possible decisions in the context of any outcome or event [44]. Decision tree classification was utilized in lubricant condition monitoring for generating the model for predicting wear conditions of equipment using UOA (used oil analysis), wear particle data, and failure data [45]. Furthermore, it has been used for classifying oil samples, and fault detection [43]. However, despite its flexibility, this model can be sensitive to missing values and outliers.

The Decision Tree works in a tree-like structure, where the topmost node is known as the root (Figure 6). The root is an attribute selected by measuring the information gain, and the feature with the maximum information gain is selected as the root node. The internal nodes are the rest of the features. The branch represents the decision rule, and the leaf nodes (nodes without any children) denote the decision.

![Decision Tree Diagram]

Figure 6. Decision tree.

2.3.7. Artificial Neural Network (ANN)

An artificial neural network (ANN) is capable of creating models using interconnected mathematical nodes [46,47]. As such, ANN can imitate some abilities of the human brain and probe the advantages of parallel processing, noise immunity, strong fault tolerance, and good memorization [48–50]. ANN can learn from samples and explore those samples to realize the relationships between the inputs and results. Therefore, ANN is very suitable for treating nonlinear and complex tribological problems that cannot be solved using traditional physical theories or regular mathematical approaches. The basic step of an ANN is presented in the below picture (Figure 7), where an artificial neural network scheme is shown for fuel consumption prediction based on various engine parameters such as torque, speed, intake air pressure, and EGR%. In a similar fashion, lubrication research was facilitated by ANN to predict lubricant performance or the need for maintenance by analyzing input parameters of the tribological system [51].

A neural network provides a layered representation of the input features in the latent space, with hidden layers and their corresponding activation functions. Activation functions are used to add non-linearity to the neural network.
2.3.7. Artificial Neural Network (ANN)
An artificial neural network (ANN) is capable of creating models using interconnections of neurons. Figure 7 shows the artificial neural network model scheme for fuel consumption prediction (adopted from Tomanik et al. [52]).

Let us assume a dataset with n examples has three input features $x_1, x_2, x_3$, and corresponding outputs $\hat{x}_1, \hat{x}_2, \hat{x}_3$. The mathematical notation, $a^{(l)}_i$ denotes the neuron in the $i$th unit of the $l$th layer. In Figure 8, the first layer is the input layer, which is defined as a vector $a^{(1)}$. All the internal layers other than the final layer are called hidden layers. The first (and only) hidden layer $a^{(2)}$ is mapped from the input layer by some weights. The final layer is obtained in the same way and then transformed into a regression or classification setting. It uses a backpropagation algorithm to learn from the error, which is an algorithm to find the gradient required to estimate the weights of the ANN. Backpropagation was first introduced in 1970 but brought to light by Rumelhart et al. in 1986 [53,54]. To introduce non-linearity, the nodes of the layers go through further transformation by an activation function. Depending on the specific problem at hand, the choice of activation function includes sigmoid, tanh, ReLU, and LeakyRelu [43].

![Figure 7. The artificial neural network model scheme for fuel consumption prediction (adopted from Tomanik et al. [52]).](image)

![Figure 8. A simple neural network. The input layer contains the input features and is denoted as $x_i$. The bias term “b” is added as an additional parameter to help the network learn a flexible decision boundary.](image)
As such, for layer 2, the following set of equations can be derived:

\[ a_1^{(2)} = g(\theta_{10}^{(1)} x_0 + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2 + \theta_{13}^{(1)} x_3) \tag{18} \]

\[ a_2^{(2)} = g(\theta_{20}^{(1)} x_0 + \theta_{21}^{(1)} x_1 + \theta_{22}^{(1)} x_2 + \theta_{23}^{(1)} x_3) \tag{19} \]

\[ h_\theta(x) = a_1^{(3)} = g(\theta_{10}^{(2)} a_0^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)}) \tag{20} \]

Here, \( \theta \) with different subscripts and superscripts denote the weights. For example, \( \theta_{ij}^{(k)} \) denotes the connection weight from the \( j \)th unit of the \( k \)th layer to the \( i \)th unit of \( (k + 1) \)th layer. \( x_0 \) and \( a_0 \) denotes the bias terms, and activation is applied through \( g(\cdot) \). This process is also called forward propagation (FP). If the dataset has \( m \) examples and there are \( K \) classes, the cost function can be written as:

\[ J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \sum_{k=1}^{K} y_k^{(i)} \log(h_\theta(x^{(i)}))_k + (1 - y_k^{(i)}) \log\left(1 - h_\theta(x^{(i)})_k\right) \tag{21} \]

Using backpropagation, the \( \theta \) values are updated with every iteration. The gradient descent algorithm ensures that we find the best set of \( \theta \) so as to minimize the cost function.

