Doktorandusz Almanach

PhD Students Almanach

Anyag- és Vegyészmérnöki Kar
Miskolci Egyetem

Faculty of Materials and Chemical Engineering
University of Miskolc

2022-1.kötet/1.volume
ISSN 2939-7294

ISSN 2939-7294

Miskolci Egyetem, Anyag- és Vegyészmérnöki Karának kiadványa

A publication of the University of Miskolc, Faculty of Materials and Chemical Engineering.

3515 Miskolc Egyetemváros

Kiadásért felelős/Responsible for publishing: Prof. Dr. Palotás Árpád

Szerkesztő/Editor: Dr. Kocserha István

Doktorandusz Almanach

PhD Students Almanach

2022/1.kötet/1.volume

Miskolc, 2022

CONTENTS

The applicability of machine learning in polyurethane design

Lucjuk-Huszár Kornélia [1], Béla Viskolcz [1,2], Attila Garami [3], Béla Fiser [1,2,4,5*]

[1] Institute of Chemistry, University of Miskolc, 3515 Miskolc, Hungary
[2] Higher Education and Industrial Cooperation Centre, University of Miskolc, 3515 Miskolc, Hungary
[3] Department of Combustion Technology and Thermal Energy, University of Miskolc, 3515 Miskolc, Hungary
[4] Ferenc Rakoczi II Transcarpathian Hungarian College of Higher Education, 90200 Beregszász, Transcarpathia, Ukraine
[5] Department of Physical Chemistry, Faculty of Chemistry, University of Lodz, 90-236 Lodz, Poland

E-mail: kemfiser@uni-miskolc.hu

Abstract. Polyurethanes (PU) are used in wide range of products (*e.g.* construction materials). The properties of polyurethane-based materials can be modified and fine-tuned by using additives (*e.g.* fillers). New synthetic recipes are developed to create better materials is almost exclusively based on trial-and-error cycles which is a time and material intensive process. By using machine learning (ML) algorithms the process can be significantly accelerated. Therefore, to develop new polyurethane types, we are proposing to combine the strength of computational tools with experimental methods and data.

1. Introduction

Due to their versatility, polyurethane-based products are present in almost all areas of our daily life. PU is a special polymeric material applied in a wide range of different products and we are in touch with such materials on a daily basis. Therefore, polyurethanes are essential, and their recycling and the design of better PU types is highly important for our society. As we know, the environmental impact of plastic pollution is serious, therefore we need to pay special attention to develop new environmentally friendly materials as soon as possible. The development process can be accelerated by using machine learning (ML) algorithms. In the current project proposal, the strength of experimental and computational tools to achieve new polyurethane types will be combined. Monomers are the building blocks of polymers which are connected through the process called polymerization and thus, polymers will form. In the case of polyurethanes, the building blocks are called isocyanate and polyol and through their reaction

*Corresponding author.

these two different types of base materials form urethane bond and those structures which contain several urethane bonds are called polyurethanes (PU).
Since the beginning of evolution, humans have used various types of tools to perform a wide range of tasks more easily. The latest step in this process was the implementation of machine learning (ML). Thanks to ML, machines can learn how to manage data more efficiently. Occasionally, after reviewing the data, we cannot get the information we need. In this case, machine learning can be used to handle situations that are too complex or time-consuming for humans. The need for machine learning grows as the data increases. ML has applications in many scientific fields, one of which its use in the development of new synthetic formulations for the manufacture of environmentally friendly polyurethanes, which will be tested in the near future as is being proposed.

2. Materials and methods

Before a machine can create a model by themselves, a learning database must first be generated (Figure 1). The database must collect and store a considerable amount of data. The power of data mining is demonstrated when the number of attributes is so large that we cannot process the data using traditional methods.

