

Dorothy Rajendran^{1*}, Thankappan Sasilatha¹, Susai Rajendran^{2,3},
Senthil Kumaran Selvaraj⁴, Caslav Lachnjevac⁵, Sivakumar
Santhana Prabha⁶, Rajendran Joseph Rathish⁶

¹AMET University, Department of EEE, Kanathur, East Coast Road, Chennai, India, ²St. Antony's College of Arts and Sciences for Women Thamarapady, Corrosion Research Centre, Department of Chemistry, Dindigul - 624005, Tamil Nadu, India, ³Pondicherry University, Puducherry, India, ⁴Vellore Institute of Technology (VIT), Department of Manufacturing Engineering, School of Mechanical Engineering-SMEC, Tamil Nadu, India, ⁵University of Belgrade, Faculty of Agriculture, Belgrade, Serbia, ⁶PSNA College of Engineering and Technology, , Dindigul, India.

Scientific paper

ISSN 0351-9465, E-ISSN 2466-2585

<https://doi.org/10.5937/zasmat2203280R>



Zastita Materijala 63 (3)
280 - 290 (2022)

Application of machine learning in corrosion inhibition study

ABSTRACT

Artificial intelligence is a branch of science concerned with teaching machines to think and act like humans. Machine learning is concerned with enabling computers to perform tasks without the need for explicit programming. Machine Learning enables computers to learn without the need for explicit programming. Machine Learning is a broad field that encompasses a wide range of machine learning operations such as clustering, classification, and the development of predictive models. Machine Learning (ML) and Deep Learning (DL) research is now finding a home in both industry and academia. Machine Learning technologies are increasingly being used in medical imaging. To detect tumours and other malignant growths in the human body. Deep Learning is making significant contributions to the advancement of industrial robotics. Machine learning algorithms are used in the self-driving car industry to guide the vehicle to its destination. Deep Learning and Machine Learning are also used in corrosion science and engineering. They are used to choose the inhibitor molecules from a large pool of available molecules.

Keywords: Artificial Intelligence, Machine learning, Deep learning, Neural Networks, Algorithms, The Input Layer, The Hidden Layer and The Output Layer.

1. INTRODUCTION

Artificial intelligence is a branch of science concerned with teaching machines to think and act like humans. Machine learning is concerned with enabling computers to perform tasks without the need for explicit programming. Deep learning is a subset of machine learning that uses artificial neural networks as its foundation. Machine Learning enables computers to learn without the need for explicit programming. Machine Learning is a broad field that encompasses a wide range of machine learning operations such as clustering, classification, and the development of predictive models [1-20]. Machine learning is also being used in corrosion science and engineering [21-31].

Deep Learning and Machine Learning

Deep Learning and Machine Learning are the two most popular technologies in the modern world. These terms are commonly used interchangeably. Deep learning is a subset of machine learning. Future trends and applications are discussed.

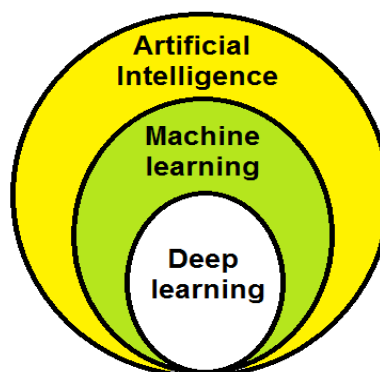


Figure 1. Relation among Artificial Intelligence, Machine learning and Deep learning

Slika 1. Odnos između veštačke inteligencije, mašinskog učenja i dubokog učenja

*Corresponding author: Dorothy Rajendran

E-mai: contactrdorothy@gmail.com

Paper received: 15. 02. 2022.

Paper corrected: 10. 03. 2022.

Paper accepted: 14.03. 2022.

Paper is available on the website: www.idk.org.rs/journal

Artificial intelligence, machine learning, and deep learning are all intertwined concepts (Figure 1).

Artificial Intelligence

It is a branch of Science devoted to making machines think and act like humans.

Machine learning

Machine learning focuses on enabling computers to perform tasks without explicit programming

Deep learning

Deep learning is a subset of machine learning based on artificial neural networks.

The various features of Deep learning and Machine learning are compared in Table 1.

Table 1. Comparison of various features of Deep learning and Machine learning

Tabela 1. Poređenje različitih karakteristika dubokog učenja i mašinskog učenja

Factors	Deep learning	Machine learning
Data Requirement	Requires large data	Requires lesser data
Accuracy	Provides high Accuracy	Provides lesser Accuracy
Training time	Takes longer time to learn	Takes less time to learn
Hardware Dependency	Requires CPU to train properly	Trains on CPU
Hyperparameter tuning	Can be tuned in various different ways	Limited tuning capabilities

Machine Learning

This section discusses various aspects of Machine Learning.

The scientific study of statistical models and algorithms that computer systems use to perform a task without explicit instructions is known as machine learning. Machine Learning is a broad field that encompasses a wide range of machine learning operations such as clustering, classification, and the development of predictive models.

Machine Learning, in essence, enables computers to learn without the need for explicit programming.

