# In silico prediction of drug-induced liver injury of drug candidates

Jingyu Lee, Myeong-sang Yu, and Dokyun Na*

*School of Integrative Engineering, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, seoul 06974, Republic of Korea*

## Introduction

Because over 45,000 cases of drug-induced liver injury (DILI) with reported causal relationships have been reported and DILI accounts for almost 20% of drug withdrawal in the market and, it is essential to consider in the early stages of drug development. Thus, development for the DILI prediction model was in high demand for successful drug discovery. To date, various research has been conducted to build an accurate DILI prediction model. However, it is still a huge challenge to develop a high-performance prediction model suitable to actual use in drug discovery.

The performance of our ensemble model was higher than of other DILI prediction models. Since many published DILI prediction models are not available for public access, we developed our optimal prediction model publicly available (http://ssbio.cau.ac.kr/software/dili).

## Results

**Table 1. 10-fold cross-validation with 10 repetitions results of the prediction models trained by different machine learning algorithms and ensemble models.**

| | Algorithm | AUC | Accuracy | Sensitivity | Specificity | F1-score |
|---|---|---|---|---|---|---|
| Individual models | Adaboost (Ada) | 0.815 ± 0.009 | 0.766 ± 0.009 | 0.790 ± 0.013 | 0.741 ± 0.012 | 0.781 ± 0.009 |
| | **Categorical gradient boosting (CBT)** | **0.875 ± 0.007** | **0.794 ± 0.009** | **0.817 ± 0.011** | **0.768 ± 0.013** | **0.807 ± 0.009** |
| | Gaussian naïve bayes (GNB) | 0.697 ± 0.008 | 0.580 ± 0.011 | 0.300 ± 0.026 | 0.894 ± 0.010 | 0.424 ± 0.027 |
| | Gaussian process (GP) | 0.862 ± 0.007 | 0.781± 0.009 | 0.818 ± 0.012 | 0.740 ± 0.014 | 0.797 ± 0.009 |
| | k-Nearest neighbor (kNN) | 0.859 ± 0.008 | 0.786 ± 0.009 | 0.770 ± 0.014 | 0.804 ± 0.013 | 0.791 ± 0.010 |
| | Linear discriminant analysis (LDA) | 0.712 ± 0.012 | 0.720 ± 0.010 | 0.727 ± 0.014 | 0.714 ± 0.017 | 0.732 ± 0.010 |
| | **Light gradient boosting (LGB)** | **0.874 ± 0.007** | **0.795 ± 0.009** | **0.818 ± 0.012** | **0.769 ± 0.009** | **0.807 ± 0.009** |
| | Logistic regression (LR) | 0.789 ± 0.010 | 0.733 ± 0.010 | 0.754 ± 0.014 | 0.710 ± 0.010 | 0.748 ± 0.010 |
| | Multilayer perceptron (MLP) | 0.826 ± 0.010 | 0.775 ± 0.010 | 0.798 ± 0.014 | 0.749 ± 0.015 | 0.788 ± 0.010 |
| | Random forest (RF) | **0.879 ± 0.007** | **0.800 ± 0.009** | **0.826 ± 0.011** | **0.771 ± 0.014** | **0.813 ± 0.009** |
| | Support vector machine (SVM) | 0.856 ± 0.005 | 0.693 ± 0.007 | 0.961 ± 0.003 | 0.394 ± 0.014 | 0.768 ± 0.004 |
| | eXtreme gradient boosting (XGB) | 0.857 ± 0.007 | 0.782 ± 0.008 | 0.802 ± 0.012 | 0.760 ± 0.012 | 0.795 ± 0.009 |
| Ensemble models | Voting (RF+LGB) | 0.878 ± 0.007 | 0.798 ± 0.009 | 0.822 ± 0.012 | 0.771 ± 0.013 | 0.811 ± 0.009 |
| | Voting (RF+LGB+CBT) | 0.878 ± 0.007 | 0.797 ± 0.009 | 0.821 ± 0.012 | 0.771 ± 0.012 | 0.810 ± 0.004 |
| | Stacking (RF+LGB) | 0.878 ± 0.005 | 0.800 ± 0.009 | 0.826± 0.012 | 0.770 ± 0.012 | 0.813 ± 0.009 |
| | Stacking (RF+LGB+CBT) | **0.878 ± 0.007** | **0.800 ± 0.009** | **0.827 ± 0.010** | **0.770 ± 0.009** | **0.824 ± 0.009** |

