# CHAPTER – 2
# LITERATURE SURVEY

# CHAPTER 2
# LITERATURE SURVEY

## 2.1 Existing System

**Katyna Sada Del Reala and Angel Rubio, "Discovering the Mechanism of Action of Drugs with a Sparse Explainable Network", 2023 [1].**

This study presents SparseGO, an interpretable neural network designed for predicting drug responses in cancer cell lines, focusing on understanding mechanisms of action (MoA). SparseGO leverages gene expression data and applies DeepLIFT, an explainable AI method, to achieve interpretable predictions by linking specific input features to model outcomes. SparseGO improves memory efficiency and optimizes resource usage, reducing reliance on high-powered GPUs. However, SparseGO's dependence on gene expression data alone limits its adaptability, and its results require further clinical validation. Our project expands on these limitations by incorporating multiple genomic features, including gene mutations and CNAs, to improve predictive accuracy. Additionally, our model includes a Flask-based interface for practical accessibility in clinical and research environments, promoting broader applicability without heavy computational requirements.

**Alexander Partin et al., "Deep Learning Methods for Drug Response Prediction in Cancer: Predominant and Emerging Trends," 2023 [2].**

This review provides an extensive overview of 61 deep learning models applied to drug response prediction, with an emphasis on architectures such as convolutional and graph neural networks (GNNs). It highlights the importance of omics data integration and discusses the increasing trend toward personalized treatment and drug repurposing. While many reviewed models achieve promising results, they often struggle with generalization across new drugs and rely on large, annotated datasets for training, limiting real-world application. The lack of standardized evaluation frameworks also complicates performance comparisons. In our project, we address these limitations by focusing on an ANN model that integrates genomic features such as mutations and CNAs to enhance prediction accuracy, complemented by a simple, accessible Flask interface for use by researchers and clinicians.

**Bara A. Badwan, Mohammad Qasem, Mohammad Anan, and Ali Hamed El-Moussawi, "Machine Learning Approaches to Predict Drug Efficacy and Toxicity in Oncology," 2023 [3].**

This paper explores various machine learning algorithms in oncology, with a focus on predicting drug efficacy and toxicity. Techniques like PCA and t-SNE for dimensionality reduction are emphasized, particularly when handling high-dimensional genomic data. Despite its success, the study points out challenges such as reliance on extensive datasets and the gap between IC50 predictions and clinical efficacy. Current models also face limitations in generalizing predictions due to data variability. Our project addresses these limitations by optimizing feature selection and employing an ANN model trained on selected genomic data. The inclusion of a Flask-based interface allows for real-time predictions, which can assist researchers and clinicians in making informed decisions without requiring substantial computational resources.

**Brent M. Kuenzi, Jisoo Park, Samson H. Fong, Kyle S. Sanchez, John Lee, Jason F. Kreisberg, Jianzhu Ma, and Trey Ideker, "Predicting Drug Response and Synergy Using a Deep Learning Model of Human Cancer Cells," 2020 [4].**

DrugCell, introduced by Kuenzi et al., is a deep learning model designed to predict cancer cell responses to therapies by integrating genetic and chemical data. It employs a Visible Neural Network (VNN) architecture that hierarchically maps biological subsystems. DrugCell is effective in predicting single-drug and combination responses but is constrained by its reliance on predefined biological hierarchies, which may limit its adaptability to novel molecular interactions. Furthermore, DrugCell's computational requirements are high, which may limit its accessibility in some clinical settings. Our project aims to address these limitations by utilizing a streamlined ANN model that integrates selected genomic features without reliance on hierarchical dependencies, and by offering a Flask-based interface that makes our model accessible and practical for both research and clinical applications.

## 2.2 Limitations of Existing System

**Katyna Sada Del Reala and Angel Rubio, "Discovering the Mechanism of Action of Drugs with a Sparse Explainable Network", 2023 [1].**

SparseGO improves computational efficiency but is restricted by its focus on gene expression data, limiting its capacity to incorporate broader genomic features such as mutations and CNAs. Additionally, SparseGO's MoA insights require experimental validation before clinical use. Our project builds on this by including a wider range of genomic data within an ANN framework and providing a user-friendly interface for enhanced accessibility.

