Abstract

Over time, many individuals have been exposed to chemical substances with potentially harmful effects on the human body, making drug toxicity a critical factor in drug development process. High toxicity remains a primary cause of drug failure during clinical trials. Therefore, toxicity testing has become a main focus in the medical field to prevent further exposure to hazardous chemicals. The study focuses on predicting the toxicity of androgen receptor ligand-binding domain (AR-LBD) compounds by implementing Long Short-Term Memory (LSTM) model optimized by Simulated Annealing (SA). The methodology includes several steps, such as dataset preparation from the Tox21 Data Challenge, model training using the SA-optimized LSTM, performance evaluation against traditional toxicity prediction methods, and validation through testing datasets. The proposed model demonstrated impressive results, achieving an F1-score of 0.7105 and an accuracy 0.9782 outperforming traditional prediction models and the baseline without SA. These results support the effectiveness of SA for optimization, improving predictive capabilities and applicabilities in real-world scenarios. Future research should explore additional hyperparameter tuning and alternative optimization techniques to further improve prediction accuracy and applicability across diverse datasets and various scenarios in toxicology.

Keywords: Toxicity Prediction, LSTM, Simulated Annealing, AR-LBD