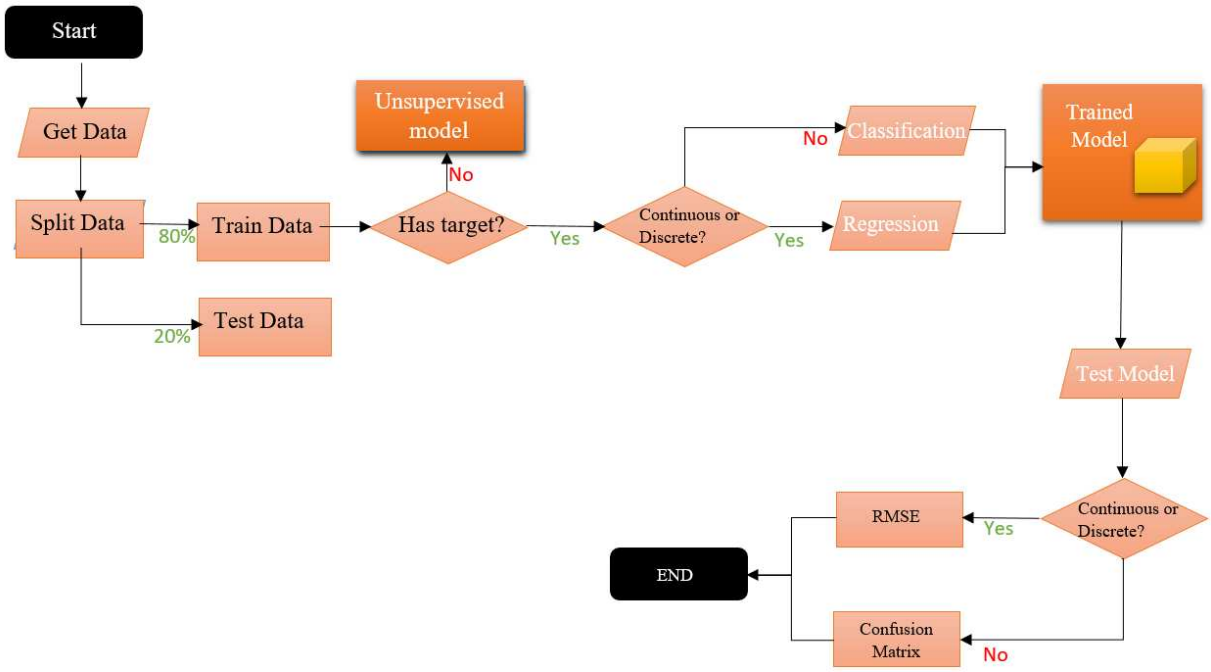

Figure 1. – General workflow of learning algorithms [1].

We will often use the phrase data table attribute instead of attribute. When each attribute is a number, instances can be mapped to points in a multidimensional space, so we can use the dimensional expression instead of attribute. The word instance also refers to object, element, record in the literature. The variable whose value you want to predict or estimate is called an explained (or to be explained) variable, also known as a class attribute (target). All other variables, other than the explained variable (class attribute) are called explanatory variables. We talk about regression when we measure the explained variable on an intervallum scale. If the explained variable has a discrete set of values, measured on a nominal or ordinal scale, we

talk about classification or clustering. Classification is often referred to in the statistical literature as discriminant analysis [2].

Using the database, the learning algorithm builds a model and not only learns the patterns "by heart", but after the learning phase, it is able to generalise from them to make "correct" decisions or predictions about unknown data.
There are several different learning algorithms available, which should be chosen according to the task at hand, in order to discover the internal structures of the database from the initial data. The choice of the right algorithm is critical, considering the data to be analysed and the properties to be predicted. Four types of learning algorithms can be distinguished: supervised, unsupervised, semi-supervised and reinforced. The supervised type includes classification and regression algorithms, while the unsupervised type includes the estimation of fidelity functions. Supervised learning means that the training data contain output labels to achieve data classification or regression. If the training data have no output labels, this type of ML is called unsupervised learning, which can achieve data clustering or dimensionality reduction, such as K-means clustering and principal component estimation. Reinforcement learning is an iteration method in which an agent takes steps to change its state and interact with the environment to maximize the value of the target reward, such as Markov decision process and active learning. Therefore, successful application of ML requires suitable algorithms and sufficient data.
Certain studies have used well established ML models to quickly estimate the properties and assess the suitability of newly created polymer targets. The candidates with the best suitability score are retained as parent polymers and the above steps are repeated until a sufficient number of potential polymer candidates with the required properties are found [3].