NOTE. Data are presented as average ± standard error of mean

**Table 2. Evaluation results of the selected individual models and developed ensemble models with the independent dataset. (367 compounds)**

| Algoritnm | Accuracy | Sensitivity | Specificity | F1-score |
|---|---|---|---|---|
| Categorical gradient boosting (CBT) | 0.77 | 0.81 | 0.68 | 0.83 |
| Light gradient boosting (LGB) | 0.79 | 0.83 | 0.67 | 0.85 |
| Random forest (RF) | 0.79 | 0.83 | 0.70 | 0.85 |
| Voting method (RF + LGB) | 0.79 | 0.85 | 0.70 | 0.85 |
| Voting method (RF + LGB + CBT) | 0.79 | 0.85 | 0.64 | 0.85 |
| Stacking method (RF + LGB) | 0.80 | 0.85 | 0.68 | 0.85 |
| Stacking method (RF + LGB + CBT) | 0.79 | 0.84 | 0.66 | 0.85 |

**Table 3. Evaluation results of the selected individual models and developed ensemble models with the test dataset for model comparision. (96 compounds)**

| Algoritnm | Accuracy | Sensitivity | Specificity | F1-score |
|---|---|---|---|---|
| Categorical gradient boosting (CBT) | 0.67 | 0.80 | 0.52 | 0.71 |
| Light gradient boosting (LGB) | 0.73 | 0.76 | 0.70 | 0.75 |
| Random forest (RF) | 0.75 | 0.82 | 0.68 | 0.77 |
| Voting method (RF + LGB) | 0.72 | 0.78 | 0.65 | 0.74 |
| Voting method (RF + LGB + CBT) | 0.71 | 0.80 | 0.61 | 0.74 |
| Stacking method (RF + LGB) | 0.75 | 0.82 | 0.67 | 0.77 |
| Stacking method (RF + LGB + CBT) | 0.77 | 0.82 | 0.72 | 0.79 |

**Table 4. Prediction results comparison with other DILI prediction models.**

| Prediction model | Accuracy | Sensitivity | Specificity | F1-score |
|---|---|---|---|---|
| **Our ensemble model** | **0.77** | **0.82** | **0.72** | **0.79** |
| Our RF model | 0.75 | 0.82 | 0.67 | 0.77 |
| DILI_MOE_RF100 | 0.59 | 0.54 | 0.67 | - |
| DILI_MOE_transp_RF100 | 0.63 | 0.56 | 0.70 | - |
| DILI_RDKiit_RF100 | 0.64 | 0.64 | 0.64 | - |
| DL_Liew | 0.55 | 0.62 | 0.47 | 0.59 |
| DL_Combined | 0.57 | 0.54 | 0.60 | 0.57 |
| ProTox - II | 0.64 | 0.32 | 1.00 | 0.48 |
| ADMETlab 2.0 | 0.72 | 0.74 | 0.69 | 0.73 |

**Figure 1. ROC curves for optimal individual models and the ensemble model. The ROC curves of the high-performing individual models are shown RF (A), LGB (B), and CBT (C). The ROC curves of the ensemble model (D) is shown.**

