Comprehensive Report on ANN Modeling of Vapor-Liquid Equilibrium (VLE) System

Introduction

Vapor-Liquid Equilibrium (VLE) plays a critical role in the design of chemical separation processes such as distillation, extraction, and azeotropic separations. Classical models, such as Raoult ' s Law and activity coefficient-based approaches (Wilson, NRTL, UNIQUAC), provide approximate predictions but often fail in highly non-ideal mixtures. The ethanol-water system is a classic example, forming an azeotrope that cannot be explained using ideal models alone.

In this study, Artificial Neural Networks (ANNs) were applied to learn VLE behavior directly from synthetic data. The ANN provides a data-driven model capable of capturing nonlinear relationships and predicting azeotropic behavior. This report outlines the dataset, neural network design, training methodology, and results, highlighting the strengths and limitations of the ANN approach in thermodynamic modeling.

1. Dataset Description

The dataset was synthetically generated to mimic experimental measurements of the ethanol-water binary system. The ethanol mole fraction in the liquid phase (x_1) was varied between 0.01 and 0.99. To ensure proper representation of azeotropic behavior, denser sampling was introduced near $x_1 \approx 0.90$, which corresponds to the known azeotropic composition.

The vapor mole fraction (y_1) was simulated using a nonlinear function that introduces non-ideal deviations from Raoult ' s law. Additionally, the system temperature (T) was generated with sinusoidal perturbations and Gaussian noise to represent experimental variability. The total pressure was kept constant at 101.325 kPa. The final dataset contained over 450 data points.

Input variables: liquid mole fraction (x_1), temperature (T), pressure (P).
Output variable: vapor mole fraction (y_1).

This dataset provides the ANN with sufficient diversity to learn azeotropic trends and temperature dependencies.

2. Artificial Neural Network (ANN) Design
The ANN was developed using TensorFlow/Keras. The chosen architecture balances complexity and computational efficiency, consisting of:

- Input layer with three neurons (x_1, T, P).
- Two hidden layers, each containing 64 neurons with ReLU activation functions. ReLU was chosen for its efficiency in training deep networks and avoiding vanishing gradient problems.
- Output layer with one neuron and sigmoid activation, restricting predictions to [0, 1], which is physically meaningful for mole fractions.

The Adam optimizer was selected due to its adaptive learning rate capabilities, allowing efficient convergence. Mean Squared Error (MSE) was used as the loss function to minimize deviations between predicted and experimental y_1 values.

Data Preprocessing: Input and output data were normalized using Min-Max scaling to improve convergence and ensure balanced feature contributions during training.

3. Training Methodology
The dataset was divided into training (80%) and testing (20%) subsets using a randomized split to avoid bias. During training, 10% of the training data was further reserved for validation to monitor overfitting.

Training parameters:
- Epochs: 200
- Batch size: 32
- Validation split: 0.1

The training process showed steady reduction in both training and validation losses, demonstrating that the model generalized well. No signs of overfitting were observed, as validation performance closely followed training performance throughout the epochs.

4. Results and Analysis

The ANN ' s predictive capability was evaluated on the test dataset. A parity plot comparing predicted and actual vapor mole fractions showed strong agreement, with most points lying close to the ideal diagonal line.

Azeotrope Prediction:

One of the critical tests for the ANN was the ability to detect azeotropy, where $y_1 = x_1$. By solving this condition using the trained model, the ANN predicted an azeotropic composition near $x_1 \approx 0.89$, which is close to the experimentally known azeotrope of ethanol-water (~0.90).

The predicted azeotropic temperature was around 350 K, consistent with literature values. In contrast, Raoult ' s Law was unable to capture azeotropy, reinforcing the advantage of ANN-based modeling.

Comparison with Traditional Models:

- Raoult ' s Law: fails to capture azeotropy due to lack of activity coefficient consideration.
- ANN: successfully modeled non-linear thermodynamic behavior from data without explicit equations.

Performance Metrics:

While exact metrics such as R^2 and RMSE were not the primary focus, visual analysis of parity plots and prediction errors confirmed high accuracy of the ANN model.

5. Conclusion

This project demonstrated the ability of Artificial Neural Networks to model complex thermodynamic phenomena. By training on synthetic VLE data, the ANN successfully learned the azeotropic behavior of ethanol-water systems, which traditional ideal models fail to capture.

The findings highlight the potential of machine learning in chemical engineering, where ANN-based models can complement or replace classical models, particularly in highly non-ideal systems. Future improvements could include training on experimental datasets, testing different ANN architectures, and extending the approach to multicomponent systems.

In summary, ANN-based thermodynamic modeling offers a robust, data-driven alternative to traditional methods, with demonstrated success in capturing non-linear, azeotropic behavior in binary mixtures.