**Title:** Toxicity Prediction of Drug Compounds using Deep Learning  
**Author:** Mihir Kumar Panigrahi

**1. Introduction:**

The integration of deep learning in pharmaceutical sciences has revolutionized drug discovery and toxicity prediction. In this paper, I present a simple deep learning-based approach to classify drug compounds as toxic or non-toxic using simulated data to demonstrate model structure and logic.

**2. Objective:**

To design a binary classification model that predicts the toxicity of compounds using molecular-level features. The model simulates a typical deep learning pipeline for AI-powered drug development.

**3. Dataset:**

Due to time and computational constraints, a simulated dataset was used that mimics the structure of real molecular descriptors (10 features, 1000 samples). This allowed us to focus on building the model pipeline with the intent to integrate real datasets like Tox21 in future iterations.

**4. Methodology:**

* Data preprocessing using StandardScaler
* Binary classification using a deep neural network (Keras Sequential API)
* **Architecture:**
  + Input Layer: 10 features
  + Dense Layer: 64 neurons, ReLU
  + Dropout Layer: 30%
  + Dense Layer: 32 neurons, ReLU
  + Dropout Layer: 20%
  + Output Layer: 1 neuron, Sigmoid activation
* Binary Cross-Entropy Loss
* Adam optimizer

**5. Novelty:**

The novelty in this approach lies in the use of dropout-based regularization to handle potential overfitting, inspired by patterns seen in reviewed academic papers. A custom toxicity prediction function was also introduced to simulate real-life applications.

**6. Evaluation:**

* Accuracy on test set: ~84–88%
* Metric: Accuracy Score
* Model performs binary prediction with relatively high performance on unseen data.

**7. Future Work:**

In future versions, the project will be extended to:

* Use real datasets like **Tox21** or **ChEMBL**
* Process chemical structures using **SMILES** strings via **DeepChem** or **RDKit**
* Explore **Graph Neural Networks**
* Include advanced dashboard visualizations using Plotly or Streamlit

**8. Conclusion:**

This paper presents a deep learning approach to drug toxicity prediction using simulated data. It lays the foundation for future models using real molecular data and more complex techniques in AI-driven drug discovery.

s