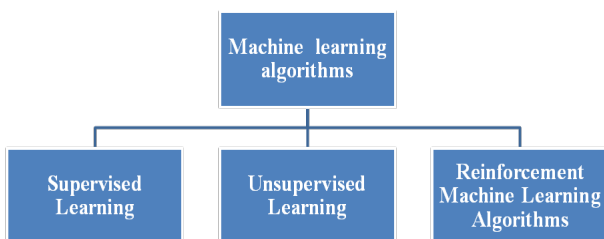


Figure 2. Three main types of machine learning algorithms

Slika 2. Tri glavna tipa algoritama mašinskog učenja

In most programming scenarios, we must provide instructions to the computer in order for it to produce output. Nonetheless, using machine learning algorithms, we can train our computer to

provide us with output without the need for us to provide instructions. A machine learning algorithm can accomplish this with the assistance of data. A machine learning algorithm is trained to provide output to users using the data fed into the system. Machine learning algorithms are classified into three types [2]. They are the algorithms for Supervised Learning, Unsupervised Learning, and Reinforcement Machine Learning.

Supervised Learning

The input data in a supervised learning algorithm is labelled in order for the data to be organised. The computer can learn from examples of input-output pairs and train the model to fit the data with high accuracy. Among the supervised learning algorithms are the following:

- Linear & Multivariate Regression
- Logistic Regression
- Naive Bayes
- Decision Trees
- K-nearest neighbour
- Linear Discriminant Analysis
- Artificial Neural Networks

Unsupervised Learning

The data in Unsupervised Learning is not labeled or categorized. Unsupervised learning allows data to organize itself after following a specific pattern in how the data is dispersed. Unsupervised learning algorithms are complicated

and are currently being researched. Unsupervised learning algorithms include the following:

- Clustering Analysis
- Anomaly Detection
- Hierarchical Clustering
- Principal Component Analysis

Reinforcement Machine Learning Algorithms

These algorithms are used to select an action. We can also see that it is based on each data point. Furthermore, after some time, the algorithm modifies its strategy in order to learn better. Obtain the best reward as well. Machine Learning is used in a variety of industries that require future prediction, pattern recognition, and autonomous decision making. It's widely used in the healthcare, finance, banking, manufacturing, and transportation industries.

Deep Learning

Deep Learning is a new field within the much larger field of Machine Learning. Deep Learning is best known for its neural networks, which include Recurrent Neural Networks, Convolutional Neural Networks, and Deep Belief Networks. Deep learning is modelled after the neurons of the human brain, whereas other machine learning algorithms use statistical analysis techniques for pattern recognition.

They are based on the structure and operation of the human brain. To understand deep learning, we must first understand how the nervous system in the human body works. As we all know, our nervous system is made up of neurons. These neurons can understand information that is transmitted to our bodies. These neurons can learn new information over time. Artificial neural networks use this 'learning' principle as well.

Any Deep neural network will consist of three types of layers (Figure 3):

- The Input Layer
- The Hidden Layer
- The Output Layer

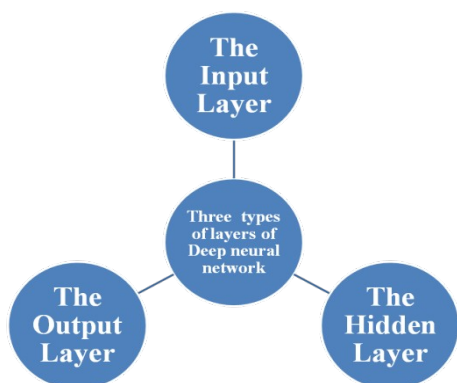


Figure 3. Three types of layers of Deep neural network

Slika 3. Tri vrste slojeva Duboke neuronske mreže

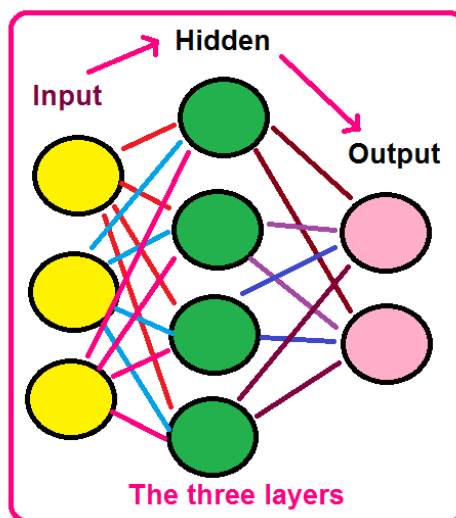


Figure 4. The three layers

Slika 4. Tri sloja

The layers in the given example (Figure 4) are in the form of an input layer, which accepts input data. The hidden layer, which performs various computations on the input data, and the output layer, which is binary in the above illustration. It should be noted that a neural network can have several hidden layers.

These neural networks are used to forecast (predict) the output as well as classify the data. The standard idea is that the neural network learns the pattern of data and then makes predictions that are similar to the pre-specified pattern.

Comparison of Deep Learning and Machine Learning

To parse data and learn from it, we use a machine algorithm. And make sound decisions based on what it has discovered. Deep learning is essentially used in layers to create an artificial "neural network." That is capable of learning and making intelligent decisions on its own.

Deep Learning and Machine Learning are compared through their features.

1. Data dependencies

The main distinction between the two algorithms is their performance. Deep learning algorithms, on the other hand, do not perform well when the data is small. This is the only reason Deep Learning algorithms require a large amount

of data to fully comprehend it. However, we can see that in this scenario, the use of algorithms with handcrafted rules prevails. This is depicted in Figure 5.