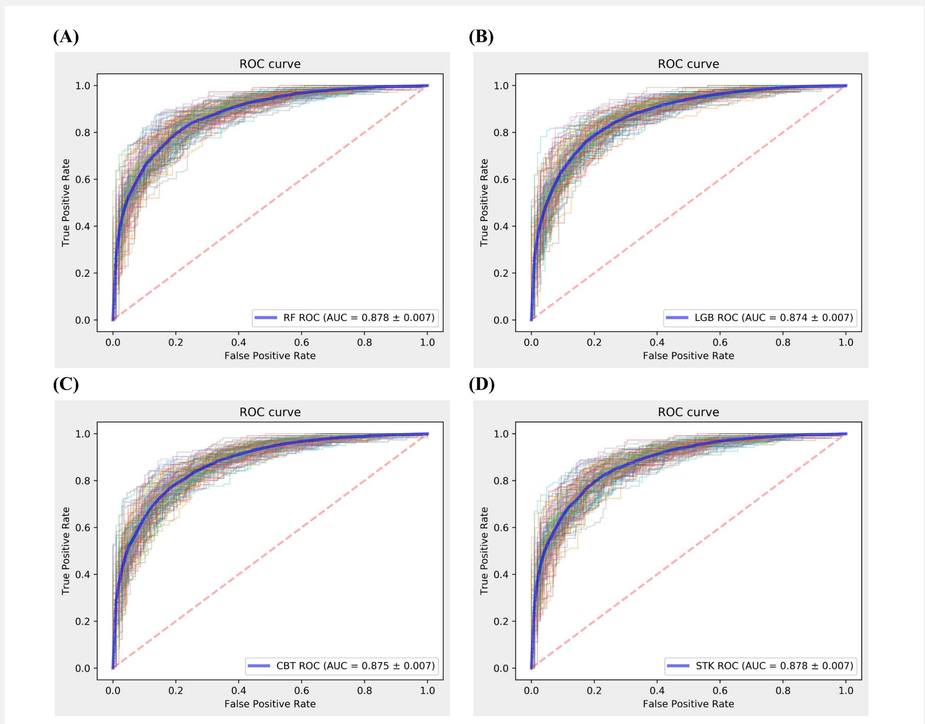



**Figure 2. Prediction results for each chemical compound are shown graphically. (96 compounds)**

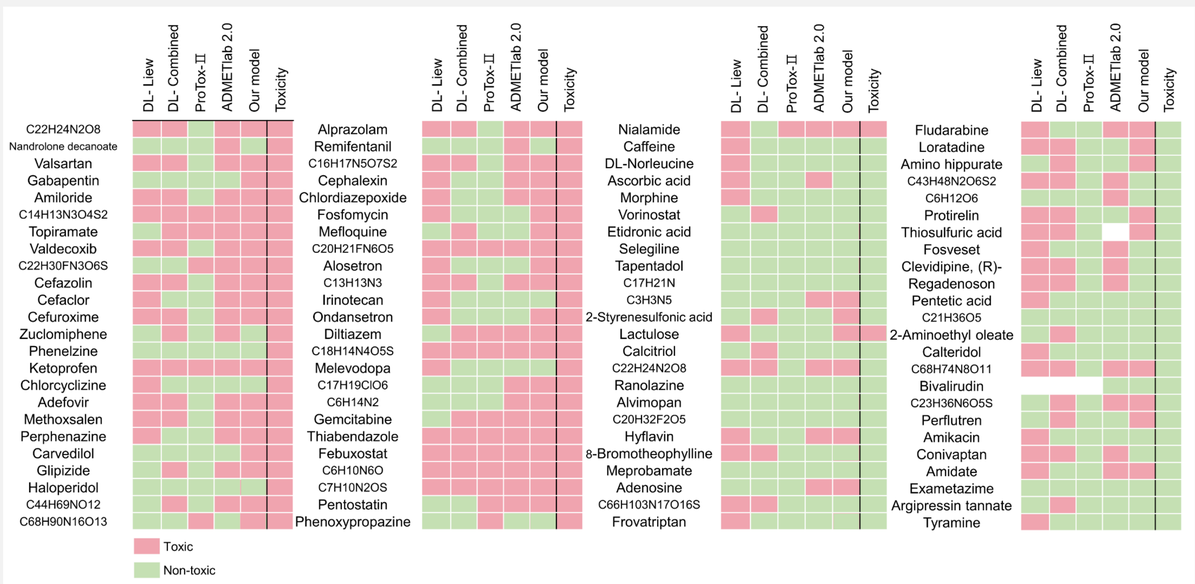
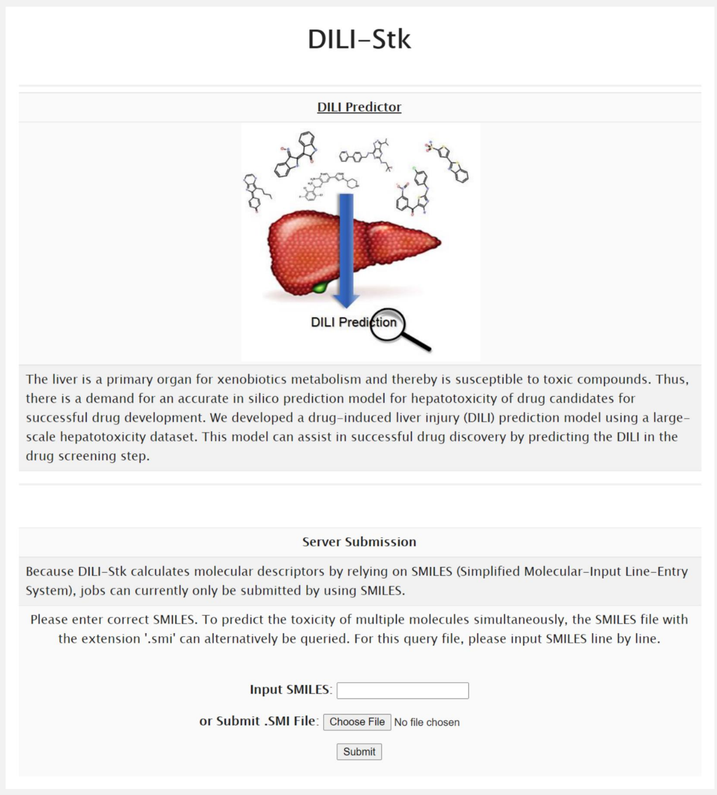



**Figure 3. DILI-Stk web server. Our web server is accessible via conventional browsers, and users are able to predict the DILI of a query compound expressed in SMILES format.**




## Conclusion

- We developed ensemble model to prediction DILI
- Our prediction model showed better performance than other previous DILI prediction models.
- We expect our DILI prediction model to have potential applications in drug discovery by predicting cases of DILI and reducing drug withdrawal rates.
- We made our optimal prediction model publicly available for practical use for drug discovery (http://ssbio.cau.ca.kr/software/dili).