Machine learning (ML) methods are becoming state-of-the-art in many fields, including materials science. The exploration of correlations between data sets has historically led to many significant results, but this requires two basic preconditions: reliable data and an inspired guess to where to look for the correlations. The increase in data processing capacity and the progress in intelligent search methodologies are said to be changing the course of some branches of science. In materials science, large databases may allow us to search for correlations between features that we would not normally look for. There are several studies where machine learning methods were used for polyurethane development [4].

Examples of applications
Polyurethane (PU) composites are increasingly applied as building materials for civil engineering structures such as road pavements, runways, parking lots and flooring to maintain building floor systems. In a recent work, polyurethane polymer concrete (PUPC) material was developed using two mixing ratios for the repair of road pavement, and runaway facilities. Three non-linear AI-based models including Gaussian process regression (GPR), classification and regression tree (CART), and support vector regression (SVR) model were applied to predict the compressive strength of PUPC mixtures as repair materials. The models are based on the flexural strength (MPa), density (kg/m^3) and PU composition (%) and these were used as input parameters. The results showed that the compressive stress-strain curves of PU-based polymer concrete exhibit linear elastic behaviour under compression. The developed models show high prediction accuracy of the strength of PUPC [5].

In another investigation, mechanical property predictions of polyurethane elastomers using the HML algorithm was carried out. The accuracy was compared against a random forest model, and it was found that HML produced significantly better predictions of the tested data. This

was attributed to the integration of an intermediate layer of variables comprising domain knowledge-based physicochemical factors which significantly improved the model relating experimental formulation variables and mechanical responses of the cured elastomers. the advantages of HML are the possibility of modelling categorical and qualitative responses of polyurethane products to formulation and processing variables, and predicting the properties of novel monomers, such as bio-based materials [6].

Machine learning approach (based on multilayer perceptron network, random forest, and support vector machine regression, respectively) was proposed for the prediction of creep properties of polyurethane elastomer considering the effect of creep time, creep temperature, creep stress, and the hardness of the material. The results showed that the three models all proposed excellent fitting ability for the training set. Moreover, the three models had different prediction capabilities for the testing set by focusing on various changing factors. The method could provide a new research idea for the accelerated representation of long-term mechanical properties of polymers [7].

Artificial neural network (ANN) technique has been used to model the temperature dependence of dynamic mechanical properties and viscoelastic behaviour of thermoplastic polyurethane (TPU) over a wide range of temperatures. The excellent agreement between the modelled and experimental data has been found over the entire investigated temperature interval, including all of the observed relaxation transitions. The multi-layer feed-forward back-propagation artificial neural network has been confirmed to be a very effective artificial intelligence tool for the modelling of dynamic mechanical properties and for the prediction of viscoelastic behaviour of the tested thermoplastic polyurethane within the whole temperature range of its service life. The ANN is used by many researchers for the analysis of the viscoelastic properties of polymers. The excellent predictive abilities of ANN models can also be used in the development of new TPU-based materials and polymer composites with the desired dynamic mechanical properties, relaxation transitions, and predictable viscoelastic behaviors [8].