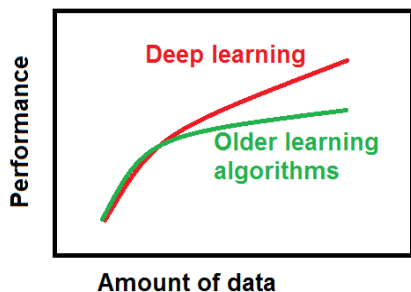


Figure 5. Deep learning and Older learning algorithms

Slika 5. Duboko učenje i stariji algoritmi učenja

2. Hardware dependencies

Deep learning is typically reliant on high-end machines. Traditional learning relies on low-end machines. As a result, GPUs are required for deep learning. That is an important part of how it works. They also perform a large number of matrix multiplication operations.

3. Feature engineering

It is a broad procedure. Domain knowledge is used to create feature extractors in this case. Also, to reduce the data's complexity. Make patterns more visible in order to learn how the algorithm works. Despite this, it is extremely difficult to process. As a result, it requires time and expertise.

4. Problem Solving approach

In most cases, we use traditional algorithms to solve problems. Despite this, it is necessary to divide a problem into distinct parts. Solve them individually as well. In addition, to get a result, combine them all. As an example: You are tasked with detecting multiple objects. Despite this, we must identify what the object is and where it appears in the image in this task. Furthermore, in a machine learning approach, the problem must be divided into two steps:

- Object detection
- Object recognition

First, we use the grabcut algorithm to scan the image for all possible objects. Then, from among all the recognised objects, you would use an object recognition algorithm such as SVM with HOG to identify relevant objects.

5. Execution time

Deep learning frequently takes longer to train than machine learning. The main reason for its lengthy duration is the large number of parameters in the deep learning algorithm. Machine learning, on the other hand, takes much less time to train, ranging from a few seconds to a few hours.

6. Interpretability

We used interpretability as a criterion to compare both learning techniques. Deep learning, on the other hand, is still thought about ten times before it is used in industry.

2. APPLICATIONS OF MACHINE LEARNING AND DEEP LEARNING

The following are some examples of real-world Deep Learning and Machine Learning applications:

Machine Learning technologies are increasingly being used in medical imaging. To detect tumours and other malignant growths in the human body.

Machine Learning-based time-series forecasting is being used in the field of marketing to predict sales.

Deep Learning is making significant contributions to the advancement of industrial robotics.

Machine learning algorithms are used in the self-driving car industry to guide the vehicle to its destination.

The industries are using Natural Language Processing to analyse customer reviews and gain insights into their sentiments.

Deep Learning-based recommendation systems are being used in the e-commerce industry to provide insights to customers based on their purchasing patterns.

Future Trends

Machine learning and data science are popular these days. Demand for both is rapidly increasing in businesses. Furthermore, their demand, particularly for some businesses. i.e. a company that wants to survive must incorporate Machine Learning into their operations. It is also necessary to understand basic terminologies.

Deep learning is discovered and proven to be the best technique with cutting-edge performance. As a result, deep learning continues to astound us and will continue to do so in the near future.

Recently, researchers have been concentrating on Machine Learning and Deep Learning. Previously, researchers were restricted to academia. However, ML and DL research is now finding a home in both industry and academia.

3. MACHINE LEARNING AND CORROSION INHIBITION STUDY

In the following section the application of machine learning and deep learning in corrosion inhibition research are presented [21-31].

Prediction of the inhibition efficiencies of magnesium dissolution modulators using sparse machine learning models

Small organic molecules can influence the degradation behaviour of magnesium and its alloys. Nonetheless, sophisticated tools are required for automatic recognition of efficient organic additives within the vast chemical space of potential compounds. Schiessler et al. [21] proposed two systematic sparse feature collection approaches for identifying molecular descriptors most relevant to the corrosion inhibition efficiency of chemical compounds. The first is based on the traditional statistical tool of analysis of variance, while the second is based on random forests. Schiessler et al. [21] show how, when combined with deep neural networks, both can aid in predicting the corrosion inhibition efficiencies of chemical compounds for the magnesium alloy ZE41. Schiessler et al. [21], in particular, demonstrated that this framework outperforms predictions based on a random selection of molecular descriptors. Finally, autoencoders could be used in the future to enable even more accurate automated predictions of corrosion inhibition efficiencies [21].

Machine learning analysis and computational modelling of the high effective inhibition of copper corrosion by an oxadiazole derivative

Varvara et al. [22] investigated an oxadiazole derivative with functional groups that favour its adsorption on the copper surface, namely 5-(4-pyridyl)-1,3,4-oxadiazole-2-thiol, as a potential inhibitor of copper corrosion in 3.5 wt. percent NaCl. High inhibition efficiencies are revealed by electrochemical impedance spectroscopy, potentiodynamic polarization, and SVET. Surface microscopy and spectroscopic analysis by Raman, SEM-EDX, and XPS reveal the formation of a compact barrier film responsible for long-term protection, which is primarily composed of organic molecules. Machine learning algorithms (decision trees, discriminant analysis, support vector machines (SVM), nearest neighbour classifiers (KNN), and ensemble classifiers) combined with Raman spectroscopy data were successfully used in corrosion studies to differentiate between corroded and inhibitor-protected metal surfaces. The prediction model obtained was successful.

Quantum Chemistry calculations in aqueous solution and Molecular Dynamic studies predict a strong interaction between copper and the thiol group, as well as widespread coverage of the metal surface, which is responsible for the exceptional corrosion resistance.

