

---

## LIST OF FIGURES

1.1	The conceptual framework of this study . . . . .	5
2.1	The illustration of multi-head attention and the scaled dot-product attention in Vaswani et al. study . . . . .	15
3.1	System Flow to Develop a Sequence-based DTA Prediction Model . . . . .	17
3.2	Sample data from Davis and KIBA datasets . . . . .	18
3.3	The distribution of Davis and KIBA datasets . . . . .	19
3.4	Illustration of DTA prediction model framework . . . . .	21
3.5	Example of modified scale process . . . . .	23
4.1	Scatter plot of comparison between actual and predicted values for samples with affinity > threshold in Davis and KIBA datasets . . . . .	34
4.2	Histogram of error distribution in high-affinity data for Davis and KIBA . .	34
4.3	The training and validation loss curves of the best model on both datasets, Davis (left) and KIBA (right) . . . . .	35
4.4	The evaluation results of the comparison of the proposed drug and receptor representations with other representations in predicting DTA (the proposed representation method is marked with *) . . . . .	36
4.5	The evaluation results of the interaction modeling experiment on the Davis and KIBA dataset . . . . .	37