Predicting Pure-Component Vapor Pressures with Graph Neural Networks

Marco Hoffmann[1, S], Dominik Gond[1], Hans Hasse[1] and Fabian Jirasek[1, C]

[1]*Laboratory of Engineering Thermodynamics, RPTU Kaiserslautern, Kaiserslautern, Germany*
fabian.jirasek@rptu.de

The vapor pressure is a crucial property for designing and optimizing various processes in chemical engineering, e.g., distillation. It is usually modeled using the Antoine equation, which must be parameterized to at least some experimental data points for each component of interest. However, since experimental data are scarce, prediction methods are paramount. Unfortunately, most currently available prediction methods for the vapor pressure are restricted to specific chemical classes and/or still require some experimental data for each component, e.g., normal boiling temperature, critical data, or acentric factor, which strongly limits their applicability [1]. Prediction methods based solely on structural formulas are rare, and their extrapolation capability to unseen components has not been studied systematically [2].

In the present work, we explore a new way of predicting the vapor pressure of pure components by using Graph Neural Networks (GNN), which only require information on the molecular structure of the component as input. Several versions of this approach were studied, including purely data-driven and hybrid models, in which physical knowledge of the properties of vapor-pressure curves is also used.

The new models were trained on vapor-pressure data for more than ten thousand components from the Dortmund Data Bank [3] and the DIPPR database [4]. As only the structural formula is needed as input, the new models have an unlimited scope. The predictive capabilities of the models were studied by retaining data from the training and using it for tests. The new approach yields clearly better results than established group-contribution methods for predicting the pure-component vapor pressure, e.g., [2]. Furthermore, a systematic analysis of the parameters of the GNN learned during the training gives insights into the importance of different molecular substructures for describing the vapor pressure.

References

1. Evangelista et al. Ind. Eng. Chem. Res., 56 (2017) 2298-2309.
2. C.-H. Tu. Fluid Phase Equilibria, 99 (1994) 105-120
3. Dortmund Data Bank, 2023, www.ddbst.com.
4. DIPPR 801 Database, 2023, www.aiche.org/dippr.