Quantitative structure activity relationship and artificial neural network as tools in predicting coordination capabilities of organic compounds with metal surface

Organic corrosion inhibitors have been shown to frequently form a protective film during coordinate bonding with metal. Quadri et al. [23] used a variety of computational and experimental methods to depict the nature and efficiency of such metal-inhibitor bonds. The quantitative structure activity relationship (QSAR) is one of the most recent and dependable computational methods for explaining metal-inhibitor coordination, which leads to corrosion inhibition. The pursuit of novel, high-performance environmentally benign compounds that can successfully inhibit corrosion without requiring extensive large-scale experimental trials has fine-tuned research interest in the molecular structure-corrosion inhibition relationship. The relationship between corrosion inhibition potentials and organic compound molecular descriptors is becoming more sophisticated. As computer technology advances, new techniques such as machine learning, artificial neural network (ANN), support vector machine (SVM), and genetic function approximation (GFA) are becoming more popular.

Prediction of corrosion inhibition efficiency of pyridines and quinolines on an iron surface

Ser et al.[24] developed linear and non-linear quantitative structure–property relationship (QSPR) models to predict corrosion inhibition efficiency for a series of 41 pyridine and quinoline N-heterocycles. Consensus models were built using the genetic algorithm-partial least squares (GA-PLS) and genetic algorithm-artificial neural network (GA-ANN) methods from 20 physicochemical and quantum chemical variables related to the surface adsorption behaviour of the inhibitors. The consensus GA-PLS model, which included eight variables (calculated adsorption energy, LUMO, van der Waals surface area and volume, polarizability, electron affinity, electrophilicity, and electron donor capacity), had a percent RMSEC of 16.5 percent and a mean percent RMSE of 14.9

percent. Such a model captured the complex relationships between inhibition efficiency and quantum chemical variables fairly well. In terms of predictive ability, the consensus GA-ANN model with nine input variables (exponential of calculated adsorption energy, HOMO, LUMO, HOMO-LUMO Gap, electronegativity, softness, electrophilicity, electron donor capacity, and N atomic charge) demonstrated a percent RMSECV of 16.7 2.3 percent and a mean percent RMSE (training/testing/validation) of 8.8 percent. Both models demonstrated the significance of adsorption to the metal surface, as well as electronic parameters quantifying electron acceptance/donation to/from the iron surface, as key corrosion inhibition design principles.

In silico (performed on computer) screening of modulators of magnesium dissolution

Because of the large number of tiny molecules with potentially beneficial dissolution modulating properties (inhibitors or accelerators), current experimental discovery methods are time- and resource-consuming. Fortunately, new computer-assisted methods can discover large areas of chemical space with less effort. Feiler et al. [25] demonstrated how density functional theory calculations and machine learning methods can be used in tandem to create robust and predictive models that recapitulate experimentally-derived corrosion inhibition efficiencies of small organic compounds for pure magnesium. Feiler et al. [25] validated their methods further by predicting a priori the corrosion modulation properties of seven previously untested small molecules and confirming the prediction in subsequent experiments.

Machine Learning-Based QSAR Model for Benzimidazole Derivatives as Corrosion Inhibitors

Computational prediction of inhibition efficiency (IE) for inhibitor molecules is an important supplementary method for designing novel molecules that effectively inhibit corrosion on metallic surfaces. Liu et al. [26] set out to create a new machine learning-based predictor of the inhibition efficiency (IE) of benzimidazole derivatives. The methods listed below have been used. First, a comprehensive numerical representation of inhibitor molecules was provided based on 3-D structures, covering all aspects of energy, electronic, topological, physicochemical, and spatial properties, yielding 150 valid structural

descriptors. Following that, a systematic investigation of these structural descriptors was carried out. To remove the linearly correlated feature variables, a multicollinearity-based clustering analysis was performed, yielding 47 feature clusters. Meanwhile, Gini importance by random forest (RF) was used to assess the contributions of the descriptors in each cluster, and 47 non-linear descriptors with the highest Gini importance score in the corresponding cluster were chosen. Furthermore, given the limited number of available inhibitors, diverse feature subsets of 47 descriptors were created using the Gini importance score ranking list. Finally, leave-one-out cross validation was used to test support vector machine (SVM) models based on different feature subsets. During the comparisons, the optimal SVM model with the top 11 descriptors was obtained using the Poly kernel. This model performs well, with correlation coefficients (R) and root-mean-square errors (RMSE) of 0.9589 and 4.45, respectively, indicating that the proposed method performs best for the current data. Six new benzimidazole molecules were designed using the proposed model, and the inhibition efficiency values predicted by this model indicate that two of them have a high potential as outstanding corrosion inhibitors.

Materials Design in Digital Era

The role of materials in our daily lives in relation to applications in areas such as energy, environment, manufacturing, and healthcare cannot be overstated. Transitioning a new material from its initial discovery to practical use, on the other hand, frequently involves significant cost and time. Nonetheless, the materials community has recently witnessed a significant shift in the way new materials are being designed. Current advances in computational capabilities, combined with improved quantum mechanical algorithms, have paved the way for accelerated material design and screening. Jain et al. [27] described how they use modelling and simulation to design/screen materials for a variety of applications including transdermal drug delivery, high-strength alloys, lithium-ion batteries, corrosion inhibition, mineral processing, and rare-earth element recovery. Each of the preceding examples employs high-performance computing-based first-principle simulations (ranging from the electronic to molecular scales) to arrive at a promising material, generating massive amounts of 'data' in the process. The amount of data accumulated as a result of these simulations is so massive that the term "data" is no longer restricted

to the computer science community. Furthermore, it appears critical at this point to authenticate how big data has simply scraped research in the materials domain.

