Machine Learning Models for Predicting Key Properties in Free Radical Emulsion Polymerization

Kiarash Farajzadehahary[1], Nicholas Ballard[1,2], José M Asua[1]
[1]POLYMAT University of the Basque Country UPV/EHU, Joxe Mari Korta zentroa, Tolosa Hiribidea 72, Donostia-San Sebastián 20018, Spain
[2]Ikerbasque, Basque Foundation for Science, Plaza Euskadi 5, Bilbao 48009, Spain

Kiarash.Farajzadeh@Polymat.eu

Predicting material properties is a critical step in advancing polymer science and enabling the design of new materials with tailored functionalities [1]. This study focuses on developing machine learning models trained on datasets compiled from the literature to accurately predict four key properties of monomers relevant to free radical emulsion polymerization: propagation rate constant [2], reactivity ratios in copolymerization [3], glass transition temperature [4], and water solubility [5]. These properties significantly influence polymerization kinetics, polymer composition, and the final characteristics of the resulting polymers.

Artificial neural networks (ANN) were employed to predict rate constant of propagation, reactivity ratios, and water solubility due to their ability to capture complex, nonlinear relationships inherent in chemical datasets. For glass transition temperature prediction, a meta-model was implemented by combining ANN with CatBoost, a gradient-boosting decision tree algorithm [6]. This hybrid approach enhanced predictive performance by leveraging the strengths of both models.

For all models, hyperparameter optimization was conducted using grid search, and training and testing were performed using 5-fold cross-validation to ensure robust and reliable evaluation. The predictive framework demonstrated excellent accuracy, achieving high R-squared values and low root mean square errors (RMSE) across both training and test datasets.

The developed models rely on chemical structure-based features such as molecular fingerprints and descriptors [7]. These inputs effectively represent monomer properties, enabling data-driven insights into how chemical structures correlate with the desired polymerization characteristics. This predictive framework reduces the dependency on extensive experimental testing, offering an efficient and reliable alternative for evaluating monomer properties.

While the current work emphasizes model development and validation, the assembled framework holds significant potential for broader applications. A key area of interest is the identification and optimization of bio-based monomers as sustainable alternatives to conventional petrochemical monomers.

Predicting key properties such as glass transition temperature (a critical parameter influencing minimum film formation temperature), water solubility (important for mass transfer considerations), and propagation rate constant and reactivity ratios (which govern reactivity and polymerization kinetics) can facilitate the selection of bio-based candidates that emulate or improve upon traditional monomers' performance in free radical emulsion polymerization [8].

The assembled model is a handy tool for researchers aiming to innovate in polymer design and applications. By providing accurate predictions for multiple key properties, these models address a gap in materials science, where experimental trial-and-error approaches can be resource-intensive and time-consuming. The integration of advanced machine learning techniques, including neural networks and ensemble methods, underscores the potential of artificial intelligence to accelerate innovation in polymer chemistry.

References

[1] C. Kuenneth, A.C. Rajan, H. Tran, L. Chen, C. Kim, R. Ramprasad, Polymer informatics with multi-task learning, Patterns 2 (2021) 100238.
[2] E. Van de Reydt, N. Marom, J. Saunderson, M. Boley, T. Junkers, A Predictive machine-learning model for propagation rate coefficients in radical polymerization, Polym. Chem. 14 (2023) 1622–1629.
[3] K. Farajzadehahary, X. Telleria-Allika, J.M. Asua, N. Ballard, An artificial neural network to predict reactivity ratios in radical copolymerization, Polym. Chem. 14 (2023) 2779–2787.
[4] S. Otsuka, I. Kuwajima, J. Hosoya, Y. Xu, M. Yamazaki, PoLyInfo: Polymer Database for Polymeric Materials Design, in: 2011 Int. Conf. Emerg. Intell. Data Web Technol., IEEE, 2011: pp. 22–29.
[5] P.G. Francoeur, D.R. Koes, SolTranNet–A Machine Learning Tool for Fast Aqueous Solubility Prediction, J. Chem. Inf. Model. 61 (2021) 2530–2536.
[6] A.V. Dorogush, V. Ershov, A. Gulin, CatBoost: gradient boosting with categorical features support, (2018).
[7] D. Rogers, M. Hahn, Extended-Connectivity Fingerprints, J. Chem. Inf. Model. 50 (2010) 742–754.
[8] F.J. Schork, Monomer transport in emulsion polymerization, Can. J. Chem. Eng. 100 (2022) 645–653.