**Alexander Partin, Thomas S. Brettin, Yitan Zhu, Oleksandr Narykov, Austin Clyde, Jamie Overbeek, and Rick L. Stevens, "Deep Learning Methods for Drug Response Prediction in Cancer: Predominant and Emerging Trends," 2023 [1].**

Reviewed models often require extensive annotated data and face generalization challenges with new drugs. The absence of standardized evaluation frameworks restricts the comparability of model performance. Our project addresses these limitations by focusing on key genomic features, like mutations and CNAs, to enhance robustness, complemented by a Flask-based interface that ensures ease of access for researchers and clinicians.

**Bara A. Badwan, Mohammad Qasem, Mohammad Anan, and Ali Hamed El-Moussawi, "Machine Learning Approaches to Predict Drug Efficacy and Toxicity in Oncology," 2023 [3].**

This study highlights the reliance on large datasets and the limited generalizability of current models. Additionally, inconsistencies in evaluation standards complicate the comparison of models across studies. Our approach uses feature selection to enhance predictive accuracy, and our web-based interface allows real-time access, simplifying use for clinical and research applications.

**Brent M. Kuenzi, Jisoo Park, Samson H. Fong, Kyle S. Sanchez, John Lee, Jason F. Kreisberg, Jianzhu Ma, and Trey Ideker, "Predicting Drug Response and Synergy Using a Deep Learning Model of Human Cancer Cells," 2020 [4].**

ADrugCell's reliance on pre-defined biological hierarchies limits its adaptability, and its computational demands restrict scalability. Our project circumvents these limitations by focusing on an ANN model that integrates genomic data without complex hierarchies and provides a Flask interface suitable for various research and clinical settings.

## 2.3 Proposed System

This project aims to advance precision medicine in oncology by developing an Artificial Neural Network (ANN) model to predict drug response, specifically the IC50 values that indicate drug efficacy in inhibiting cancer cell growth. Traditional methods often struggle to capture the complex and nonlinear relationships between genomic features and drug efficacy, leading to less effective treatment options. Leveraging a merged dataset that combines the Genomics of Drug Sensitivity in Cancer (GDSC) data and other relevant sources, this model incorporates critical genomic features, such as gene mutations, gene expression profiles, and copy number alterations (CNAs), to improve prediction accuracy and support tailored cancer treatments.

The ANN model architecture features multiple hidden layers with ReLU activation functions, which enable it to capture intricate relationships within the genomic data. Optimized using the Adam optimizer and tuned hyperparameters, this model minimizes Mean Squared Error (MSE) to achieve high prediction accuracy. Early stopping techniques prevent overfitting, ensuring the model's predictions are generalizable and reliable across various cancer cell profiles. Key performance metrics, such as R-squared ($R^2$) and Root Mean Squared Error (RMSE), validate the model's effectiveness in accurately predicting drug response.

A user-friendly, Flask-based web interface allows users to input genomic data, select specific cancer types, and receive IC50 predictions with real-time visualization. The interface provides dropdown menus for easy data entry, eliminating the need for file uploads and making the platform accessible for both clinical and research purposes. Results are displayed in a comparative graph format that shows IC50 predictions for three different drugs, allowing users to make informed decisions about the most effective treatment options.

This project includes a comprehensive data preprocessing pipeline, which covers normalization, feature scaling, and careful feature selection. These preprocessing steps ensure the model effectively captures the most relevant information from the merged datasets, enhancing prediction accuracy. By addressing the limitations of existing models, such as limited input features and accessibility issues, this system provides a scalable solution for personalized cancer treatment. The framework is designed to accommodate future expansions with additional genomic datasets, underscoring its potential in advancing precision oncology by minimizing trial-and-error in therapeutic decisions.