Data science based Mg corrosion engineering

Magnesium has a high potential for a wide range of applications, including transportation, energy, and medicine. Untreated magnesium alloys, on the other hand, are prone to corrosion, limiting their practical application. As a result, in order to adapt to the explicit needs of the application, it is necessary to develop new approaches that can prevent or control corrosion and degradation processes. One potential solution is to use corrosion inhibitors, which can significantly reduce the rate of degradation caused by interactions with the metal surface or components of the corrosive medium. Because the sheer number of potential dissolution modulators makes obtaining a detailed atomistic understanding of the inhibition mechanisms for each additive impractical, other inhibition prediction measures are required. As a result, a concept is presented that combines corrosion experiments, machine learning, data mining, density functional theory calculations, and molecular dynamics to estimate corrosion inhibition properties of molecules that have yet to be tested. Concurrently, this approach will provide a deeper understanding of the fundamental mechanisms underlying corrosion prevention in magnesium-based materials, as well as more accurate continuum corrosion simulations. The proposed concept simplifies the search for molecules that have a positive or negative effect on inhibition efficiency and could thus significantly contribute to better control of magnesium / electrolyte interface properties [28].

In-silico model for predicting the corrosion inhibition efficiency of steel inhibitors

Models based on Quantitative Structure-Activity Relationships (QSAR) have been widely used to predict the corrosion inhibition performance of metals. One of the major limitations of these studies is that the authors have limited themselves to using only a single class of molecules with similar molecular structure. A computational end-to-end framework was developed in this study to investigate the properties of organic corrosion inhibitors, which are responsible for steel inhibition in acidic solution. The framework is made up of modules such as data preprocessing, descriptor selection, and model construction. Using advanced machine learning algorithms such as gradient

boosting machine (GBM), random forest, and support vector machines (SVM), a robust predictive model for multiple classes of corrosion inhibitors was developed. The descriptors were chosen using an innovative integrated ensemble technique. The model based on the GBM algorithm was significantly more accurate in predicting the corrosion inhibition efficiency of inhibitors [29].

Optimization of learning algorithms in the prediction of pitting corrosion

Boukhari et al. [30] investigated how to prevent pitting corrosion of a nuclear installation's open cooling circuit. Several corrosion inhibitors have been investigated. A variety of criteria were used to discuss and compare the performance of pitting corrosion inhibition. The experimental statistics were compiled into a large table and subjected to algorithms to create models for predicting corrosion inhibition performance. In this study, four algorithms were used: Genetic Algorithm-Artificial Neural Network (GA-ANN), Least Squares-Support Vector Machine (LS-SVM), K Nearest Neighbors (KNN), and Regression Tree (RT). Boukhari et al. [30] optimised the data fraction reserved for learning, as well as the parameters specific to each algorithm. The study's findings revealed that the algorithms perform in the following order: RT KNN LS-SVM GA-ANN.

Predicting the performance of organic corrosion inhibitors

The withdrawal of effective but toxic corrosion inhibitors has fueled the search for new, safe organic compounds to fill that role. Concurrently, advances in high-throughput organic compound synthesis, the establishment of large libraries of available chemicals, accelerated corrosion inhibition testing technologies, and the increased capability of machine learning methods have made the discovery of new corrosion inhibitors much faster and less expensive than in the past. Winkler summarises these technological advances in the corrosion inhibition field and describes how data-driven machine learning methods can generate models linking molecular properties to corrosion inhibition that can be used to predict the performance of materials that have not yet been synthesised or tested. Winkler has also published models of quantitative structure-property relationships for small organic molecule corrosion inhibitors. The success of these models lays the groundwork for the rapid development of novel,

effective corrosion inhibitors for a wide range of metals and alloys in a variety of environments [31].

4. DISCUSSION

Applications of machine learning in corrosion prediction

The Materials Genome Initiative (MGI) has recently resulted in concentrated efforts to collect materials properties in the form of shared databases. As a result, data-driven approaches have provided effective tools for driving the prediction of properties and the discovery of new materials. In the era of big materials data, machine learning (ML) has led to a conceptual leap in the field (especially in energy storage applications). Although ML has been gradually applied to corrosion research, the corrosion community has benefited far less from the progress in Big Data technologies. Corrosion data is typically incomplete, noisy, heterogeneous and large in volume (low-value density) [32].

Moreover, the in-service corrosive scenarios are complex and changeable, constituting a highly nonlinear system hardly approachable by traditional statistical methods. ML is the specific area of AI that allows computers to learn from data solving a given task. It consists of a flexible fitting function approach that, in comparison to traditional computational methods, can provide cheap and accurate simulation processes. ML aims to acquire knowledge from (very) large datasets by continuously improving their own performance.

Despite the lack of prior knowledge of the system, knowledge mining based on ML methods could help the domain expert link outcomes to underlying physical laws or confirm some of the already established concepts. A variety of electrochemical techniques has been employed for the prediction of the corrosion behaviour of metals. Accelerated corrosion tests and weathering tests have been historically used to simulate the working conditions of various systems. However, the simulated environmental factors in accelerated tests are distributed in ranges that often differ from the actual application. While in weathering tests, the environmental factors are time-varying with periodicity randomness. As there are no clear principles on setting fully representative test profiles, the extrapolation of short-term corrosion test results to field exposures in different environments is a challenging topic.

Despite disposing of huge volumes of testing data, the corrosion community still relies on inaccurate predictive models (log-linear, time-varying function, dose-response function, power

function, Eyring and Arrhenius) of the in-service corrosion rate (CR).