There are many variations of ANN based on the nature of the task. A few popular algorithms include the Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), and Generative Adversarial Network (GAN), etc. [55–57].

3. Application of Machine Learning Algorithms for the Lubricant Industry

The lubricant industry is heading towards automation through condition monitoring and useful life prediction. This shift is not only happening on the experimental side to choose a lubricant; rather, it starts with the extraction of the crude oil from the oil bore [58]. The geological data and drill bit condition monitoring were a primary focus for safe oil extraction. When the crude oil is purified and assigned as lubricants, monitoring is very important to ensure a safe operation for the machine with minimum breakdown. For that, the condition of the lubricant is critical. In addition, at the design phase, the prediction of COF and wear is important to avoid unwanted machine breakdown due to lubrication failure. If the algorithm is trained well, it will predict with higher accuracy, which would be beneficial to designing lubricants with superior qualities. In addition, there have been studies where experimental parameter prediction was a key concern to identify failure conditions. Also, ML algorithms have been utilized for film thickness prediction, which dictates lubricant performance significantly. Overall, the recent progress of ML algorithms for lubrication could be segmented into the following four categories:

1. Predicting experimental parameters;
2. Predicting film thickness;
3. Predicting COF and wear;
4. Lubricant condition monitoring.

Each of these categories is discussed in the following sub-sections with a state-of-the-art literature review in tabulated format.

3.1. Predicting Experimental Parameters

Experimental parameter prediction is an important aspect where ML could be utilized. In a few earlier expeditions, wear profiles were inspected and trained against their operating conditions in order to enable the algorithm to predict the experimental conditions for any similar wear profile. This prediction is important because it can tell if any experimental condition has gone wrong. If there is significant wear observed, then the prediction can illustrate the condition behind that wear so that the operator can improvise such extreme
conditions quickly to minimize that condition deliberately. Also, performing tests under critical conditions might be more convenient for computer simulation than real experiments.

In the recent past, Gong et al. [59] proposed an approach for simulating the tribochemistry, wear, and stress of a general fully formulated oil and found that tribofilm growth can be promoted by temperature and shear. Information from such simulations could be helpful in training ML models to predict experimental conditions. Also, molecular dynamic simulation can help predict the lubricant’s performance and evaluate its lubrication mechanisms. These techniques have been proven helpful to complement experimental observation to evaluate lubricant performance to some extent. Therefore, parameters from simulations are often used as descriptors to evaluate lubricant performance [60]. In Table 1, a few examples are illustrated where researchers have investigated the experimental conditions against some output results using popular algorithms such as discriminant analysis, Bayesian modeling, and the artificial neural network (ANN).
Table 1. Predicting experimental parameters using ML techniques.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multilayer neural network</td>
<td>Four particle descriptors were used-&lt;br&gt;1. Representative diameter (0.3 µm to 21 µm)&lt;br&gt;2. Elongation (1 to 4.7)&lt;br&gt;3. Roundness (1.1 to 6.5)&lt;br&gt;4. Reflectivity (0.05–0.92)</td>
<td>• Paraffin oil H60&lt;br&gt;• H60 with 0.1 wt.% steric acid&lt;br&gt;• H60 with 1.0 wt.% TCP&lt;br&gt;• H60 with 0.5 wt.% DBDS&lt;br&gt;• H60 with 1.0 wt.% ZndTP</td>
<td>Data was taken from pin-on-disk experiments under 58.8 N and 176 N loads. The wear particle features observed from the microscopic images were analyzed using ML to predict the unknown experimental conditions.</td>
<td>[61]</td>
</tr>
<tr>
<td>Fuzzy kohonen neural network, multi-layer perception with backpropagation</td>
<td>Six types of wear particles were analyzed: rubbing, cutting, spherical, laminar, fatigue chunk, and severe sliding. Wear particles were characterized based on: Area, length, roundness, fiber dimension, fractal dimension, Height aspect ratio (HAR), Average roughness (R_a), RMS roughness (R_q), and spectral moment analysis (γ^2).</td>
<td>N/A</td>
<td>Six common types of metallic wear were studied using three-dimensional images obtained from laser scanning confocal microscopy. This study showed a way that a neural network could be effectively implemented to classify wear particles.</td>
<td>[62]</td>
</tr>
<tr>
<td>Discriminant analysis</td>
<td>Two types of markers were used-&lt;br&gt;1. Macro (additive elements Ca, Zn, P)&lt;br&gt;2. Trace (wear metal elements of Ag, Al, Ba, Cd, Cr, Cu, Fe, Mg, Mo, Na, Pb, and Sn)</td>
<td>• ZIC A (SK) 5w30&lt;br&gt;• Kixx G1 (GS Caltex) 5w30&lt;br&gt;• HD premium gasoline (Soil) 5w20&lt;br&gt;• ZIC XQ RV/SUV (SK) 5w40&lt;br&gt;• Kixx RV (GS Caltex) 10w40&lt;br&gt;• HD premium diesel (SK) 5w30</td>
<td>Six common automobile gasoline and diesel engine oils were collected from seventy-six sources, analyzed using inductively coupled plasma-optical emission spectrometry (ICP-OES), and then statistically compared. Brand of oil, engine type, and source vehicle type were predicted using ML with 92.1%, 82.9%, and 92.1% accuracy, respectively. This oil analysis would be beneficial for forensic analysis.</td>
<td>[41]</td>
</tr>
</tbody>
</table>
### Table 1. Cont.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian modeling and transfer learning approach</td>
<td>Descriptors were selected based on the physical models for estimating dissipation in 2D materials- 1. Lattice constant 2. c/a ratio 3. X-M-X angle 4. Interlayer spacing 5. M-X bond length 6. X-X distance 7. Hexagonal width 8. Covalent radii of M/X atoms 9. In-plane stiffness 10. Cohesive energy per unit cell 11. Thermal conductivity 12. Exciton binding energy 13. Average atomic mass 14. Bandgap energy</td>
<td>• Gr  • BN  • h-Gr  • f-Gr  • MoS2  • MoSe2  • MoTe2  • NbS2  • NbSe2  • NbTe2  • WS2  • WSe2  • WTe2  • TiS2  • TiSe2</td>
<td>Five experimental data sets (from the literature) and ten molecular dynamic simulation data sets of the maximum energy barrier (MEB) of the potential surface energy were obtained. MEB is correlated to intrinsic friction (friction for a crystalline material with no defects) obtained from ten different 2D materials. Less than an 8% difference was observed in MEB values as predicted via the ML model and the PES profiles obtained from the MD simulation.</td>
<td>[60]</td>
</tr>
<tr>
<td>Back propagation neural network (BPNN), Generic algorithm (GA)</td>
<td>Content of additives for 25 combinations</td>
<td>• Base oil: trimethylolpropane trioleate  • Additives: AW 34, AW 38, and AW 12 (These were phosphonium-based additives, added up to 1% compared to the base lubricants).</td>
<td>A four-ball test was carried out to compare the results from the experiment and modeling. The content of the three functional additives varied between 0 and 0.7 wt.%, with a total additive wt.% of 1% in the bio-based base lubricant. The content of those additives was used as the input variable for the algorithm. The model was developed by combining the BPNN with GA and provided theoretical guidance for compounding optimization of lubricant additives. The predicted output matched above 90% of the experimental values.</td>
<td>[63]</td>
</tr>
</tbody>
</table>
3.2. Predicting Film Thickness

