FOPAM 2026
Proceedings of the 3rd Foundations of Process/Product Analytics and Machine Learning
Decatur, USA, July 26-30 (2026)

Tennet-SAC: A Physics-Embedded Machine Learning Model for Activity Coefficient of Multicomponent Liquid Mixtures

Shiang-Tai Lin

National Taiwan University, Department of Chemical Engineering

ABSTRACT

We present TeNNet-SAC (Thermodynamics-embedded Neural Network for Segment Activity Coefficient) [1], a physics-embedded machine learning framework for predicting activity coefficients in multicomponent liquid mixtures directly from SMILES. Building upon the concept of segment-based thermodynamic foundation of COSMO-SAC model [2], TeNNet-SAC preserves physical interpretability while eliminating the need for quantum chemical calculations. The model comprises three components: (i) a σ-profile predictor that infers molecular surface charge distributions from SMILES, (ii) a geometry predictor for molecular volume and surface area, and (iii) a Γ predictor that computes segment activity coefficients. The σ-profile and geometry predictors are trained on 39,745 chemically diverse quantum-calculated structures, ensuring broad chemical coverage. The Γ predictor is designed to enforce thermodynamic consistency and is pretrained on one million synthetic data points to reproduce segment activity coefficients from the COSMO-SAC model, followed by fine-tuning using 34,371 curated experimental data points from VLE measurements of 397 binary mixtures. TeNNet-SAC achieves accuracy comparable to—and after fine-tuning, exceeding—the COSMO-SAC model, while inherently satisfying thermodynamic consistency and naturally extending to multicomponent systems. To facilitate adoption, the full implementation is openly available on GitHub [3], along with an easy-to-use web interface for rapid prediction without specialized software [4]. In addition, both the code and the web platform provide functionality to generate NRTL parameters directly, enabling seamless integration with Aspen Plus for process simulation. This combination of physical rigor, scalability, and accessibility makes TeNNet-SAC a practical tool for phase equilibrium modeling and process design.

REFERENCES

1. Yang Y, Lin ST. Physics-Embedded Machine Learning Model for Phase Equilibrium Prediction in Multicomponent Systems. *J Chem Info Mod* 65:10180–10193 (2025). https://doi.org/10.1021/acs.jcim.5c01804
2. Lin ST, Sandler SI. A priori phase equilibrium prediction from a segment contribution solvation model. *Ind Eng Chem Res* 41:899-913 (2002). https://doi.org/10.1021/ie001047w
3. https://github.com/yueyue2299/TeNNet-SAC
4. https://huggingface.co/spaces/stlin/tennetsac