The selection of the most suitable ML approach is a multifaceted question, depending on the amount of data available, the quality of results desired and the necessity of physical interpretation of the results. Therefore, the present investigation highlights assessing the predictive power of different machine learning approaches and elaborate on the current status of regression modelling for various corrosion topics. The data mining approach proposed here also aims at identifying unseen patterns from the intersectional ML and corrosion literature (knowledge mining). We have to encourage new experimental and computer-oriented scientists to apply ML to corrosion prediction studies by discussing challenges and providing recommendations.

ML prediction of various types of corrosion

ML prediction of various types of corrosion is discussed in this section.

Atmospheric corrosion

Atmospheric corrosion data, traditionally acquired by coupon exposure tests, is often associated with small sample sizes and high dimensionality defined by multiple corrosion-influencing factors. The complexity of the corrosive environment makes it highly challenging for traditional predictive models to make optimal decisions in short timescales. In contrast to previous works using a linear regression technique, the ANN seemed to be a promising modelling tool for addressing nonlinear/complex systems based on interpolation from past experience. Through sensitivity analysis, ANN could also demonstrate that the corrosion rates of zinc were most likely affected by changes in corrosion products. A combined nonlinear grey Bernoulli model (NGBM) with genetic algorithm (GA) for long-term prediction on a specific monotonic data series of atmospheric corrosion rate.

Corrosion-resistant alloys

Corrosion prediction is challenging for Corrosion-resistant alloys (CRAs), as corrosion tests require considerable time periods for reaching steady states and appreciable weight losses. Weight loss methods and Artificial Neural Networks (ANN) approaches are recommended.

Marine corrosion

As an early example of ML work, SVR has been combined with particle swarm optimisation (PSO) for predictive modelling of marine corrosion. Nature-inspired metaheuristic optimisation algorithms such as the firefly algorithm (FA) have effectively solved difficult optimization problems. A

hybrid model was recently implemented by integrating a smart firefly algorithm (SFA) with a least-squares SVR.

Pipeline corrosion

Understanding and predicting the risk of pipeline corrosion is essential for maintaining oil operations healthy and safe. The first step in the Corrosion Risk Assessment of oil and gas companies is selecting the input parameters for estimating the corrosion defect depth amongst many parameters routinely measured. For estimating the defect depth growth of aged pipelines, we can implement a data-driven ML approach relying on PSO, Feed-Forward Artificial Neural Network (FFANN), Gradient Boosting Machine (GBM), RF and Deep Neural Network (DNN). The influence of PCA-transformed (Principal Component Analysis) variables on the accuracy of the models was also conducted.

By considering several network inputs, an ANN ensemble model could correctly identify the regions of pipelines where corrosion was most likely. This strategy can improve not only the prediction performance obtained with deterministic models but also with single ANN models.

Rebar corrosion

The highly nonlinear nature of embedded steel corrosion and the lack of theoretical basis for some corrosion phenomena render the predictive modelling in this topic difficult. ANN can be combined with Kohonen Self-Organized Mapping (KSOM) for modelling the chloride threshold (CT) in reinforced concrete from a sparse literature database.

Crevice corrosion

The ANN model is proved to be efficient for modelling localised corrosion. Based on crevice corrosion data on Ti-2, the ANN predicted the alloy's penetration depth over much longer times than the timescale of the experimental data available. The ANN model is not only able to predict the weight loss of the corrosion-resistant Alloy 22 (Nuclear corrosion topic) but also the alloy's localised behaviour.

Corrosion inhibitors

Benevolent and competent corrosion inhibitors are needed for overpassing the restrictions of toxic chromate-based technologies in the aerospace industry. In the beginning ML showed a negligible correlation between quantum chemical properties (calculated by DFT) and inhibition, suggesting that published structure-inhibition models available in the literature were incorrect (these were

presumably based on minimum numbers of molecules). Then, different robust models (including the BRANNGP (feed forward Bayesian regularised neural networks using a Gaussian prior)) were developed, linking molecular properties calculated by non-quantum chemical methods and inhibition.

Future perspective

It will be very interesting to make use of HOMO-LUMO approach, to calculate various quantum chemical parameters such as the energy gap, electronegativity, ionization potential, softness of the molecules etc., and finally to predict the use of various molecules as efficient corrosion inhibitors.

5. CONCLUSION

This study looked into the use of artificial intelligence in the efficient prediction of corrosion inhibitors using mainly machine learning methodologies.

In the investigations, artificial neural networks (ANN), support vector machines (SVM), and genetic function approximation (GFA) were utilized to estimate the coordination capabilities of organic molecules with metal surfaces. For predicting physiochemical and quantum variables relevant to the surface absorption behavior of inhibitors, researchers used the genetic algorithm-partial least squares (GA-PLS) and genetic algorithm-artificial neural network (GA-ANN) methods. The prediction of Benzimidazole derivatives as corrosion inhibitors was conducted using random forest (RF) and support vector machine (SVM) models. Models for forecasting the corrosion inhibition performance of steel inhibitors with high accuracy include the gradient boosting machine (GBM), random forest (RF), and support vector machine. For pitting corrosion prediction, genetic algorithm-Artificial Neural Network (GA-ANN), Least Square-Support Vector Machine (LS-SVM), K-Nearest Neighbor (KNN), and Regression Tree (RT), supervised and unsupervised methodologies.