ML has been implemented to predict the film thickness of lubricants. This is important for gearboxes, where oil lubrication restricts the real contact between gear teeth and reduces friction, heat buildup, vibration, and corrosion. Ali et al. [64] prescribed a lambda ratio (λ) to predict the fatigue life of the gears, where λ is the ratio between film thickness and RMS surface roughness. Basically, the factors that affect the λ ratio are load, temperature, surface roughness, and gear speed [65]. The film thickness could be derived using the following equation [64]:

\[ h = \frac{k(\eta_0u)^{0.7}R^{0.43}}{\alpha^{0.13}m} \]  

(22)

Here, \( \eta_0 \) is the dynamic viscosity of the lubricant in Pa·s, \( k = 1.6 \alpha^{0.6} E^{0.03} \), \( \alpha \) = pressure-viscosity co-efficient in \( \text{mm}^2/\text{N} \), \( R \) = equivalent radius in m, and \( \omega \) = load applied normal to the line of contact in N/m. It was reported that gears usually operate in the elastohydrodynamic and boundary regimes [64]. In the boundary regime, the film thickness becomes lower, and contact between the tribo-pairs is likely to happen [3,66,67]. Therefore, knowledge about film thickness is very important to design and monitor oil performance. As per their studies, the network simulation conducted by MATLAB Simulink attained 100% success in predicting and classifying at high speed [64]. Table 2 represents individual cases where researchers have investigated the film thickness and highlighted the lubrication regime per the Strubeck curve.

