1. Introduction

In the medical field, a drug refers to a chemical substance that can be used to prevent, diagnose, treat, ease disease symptoms, and restore physiological functions in living organisms [1, 2, 3]. Over time, many individuals have been affected by chemical substances that pose potential harm to the human body. Drug toxicity is a critical aspect of drug development, along with the identification of adverse effects of substances on humans, plants, animals, and the environment. However, high toxicity remains a major factor contributing to drug failure in clinical trials [4]. Therefore, toxicity testing is an essential consideration in the medical field to minimize exposure to harmful chemical substances.

The increasing rates of drug withdrawals, rising costs, late-stage failures, and significant recalls represent some of the major challenges in drug toxicity testing. These issues are critical in pharmaceutical research and development (R&D). Data from CMR International demonstrated that between 2006 and 2010, approximately 22% of clinical development failures and 54% of preclinical development failures were attributed to toxicity [5]. Therefore, in addition to minimizing exposure to harmful chemical substances, toxicity testing is a critical step in the early phases of drug research and development. This approach aims to prevent toxic drugs from progressing to clinical trials while also reducing the costs and labor associated with preclinical drug testing [6].

Traditionally, toxicity testing or measurement has been conducted through in vivo methods, where tests are performed directly on living organisms to observe their physiological reactions to the tested substances [7]. This approach is commonly used for toxicity assessments, with animals serving as the primary test subjects. However, this method is highly expensive and challenging to implement, requiring two to three years per chemical substance for comprehensive toxicity evaluation [8]. Additionally, traditional toxicity prediction methods often struggle with accurately classifying compounds, particularly the compounds with similar structures, leading to high rates of false positives and negatives. Moreover, the results from conventional in vivo testing offer limited insights into toxicity reactions in the human body due to interspecies differences and the diverse nature of disease models [6].

In response to this, the utilization of computational or in silico toxicity testing is recommended as it can reduce costs and uncertainties associated with animal testing [6], [8]. The in silico method offers several advantages, including cost reduction, greater time efficiency, and a unique approach to visualize toxicity testing results through the virtual structure of the tested compounds. In toxicology, the application of machine learning (ML) approaches has become increasingly widespread. These approaches enable the evaluation of large numbers of compounds in a time-efficient manner without involving animals as test subjects, in contrast to traditional in vivo methods [9].

Several researchers have conducted studies using machine learninig (ML) to predict the toxicity of various compounds. In 2020, Elizeta Semenova and colleagues developed a Bayesian neural network model to predict the toxicity of compounds that have potential to cause liver injury. Their findings demonstrated that neural network methods are flexible and can reduce uncertainty when predicting compound toxicity [10]. In 2019, Abdul Karim and colleagues conducted research on toxicity prediction using multimodal deep learning. This method utilized multiple types of homogeneous neural networks and diverse data representations. Their proposed method achieved significantly higher accuracy, ranging from 0.84 to 0.88, on the IGC50 benchmark dataset [11].

In 2021, Abdul Karim and his colleagues conducted a study that developed a deep learning ensemble framework to quantitatively predict compound toxicity levels. This research achieved significant improvements in quantitative toxicity predictions by integrating four quantitative toxicity datasets: LD50, IGC50, LC50, and LC50-DM. Among these datasets, the method used demonstrated enhanced performance, achieving improvements of 5.46%, 16.67%, and 6.34% in root-mean-square errors; 6.41%, 11.80%, and 12.16% in mean absolute errors; and 5.21%, 7.36%, and 2.54% in coefficients of determination for three out of the four datasets [12]. Literature reviews indicate that the development of such methods has delivered promising results, yet the models may not generalize well accross diverse datasets or chemical classes, limiting the applicability in real-world scenarios. Therefore, the use of Simulated Annealing (SA) algorithms combined with Long Short-Term Memory (LSTM) networks in this study aims to improve toxicity predictions while optimizing computational efficiency and applicability of toxicity predictions in real-world scenarios.

This study focuses on predicting the toxicity of androgen receptor ligand-binding domain (AR-LBD) compounds using the LSTM model optimized with SA. LSTM is a specialized recurrent neural network (RNN) with a unique design structure. It features self-looping mechanisms that enable precise and consistent control over extended time periods [13]. In this case, SA is applied as an optimization technique to identify the optimal parameters for the LSTM model. Simulated Annealing excels at identifying the best solutions while reducing uncertainties in toxicity prediction. The integration of these methods aims to deliver improved prediction accuracy and greater computational efficiency.