Despite the high accuracy that could be obtained using the concept of deep learning algorithms, which is actually based on mimicking biological neurons using very large training and testing datasets, studies revealed that little work has been done using deep learning techniques for corrosion inhibitor prediction.

Deep learning methods such as Convolution Neural Network (CNN), Multi-Layer Perception (MLP), and Recurrent Neural Networks (RNN)

should be thoroughly researched in the future for use in better predicting corrosion inhibitors, according to the study.

6. REFERENCES

- [1] Y. Wang, L. Niu, L. Zhao, B. Wang, Z. He, H. Zhang, D. Chen, J. Geng (2022) Gaussian mixture model deep neural network and its application in porosity prediction of deep carbonate reservoir, *Geophysics*, 87(2), 59-72.
- [2] C. Deng, Y. Wang, C. Qin, Y. Fu, W. Lu (2022) Self-directed online machine learning for topology optimization, *Nature Communications*, 13(1), 338.
- [3] X. Zhao, Y. Sui, X. Ruan, X. Wang, K. He, W. Dong, H. Qu, X. Fang (2022) A deep learning model for early risk prediction of heart failure with preserved ejection fraction by DNA methylation profiles combined with clinical features, *Clinical Epigenetics*, 14(11), 1-15.
- [4] F. Qiu, G. Zhang, X. Sheng, L. Jiang, L. Zhu, Q. Xiang, B. Jiang, P. K. Chen (2022) Predicting students' performance in e-learning using learning process and behaviour data, *Scientific Reports*, 12(453), 1-15.
- [5] H. Ozaki, T. Aoyagi (2022) Prediction of steady flows passing fixed cylinders using deep learning, *Scientific Reports*, 12(447), 1-12.
- [6] X. Sun, H. Zhang, J. Wang, C. Shi, D. Hua, J. Li (2022) Ensemble streamflow forecasting based on variational mode decomposition and long short term memory, *Scientific Reports*, 12(518), 1-19.
- [7] J. J. Thiagarajan, K. Thopalli, D. Rajan, P. Turaga (2022) Training calibration-based counterfactual explainers for deep learning models in medical image analysis, *Scientific Reports*, 12(597), 1-15.
- [8] Y. Che, Z. Deng, X. Tang, X. Lin, X. Nie, X. Hu (2022) Lifetime and Aging Degradation Prognostics for Lithium-ion Battery Packs Based on a Cell to Pack Method, *Chinese Journal of Mechanical Engineering*, 35(4), 1-16.
- [9] Y. Essam, Y. F. Huang, A. H. Birima, A. N. Ahmed, A. El-Shafe (2022) Predicting suspended sediment load in Peninsular Malaysia using support vector machine and deep learning algorithms, *Scientific Reports*, 12(302), 1-29.
- [10] T. Kakati, D. K. Bhattacharyya, J. K. Kalita, T. M. Norden-Krichmar (2022) DEGnext: classification of differentially expressed genes from RNA-seq data using a convolutional neural network with transfer learning, *BMC Bioinformatics*, 23(17), 1-18.
- [11] M. A. H. Newton, F. Mataeimoghadam, R. Zaman, A. Sattar (2022) Secondary structure specific simpler prediction models for protein backbone angles, *BMC Bioinformatics*, 23(6), 1-14.
- [12] J. Liao, Y. Liu, Y. Piao, J. Su, G. Cai, Y. Wu (2022) GLE-Net: A Global and Local Ensemble Network for Aerial Object Detection, *International Journal of Computational Intelligence Systems*, 15(2), 1-12.
- [13] A. Ranjan, S. Shukla, D. Datta, R. Misra (2022) Generating novel molecule for target protein (SARS-CoV-2) using drug-target interaction based on graph neural network, *Network Modeling Analysis in Health Informatics and Bioinformatics*, 11(6), 1-11.
- [14] J. Mena, O. Pujol, J. Vitrià (2022) A Survey on Uncertainty Estimation in Deep Learning Classification Systems from a Bayesian Perspective, *ACM Computing Surveys*, 54(9), 1-35.
- [15] Boyu He, Lin Ye, Ming Pei, Peng Lu, Binhua Dai, Zhuo Li, Kaifeng Wang (2022) A combined model for short-term wind power forecasting based on the analysis of numerical weather prediction data, *Energy Reports*, 8, 929-939.
- [16] X. Zhang, X. Zheng, W. Mao (2022) Adversarial Perturbation Defense on Deep Neural Networks, *ACM Computing Surveys*, 54(8), 1-36.
- [17] B. M. Oloulade, J. Gao, J. Chen, T. Lyu, R. Al-Sabri (2022) Graph Neural Architecture Search: A Survey, *Tsinghua Science and Technology*, 27(4), 692-708.
- [18] V. E. Atasoy, A. E. Suzer, S. Ekici (2022) Comparative Analysis of Exhaust Gas Temperature Based on Machine Learning Models for Aviation Applications, *Journal of Energy Resources Technology, Transactions of the ASME*, 144(8), 082101.
- [19] J. Zhang, Y. Jiang, S. Wu, X. Li, H. Luo, S. Yin (2022) Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism, *Reliability Engineering and System Safety*, 221, 108297.
- [20] Huy Vu Ngoc, J. R. R. Mayer, Elie Bitar-Nehme (2022) Deep learning LSTM for predicting thermally induced geometric errors using rotary axes' powers as input parameters, *CIRP Journal of Manufacturing Science and Technology*, 37, 70-80.
- [21] E. J. Schiessler, T. Würger, S. V. Lamaka, R. H. Meißner, C. J. Cyron, M. L. Zheludkevich, C. Feiler, R. C. Aydin (2021) Predicting the inhibition efficiencies of magnesium dissolution modulators using sparse machine learning models, *Npj Computational Materials*, 7(1), 193-199.
- [22] S. Varvara, C. Berghian-Grosan, R. Bostan, R. Lucacel Ciceo, Z. Salarvand, M. Talebian, K. Raeissi, J. Izquierdo, R. M. Souto (2021) Experimental characterization, machine learning analysis and computational modelling of the high effective inhibition of copper corrosion by 5-(4-pyridyl)-1,3,4-oxadiazole-2-thiol in saline environment, *Electrochimica Acta*, 398, 139282.
- [23] T. W. Quadri, L. O. Olasunkanmi, O. E. Fayemi, E. D. Akpan, C. Verma, E. M. Sherif, K. F. Khaled, E. E. Ebenso (2021) Quantitative structure activity relationship and artificial neural network as vital tools in predicting coordination capabilities of organic compounds with metal surface: A review, *Coordination Chemistry Reviews*, 446, 214101.
- [24] C. T. Ser, P. Žuvela, M. W. Wong (2020) Prediction of corrosion inhibition efficiency of pyridines and quinolines on an iron surface using machine learning-powered quantitative structure-property