### Table 2. Predicting film thickness and lubrication regime using ML techniques.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network and logistic regression</td>
<td>Six variables in this study were- 1. Specimen speed (rpm) 2. Load (kN) 3. Temperature of oil 4. Friction coefficient 5. Temperature of bearing 6. Contact potential (mV)</td>
<td>Hydrodynamic journal-bearing data was used</td>
<td>A total of 888 high-speed torque measurements were taken in different experiments. Those experiments were carried out on a journal-bearing test rig. Up to 99.25% accuracy was observed for predicting lubricant regime class using an artificial neural network and logistic regression.</td>
<td>[32]</td>
</tr>
<tr>
<td>Feed-forward algorithm, Recurrent Elman neural network algorithm, Levenberg-Marquardt back-propagation algorithm</td>
<td>The following information was given as inputs from previous works- 1. RMS roughness 2. Temperature 3. lambda ratio (λ)</td>
<td>Mobil MOBILGEAR 636</td>
<td>MATLAB Simulink was used to run the algorithms to monitor the oil film thickness and detect the lubrication regime for the gearbox. Levenberg-Marquardt’s back-propagation algorithm was used to reduce the error.</td>
<td>[64]</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Loading conditions were utilized, such as- 1. Vibration, 2. Temperature 3. Torque</td>
<td>Any</td>
<td>The condition of wear and film thickness were predicted using ANN for a gearbox. MATLAB was used to create the simulation model. Hammock-Dowson’s equation was utilized in the model’s development.</td>
<td>[68]</td>
</tr>
</tbody>
</table>

3.3. Predicting COF and Wear

The prediction of COF and wear rate is important from a research and development point of view. A lubricant may provide good performance at one operating condition but may not be effective at other conditions. Predicting their best fit-in situation is important to formulate novel lubricants with superior performances. ML has started to ease this task by helping to predict the COF and wear values for extreme conditions, provided that the algorithm is trained with the data that was obtained from tests previously. In Table 3, several examples have been presented where ML algorithms have been successfully utilized to predict the coefficient of friction and wear.
Table 3. Predicting coefficient of friction and wear for arbitrary experimental conditions using ML techniques.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple, multivariable, linear, polynomial, and support vector machine regression</td>
<td>The following parameters were calculated under multiple experimental conditions and used as inputs for training the models- 1. Bearing pressure distribution 2. Load carrying capacity 3. Friction</td>
<td>• Mono-grade SAE 30 • Multi-grade SAE 10W-40 • Bio lubricant AWS 100</td>
<td>The operating condition was predicted for a thrust bearing with three lubricants. The ASW 100 biolubricant exhibited the lowest friction and better load-carrying capabilities among the three lubricants. The friction prediction using different algorithms varied between 70 and 93%. For example, simple linear regression (79%), simple 2nd order polynomial regression (93%), multivariable linear regression (83%), multivariable 2nd order polynomial regression (93%), and quadratic SVM regression (90%) successfully predict the friction values.</td>
<td>[69]</td>
</tr>
<tr>
<td>Bayesian regularization neural network</td>
<td>Molecular connectivity indices were used as the descriptors of 2D BRNN–QSTR models. The descriptors were categorized as: quantum descriptors, 3D descriptors, 2D topological descriptors. Some of the important descriptors were- 1. e IAC_ Mean [70] 2. Balaban indices [71] 3. Wiener index [72]</td>
<td>Thirty-six compounds were tested as additives. Such as- benzoxazole, benzimidazole, benzothiazole, dihydrothiazole piperazine, and piperidine were used as lubricant additives, and their tribological data were examined.</td>
<td>The authors concluded that elements S and P in the lubricant significantly improved tribological properties. Also, minimal carbon chain length is needed in the case of additives. However, extending the carbon chain length might not always be beneficial; rather, there could be a threshold number of carbons in the chain that would facilitate the lowest wear.</td>
<td>[73]</td>
</tr>
<tr>
<td>Linear regression</td>
<td>• Low orbital energy • Dipole moment</td>
<td>Thirty-six lubricant data were used for wear estimation. Among them, 29 groups were used for training purposes whereas 7 groups were used to perform the prediction.</td>
<td>For COF and wear estimation, a standard four-ball tester was used at 1450 rpm over 30 min. A linear regression algorithm was used to predict the wear and COF of seven test lubricants. The database was established in the MySQL database.</td>
<td>[74]</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Neural network predictor consisted of: one input neuron, ten hidden layer of neurons, and three output neurons with non-linear activation function</td>
<td>Mobil 0W-40</td>
<td>The load-carrying capacity of the journal bearing was investigated with a steel shaft with varying surface textures. The developed neural network was able to learn quickly and predict the load-carrying capacity.</td>
<td>[75]</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed forward artificial neural</td>
<td>Four dimensional inputs were-</td>
<td>PAO-6</td>
<td>The experiments were carried out on the tribological testing equipment named a mini-traction machine (MTM) using the ball-on-disk method. The friction coefficient was predicted using the feed-forward artificial neural network, which was simulated in MATLAB software.</td>
<td></td>
</tr>
<tr>
<td>network</td>
<td>average velocity: 1 to 4 m/s, slide roll ratio (SRR): 2–200%, bath temperature: 40–100 °C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>contact load: 5–28 N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Seven inputs from the lubricant properties were:</td>
<td>Jatropha, Pongam, Neem, Mahua, and mineral oil (Servocut 945)</td>
<td>A tool-chip tribometer was used to simulate a similar aspect of a metal-cutting operation with a pool of oil samples. The prediction was made on the surface roughness and COF using artificial neural network for bio-lubricants.</td>
<td>[77]</td>
</tr>
<tr>
<td></td>
<td>• Kinematic viscosity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Viscosity index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Saponification value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Acid value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Flash point</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pour point</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Iodine value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inputs taken from the machining conditions were:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Operational: sliding, cutting, post sliding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Time: 5–90 s (increment of 5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• RPM: 100, 150, and 200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Loading conditions were utilized, such as-</td>
<td>-</td>
<td>The condition of wear and film thickness were predicted using ANN for a gearbox. MATLAB was used to create the simulation model. Hammock–Dowson’s equation was utilized in the model’s development.</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>Vibration, Temperature</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>torque</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed forward back propagation algorithm, Artificial neural network</td>
<td>Load, velocity and additive oil rate</td>
<td>Lubricant (Mineral oil, SAE 20W50) with polytetrafluoroethylene (PTFE) based additives at 1, 3, 5, 10, and 15% concentration</td>
<td>Two loads were implemented (153 and 253 N) for rotational speeds of 30, 60, 120, 180, 240, 300, 360, 420, 480, 540, 600, 660, 720, 780, 840, 900, 960, 1020, 1080, 1140, and 1200 rpm. The average COF from three tests was used to train the algorithm, and the artificial neural network reproduced the COF outputs with 98% accuracy.</td>
<td>[78]</td>
</tr>
<tr>
<td>Logistic curve</td>
<td>$\lambda$ ratio, $X = \left(1 + \lambda^k\right)^{-a}$ where $k \approx \frac{3}{2}$ and $a \approx \frac{4}{3}$</td>
<td>SAE 10 W-40, XHVI 5.2, Few other oil data from the literature [79,80]</td>
<td>A new equation: $X = \left(1 + \lambda^k\right)^{-a}$ is proposed that enables the simple calculation of the proportion of mixed/boundary friction in a contact as a function of $\lambda$ ratio. When the $X$ is plotted against $\lambda$, it represents a logistic or sigmoid curve. For a varied $\lambda$, this curve could be useful to predict friction under boundary and mixed lubrication regimes.</td>
<td>[81]</td>
</tr>
</tbody>
</table>
3.4. Lubricant Condition Monitoring

