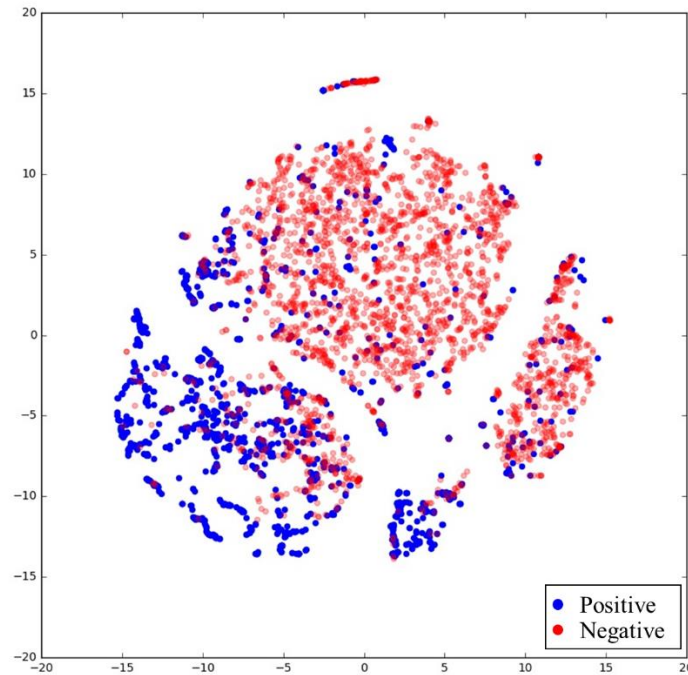
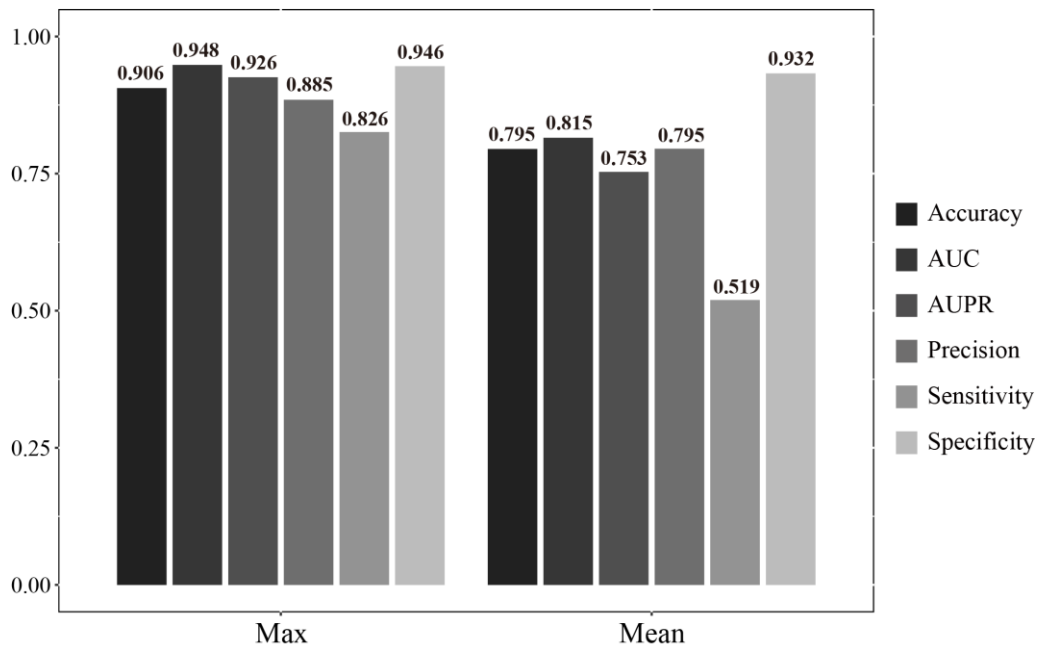


Figure S1 – Data distribution of training dataset



Blue points represent the positive class and red points represent the negative. The distribution was plotted using t-SNE algorithm.

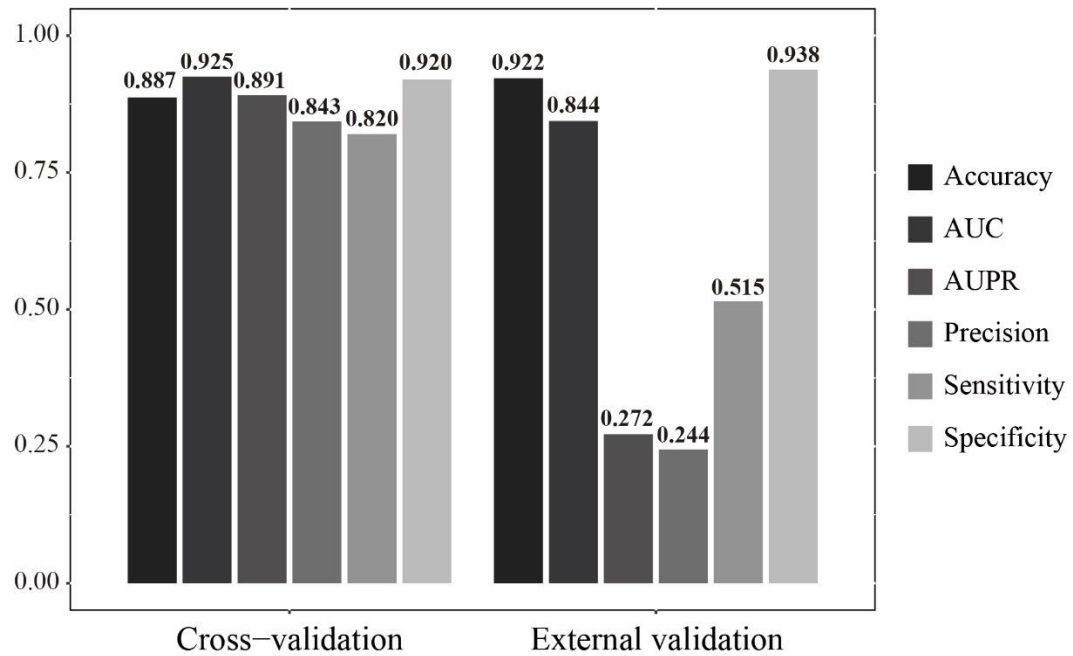
Figure S2 – Performance comparison in cross-validation using different features



Performances of prediction models constructed with features obtained by each maximum and mean values. Models were trained with random forest algorithm under same condition.

Prediction model with features from maximum values outperformed and applied to further evaluation.

Figure S3 – Prediction ability for drug repositioning



Prediction model performance trained with only drugs in the test set. Constructed model showed similar performance compared to models with entire dataset.