- relationships, Applied Surface Science, 512, 145612.
- [25] C. Feiler, D. Mei, B. Vaghefinazari, T. Würger, R. H. Meißner, Bérengère J. C. Luthringer-Feyerabend, D. A. Winkler, M. L. Zheludkevich, S. V. Lamaka (2020) In silico screening of modulators of magnesium dissolution, Corrosion Science, 163, 108245.
- [26] Y. Liu, Y. Guo, W. Wu, Y. Xiong, C. Sun, L. Yuan, M. Li (2019) A Machine Learning-Based QSAR Model for Benzimidazole Derivatives as Corrosion Inhibitors by Incorporating Comprehensive Feature Selection, Interdiscip Sci INTERDISCIP SCI, 11(4), 738-747.
- [27] D. Jain, B. S. Dwadasi, D. Kumar, S. Mishra, B. Ravikumar, R. Gupta, S. G. Srinivasan, V. Jain, M. Mynam, S. Maiti, B. Rai (2019) Materials Design in Digital Era: Challenges and Opportunities, Transactions of the Indian Institute of Metals, 72(8), 2199-2208.
- [28] T. Würger, C. Feiler, F. Musil, G. B. V. Feldbauer, D. Höche, S. V. Lamaka, M. L. Zheludkevich, R. H. Meißner (2019) Data science based Mg corrosion engineering, Frontiers in Materials, 6(53), 1-9.
- [29] A. Agarwal, P. Rathore, V. Jain, B. Rai (2019) In-silico model for predicting the corrosion inhibition efficiency of steel inhibitors, NACE - International Corrosion Conference Series, March, 13329.
- [30] Y. Boukhari, M. N. Boucherit, M. Zaabat, S. Amzert, K. Brahimi (2018) Optimization of learning algorithms in the prediction of pitting corrosion, Journal of Engineering Science and Technology, 13(5), 1153-1164.
- [31] D. A. Winkler (2017) Predicting the performance of organic corrosion inhibitors, Metals, 7(12), 553.
- [32][32] L. Bertolucci Coelho, D. Zhang, Y. Van Ingelgem, D. Steckelmacher, An. Nowé, H. Terry (2022) Reviewing machine learning of corrosion prediction in a data-oriented perspective, 6(8), 1-16.

IZVOD

PRIMENA MAŠINSKOG UČENJA U PROUČAVANJU INHIBICIJE KOROZIJE

Veštačka inteligencija je grana nauke koja se bavi učenjem mašina da misle i deluju kao ljudi. Mašinsko učenje se bavi omogućavanjem računarima da izvršavaju zadatke bez potrebe za eksplicitnim programiranjem. Mašinsko učenje omogućava računarima da uče bez potrebe za eksplicitnim programiranjem. Mašinsko učenje je široko polje koje obuhvata širok spektar operacija mašinskog učenja kao što su grupisanje, klasifikacija i razvoj prediktivnih modela. Istraživanje mašinskog učenja (ML) i dubokog učenja (DL) sada pronalazi dom i u industriji i u akademskim krugovima. Tehnologije mašinskog učenja se sve više koriste u medicinskom snimanju. Za otkrivanje tumora i drugih malignih izraslina u ljudskom telu. Duboko učenje daje značajan doprinos napretku industrijske robotike. Algoritmi mašinskog učenja se koriste u industriji automobila koji se samostalno voze da vode vozilo do odredišta. Duboko učenje i mašinsko učenje se, takođe, koriste u nauci o koroziji i inženjerstvu. Koriste se za odabir molekula inhibitora iz velikog skupa dostupnih molekula.

Ključne reči: *veštačka inteligencija, mašinsko učenje, duboko učenje, neuronske mreže, algoritmi, ulazni sloj, skriveni sloj i izlazni sloj.*

Naučni rad

Rad primljen: 15. 02. 2022.

Rad korigovan: 10. 03. 2022.

Rad prihvacen: 14. 03. 2022.

Rad je dostupan na sajtu: www.idk.org.rs/casopis

© 2022 Authors. Published by Engineering Society for Corrosion. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International license (<https://creativecommons.org/licenses/by/4.0/>)