Lubricant plays an important role in reducing friction, wear, and associated costs in machinery. Therefore, it is crucial to monitor the lubricant’s conditions regularly. In small-scale applications, periodic checks by maintenance technicians can ensure the safe usage of a lubricant. However, on sites such as wind turbines or in space, regular checking might be difficult. Also, the opportunity cost of inspection could be higher if the lubricant condition is monitored periodically by the shutdown of the running machines. In order to reduce these costs, lubrication condition monitoring (LCM) plays a very vital role. Manghai et al. [82] reviewed brake fault monitoring and highlighted the significance of ML in that process. For lubricant fault monitoring, a few classes of parameters hold significance. In Table 4, such parameters have been tabulated for reference.

Table 4. Common parameters used in lubrication monitoring (adapted from [21]).

<table>
<thead>
<tr>
<th>Class</th>
<th>Parameter</th>
<th>Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td>Viscosity at 40 °C and 100 °C, Thermal stability, Temperature, Density, Vibration</td>
<td>[83–85]</td>
</tr>
<tr>
<td>Chemical</td>
<td>Total Base Number (TBN), Total Acid Number (TAN), Flashpoint</td>
<td>[83–85]</td>
</tr>
<tr>
<td>Additives</td>
<td>Barium (Ba), Boron (B), Calcium (Ca), Magnesium (Mg), Molybdenum (Mo), Phosphorus (P), Sodium (Na), Silicon (Si), Zinc (Zn)</td>
<td>[86,87]</td>
</tr>
<tr>
<td>Contamination</td>
<td>Boron (B), Coolant, Soot/carbon, Potassium (K), Silicon (Si), Sodium (Na), Vanadium (V), Water (H₂O)</td>
<td>[88–90]</td>
</tr>
<tr>
<td>Elemental</td>
<td>Aluminum (Al), Chromium (Cr), Copper (Cu), Iron (Fe), Lead (Pb), Nickel (Ni), Silver (Ag), Titanium (Ti), Tin (Sn), Vanadium (V)</td>
<td>[91–94]</td>
</tr>
</tbody>
</table>

Continuous lubrication condition monitoring is especially important for large-scale machine tribo-pairs, wind turbine gearsets, and large plants where regular maintenance might be difficult or may interrupt operating hours. Such interruption could lead to a significant reduction in production, and therefore, automatic condition monitoring could be a great tool to implement. In recent years, companies like SKF have offered remote diagnostic services for machinery that help monitor machine health regularly [95]. Schaeffler Group also offers condition monitoring services for customers, where vibration and temperature are used [96]. A vibration and temperature monitoring system involves a set of tools to measure one or more parameters in order to identify changes in the behavior of machinery. The key advantage of such monitoring is that it helps schedule maintenance activities based on predictive analysis. In addition to temperature and vibration, other parameters such as load, material hardness, and lubricant properties also play an important role and can help predict the service life of lubricants. In some recent research articles in this area, ML algorithms such as ANN, decision trees, Naïve Bayes, and SVM have been utilized. In Table 5, lubricant condition monitoring and surface condition prediction techniques are summarized based on the available literature.
### Table 5. Predicting lubricant life or other critical information using ML techniques.