Artificial neural network was also used in combination with surface response methodology (SRM) models, to study the influences of the molar mass of the polyol on the polymerization kinetics of polyurethanes. A proper kinetic model (autocatalytic) was selected to describe the experimental data. A high level of reliability of the predicted calorimetric curves was obtained due to an excellent agreement between theoretical and modelled results, enabling creating a 3D surface response to predict the reaction kinetics. It was possible to observe that the polymerization kinetics is affected by the OH group's association phenomena. The applied methodology can be extended for other materials or properties of interest [9].

By replacing synthetic materials with bio-alternatives in the PU formulation and keeping (or improving) their mechanical properties compared to conventional polyurethanes would be ideal. To expedite the procedure machine learning algorithms will be used to predict the mechanical and acoustic properties of polyurethanes prepared by using bio-based formulations.

3. Conclusion

Machine learning algorithms can be divided into four categories: supervised, unsupervised, semi-supervised and reinforced. The selection of the right algorithm is crucial, and during this process various factors have to be taken into consideration (*e.g.* properties). All in all, six general steps have to be followed when machine learning algorithms are used to develop new polyurethane types:

- Data collection on polyurethanes (and processing the data)
- Learning database creation
- Selection of the proper algorithms
- Model creation by applying the algorithms
- Applying the model to create new polyurethane formulations
- Testing the formulations by preparing the new PUs

The process of polymer development can be significantly expedited by using ML algorithms and thus, the polymeric materials could reach the users earlier.

Acknowledgement

This research was supported by the European Union and the Hungarian State, co-financed by the European Regional Development Fund in the framework of the GINOP-2.3.4-15-2016-00004 project, which aims to promote cooperation between higher education and industry. Further support was provided by the National Research, Development and Innovation Fund (Hungary) within the TKP2021-NVA-14 project. The GITDA (Governmental Information-Technology Development Agency, Hungary) is gratefully acknowledged for allocating the computing resources used in this work.

References

[1] Dey, A. Machine Learning Algorithms: A Review. *Int. J. Comput. Sci. Inf. Technol.* **7**, 1174–1179 (2016).

[2] Bodon, F. & Búza, K. Adatbányászat, Budapesti Műszaki és Gazdaságtudományi Egyetem, Egyetemi jegyzet (2014), http://www.cs.bme.hu/nagyadat/bodon.pdf.

[3] Kim, C., Batra, R., Chen, L., Tran, H. & Ramprasad, R. Polymer design using genetic algorithm and machine learning. *Comput. Mater. Sci.* **186**, 110067 (2021).

[4] Crivello, J. C., Joubert, J. M. & Sokolovska, N. Supervised deep learning prediction of the formation enthalpy of complex phases using a DFT database: The σ−phase as an example. *Comput. Mater. Sci.* **201**, 110864 (2022).

[5] Haruna, S. I. *et al.* Gaussian process regression model for the prediction of the compressive strength of polyurethane-based polymer concrete for runway repair: A comparative approach. *IOP Conf. Ser. Earth Environ. Sci.* **1026**, (2022).

[6] Menon, A., Thompson-Colón, J. A. & Washburn, N. R. Hierarchical machine learning model for mechanical property predictions of polyurethane elastomers from small datasets. *Front. Mater.* **6**, 1–12 (2019).

[7] Yang, C., Ma, W., Zhong, J. & Zhang, Z. Comparative study of machine learning approaches for predicting creep behavior of polyurethane elastomer. *Polymers.* **13**, 1768 (2021).

[8] Kopal, I., Harničárová, M., Valíček, J. & Kušnerová, M. Modeling the temperature dependence of dynamic mechanical properties and visco-elastic behavior of thermoplastic polyurethane using artificial neural network. *Polymers.* **9**, 519, (2017).

[9] Dall Agnol, L., Ornaghi, H. L., Monticeli, F., Dias, F. T. G. & Bianchi, O. Polyurethanes synthetized with polyols of distinct molar masses: Use of the artificial neural network for prediction of degree of polymerization. *Polym. Eng. Sci.* **61**, 1810–1818 (2021).