<table>
<thead>
<tr>
<th>Applied Algorithm</th>
<th>Input Parameters/Descriptors</th>
<th>Involved Lubricants</th>
<th>Remarks</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial neural network</td>
<td>-</td>
<td>Solid lubricants</td>
<td>Solid lubricant performance was studied using the artificial neural network technique.</td>
<td>[46]</td>
</tr>
<tr>
<td>Artificial neural networks (ANNs)</td>
<td>Vibration signals extracted from rotating parts were utilized to investigate the effect of oil quality on gearboxes.</td>
<td>Energo® GR-XP, SynthetiC (SHC), SPARTAN EP gear oils, Gravis SP series, Omala Oil 320 EP gear oil</td>
<td>ANN was employed to predict the vibration parameters of the experimental test rig. The wavelet features were extracted from the vibration data of the gearbox, and the gear fault was classified. The results suggest that the proposed RBNN is more predictive than the MNN. On the other hand, the Omala oil 320 was predicted to be reliable for such gearbox applications compared to other oils.</td>
<td>[51]</td>
</tr>
<tr>
<td>Short Time Fourier Transform (STFT), Fuzzy-based classification algorithm</td>
<td>Hardness, Normal force, Lubricant, Lubrication interval</td>
<td></td>
<td>STFT was used to extract relevant frequencies. Relationship prediction was possible between the physical failure mode and the calculated characteristics.</td>
<td>[97]</td>
</tr>
<tr>
<td>Artificial neural networks (ANNs), adaptive neural-based fuzzy inference systems (ANFIS), and fuzzy clustering method (FCM)</td>
<td>Sliding distance (m), Type, Load (N), Cooling rate (°C/s), Volume loss (cm³)</td>
<td>Dry condition</td>
<td>Three cooling rates (0.5, 2, and 3 °C/s) were applied on six specimens made from pre-alloyed Astaloy 85 Mo and Distaloy AB powders through the powder metallurgy (PM) method.</td>
<td>[98]</td>
</tr>
<tr>
<td>Decision tree, Naïve Bayes</td>
<td>Statistical features were given as the inputs</td>
<td></td>
<td>The vibration signal was monitored from the actual gearbox of the automobile, with simulated fault conditions within the gear and bearings. The fault detection was performed using two algorithms, and classification accuracy was observed at more than 80% for all cases.</td>
<td>[43]</td>
</tr>
<tr>
<td>Random Forest (RF)</td>
<td>Gearbox input load moment and force measurement, torque arm displacement and angular misalignment, wind direction and speed, generator mount displacement, generator rotational speed, and blade pitch angle</td>
<td>None</td>
<td>The authors highlight the need for machine learning in wind turbine drivetrain load monitoring. A RF model is employed to determine sensor positions that significantly influence the accuracy of virtual load sensing in wind turbine transmission.</td>
<td>[99]</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Lubricant properties were used as inputs. Data were observed and feed based after a certain life interval</td>
<td>-</td>
<td>The remaining utilizable life of the lubricants (RULL) was predicted using ANN.</td>
<td>[100]</td>
</tr>
<tr>
<td>Applied Algorithm</td>
<td>Input Parameters/Descriptors</td>
<td>Involved Lubricants</td>
<td>Remarks</td>
<td>Ref.</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------</td>
<td>--------------------</td>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>Artificial Neural Network (ANN)</td>
<td>Bearing shaft temperature, axial load, rotational speed, lubricant volume flow rate Bearing and resistor voltage, temperature and pressure of storage container, and temperature of chamber input</td>
<td>Not disclosed</td>
<td>The authors used the Long Short-Term Memory (LSTM, a variation of ANN) autoencoders to reconstruct time series data without the white etching cracks (WEC) in order to identify the anomalies in the new unseen data based on the reconstruction error.</td>
<td>[101]</td>
</tr>
<tr>
<td>ML technique was not used</td>
<td>Oil samples were analyzed by dint of: Viscosity Viscosity index Element analysis Total acid number (TAN)</td>
<td>Poly alfa olefin (PAO) Mineral oil</td>
<td>Wind turbine gearbox oil was monitored, and the degradation was studied.</td>
<td>[102]</td>
</tr>
<tr>
<td>Backpropagation feed-forward neural network algorithm</td>
<td>Input parameters were- Laser power Scanning velocity and energy density</td>
<td>The lubricant was not used, rather surface texturing was studied where the output parameters were average roughness (Sa), RMS roughness (Sq), Skewness (Ssk), Kurtosis (Sku)</td>
<td>Surface roughness parameters were predicted using a neural network. Laser surface processing parameters were used to evaluate the obtained surface properties.</td>
<td>[103]</td>
</tr>
<tr>
<td>Support Vector Machine (SVM), Real-time audio signal analysis,</td>
<td>The acoustic emission pattern was filtered and used for training the algorithm</td>
<td>CAW33 type 22,205 bearings were used for the experimental tests.</td>
<td>Bearing failure prediction plays an important role in safety. The SVM classifier provided at least 92% mean accuracy.</td>
<td>[104]</td>
</tr>
<tr>
<td>Back propagation neural network, Modular neural network, Radial basis neural network</td>
<td>A wide range of inputs was used for each of the three algorithms</td>
<td>Used and unused engine oils</td>
<td>Two types of vehicle engines were used for the analysis of oil quality and performance. The input layer consisted of 1 neuron, whereas the hidden layer consisted of 10 neurons and the output layer consisted of 4 neurons. This research provided helpful information for industrial applications for vehicle oil analysis and fault detection.</td>
<td>[105]</td>
</tr>
<tr>
<td>Artificial neural network</td>
<td>Leakage oil quantity and some other parameters were involved</td>
<td>-</td>
<td>The effect of slippers (which enhance the efficiency of axial piston pumps and motors) on lubrication was studied under different surface roughnesses and conditions.</td>
<td>[106]</td>
</tr>
</tbody>
</table>
4. Future of ML in the Lubrication Industry

Lubricant design and development over the last century were somewhat dependent on trial and error [107]. Now that ML has been introduced in the lubricant industry, future prospects will undoubtedly be influenced by this technique. Lubricant performance depends on the operating conditions, lubricant properties, and material-pair properties. Therefore, there is scope to incorporate a variety of parameters under machine learning other than those that have already been considered. In recent studies, it was observed that friction and wear can be correlated not only to sliding speed, sliding distance, and normal load but also to properties such as hardness, yield strength, tensile strength, and ductility [108]. It is very likely that ML descriptors will incorporate these variables to project COF and wear and may be able to further predict the associated real-time cost of energy due to such interaction.

Moreover, it is known that to reduce friction and wear, lubricant is introduced between tribo pairs, where the shear strength of the lubricant needs to be less than that of the mating surfaces. Otherwise, there will be abrasion due to the lubricants. Therefore, when using additives or solid lubricants, the shear modulus could be an important parameter. Similarly, the hardness of both mating surfaces is critical to determining the adhesive or plowing wear. ML can be utilized to identify correlations between different properties of lubricants, such as film thickness, viscosity, and COF, which can then be used to optimize lubrication performance for specific applications. As such, the use of ML in tribology has initiated a new branch of studies called “Triboinformatics” [109]. Triboinformatics aims to enhance understanding, prediction, and optimization of tribological processes through the utilization of innovative data-driven approaches. It involves the development of models, algorithms, and databases to capture, store, analyze, and visualize tribological data for various applications, such as lubricant design, material selection, and predictive maintenance. For instance, ML can be used to predict the lubrication properties of a specific set of lubricants under different operating conditions and can help determine the most effective lubricant to use for a particular application. Recent trends in ML for the lubrication industry include the development of efficient models that can forecast the performance of lubricants with a high degree of accuracy [110]. These models can be trained using large datasets, which allows for more accurate predictions of the performance of lubricants in real-world applications. Overall, by using ML algorithms, it is possible to gain insights into lubricant properties that could otherwise be difficult to predict and analyze [74,111,112].

In addition to lubricant design, ML can play a significant role in the predictive maintenance of machines and equipment that require lubrication. The analysis of machine data is one of the fastest-growing areas in the industrial Internet of Things (IoT) and data analytics [113]. By analyzing real-time sensor data, machine-learning algorithms can predict when a machine requires lubrication or maintenance. Another exciting prospect of ML in the lubrication industry is its potential to reduce environmental impact. With the growing concern about climate change, companies are looking for ways to reduce their carbon footprint. ML can be used to optimize lubricant use, minimize waste, and reduce the environmental impact of manufacturing processes. By identifying patterns in data and analyzing performance, ML algorithms can optimize lubrication systems to minimize waste and improve efficiency. Furthermore, with the rise of Industry 4.0, the lubrication industry is rapidly moving towards a more connected, digital future [114]. Companies can leverage the vast amounts of data generated by sensors and equipment to optimize performance, reduce costs, and improve quality using ML techniques. As a result, valuable insights into the operations are possible, which not only predict failures but also make data-driven decisions to improve the bottom line.

In summary, ML is transforming the lubrication industry by providing a more accurate and efficient way of designing lubricants, optimizing lubrication systems, predicting maintenance needs, reducing waste, and improving environmental sustainability [115]. As the technology continues to evolve, ML will play a critical role in the future of lubrication, helping companies to remain competitive and reduce their carbon footprint.
5. Conclusions

ML is a field with versatile applications. With supervised and unsupervised umbrella terms, ML holds a wide variety of algorithms, each with its own unique advantages. Therefore, understanding these algorithms is important for tribologists. Along with the discussions on ML algorithms, this review offered the following conclusions:

- ML techniques are being increasingly utilized in the lubrication research and lubrication industry to enhance lubricant design and optimization and predict maintenance needs;
- Various ML algorithms, such as artificial neural networks, support vector machines, and decision trees, have been successfully applied to predict lubricant properties, such as viscosity, COF, and wear, under different operating conditions;
- ML can assist in identifying correlations between lubricant properties and performance, enabling the optimization of lubrication solutions for specific applications;
- From the literature studies, it was observed that SVM, linear regression, and Bayesian regression have been utilized several times, whereas the number of studies involving artificial neural networks is significant. The ANN algorithm has been improvised for all four aspects discussed with significant accuracy;
- The future of ML in the lubrication industry holds great promise, including advancements in lubricant design, predictive maintenance, environmental sustainability, and optimization of manufacturing processes;
- Lubricant condition monitoring is being improved through the use of ML, allowing for real-time analysis of lubricant parameters and early detection of potential issues in lubrication systems;
- By leveraging ML, companies can make data-driven decisions, reduce costs, improve efficiency, and minimize the environmental impact of lubrication processes;
- For experimental condition prediction or lubricant film estimation, ANN showed significant advantages. Moreover, for COF, wear prediction, and lubricant condition monitoring, researchers relied mostly on ANN. It is because of the built-in feature of neural networks, where the input and output layers are separated by various intermediate layers as per the designers’ requirements. Therefore, the future of the lubricant industry’s prediction ability will depend largely on how efficiently tribologists can develop efficient neural networks that can take tribo-informatics to the next level.

In conclusion, this research provided a thorough examination of seven machine learning algorithms applied in the field of tribology. The selection of tribological examples in this review was meticulously based on the specific algorithms employed, ensuring comprehensive coverage of the four discussed categories. It is important to acknowledge that the boundaries set for this review may have excluded certain articles that explore alternative algorithms and other intriguing topics within the broader realm of machine learning in lubrication. Therefore, considering the rapid advancement of this field, there is immense potential for future review articles to delve into those unexplored aspects and provide valuable insights for aspiring machine learning enthusiasts. Overall, this review will be helpful for researchers to understand the machine learning perspective better from a tribological point of view. The incorporation of ML in tribological studies will reduce machine breakdown and operating costs, help reduce carbon emissions, and unlock many possibilities. This will, in turn, help save energy and build a sustainable future.

Author Contributions: Conceptualization, M.H.R. and P.L.M.; methodology, M.H.R. and S.S.; formal analysis, M.H.R. and S.S.; investigation, M.H.R.; resources, P.L.M.; writing—original draft preparation, M.H.R.; writing—review and editing, M.H.R. and P.L.M.; supervision, P.L.M.; project administration, P.L.M. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge the support and facilities from the Department of Mechanical Engineering at the University of Nevada at Reno.

Data Availability Statement: Data will be available based on request.
Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Spikes, H. Friction modifier additives. Tribol. Lett. 2015, 60, 5. [CrossRef]
15. Rosenkranz, A.; Marian, M.; Profito, F.J.; Aragon, N.; Shah, R. The Use of Artificial Intelligence in Tribology—A Perspective. Lubricants 2021, 9, 2. [CrossRef]


78. Durak, E.; Salman, O.; Kurbanoglu, C. Analysis of effects of oil additive into friction coefficient variations on journal bearing using artificial neural network. *Ind. Lubr. Tribol.* **2008**, *60*, 309–316. [CrossRef]

79. Dawczyk, J.; Morgan, N.; Russo, J.; Spikes, H. Film thickness and friction of ZDDP tribofilms. *Tribol. Lett.* **2019**, *67*, 34. [CrossRef]


81. Taylor, R.I.; Sherrington, I. A simplified approach to the prediction of mixed and boundary friction. *Tribol. Int.* **2022**, *175*, 107836. [CrossRef]


102. Coronado, D.; Wenske, J. Monitoring the oil of wind-turbine gearboxes: Main degradation indicators and detection methods. Machines 2018, 6, 25. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.