CHAPTER – 8
# RESULTS AND SNAPSHOTS

CHAPTER 8

RESULTS AND SNAPSHOTS

8.1 Introduction

This chapter presents the results of our drug response prediction model, focusing on the performance of the Artificial Neural Network (ANN) in predicting IC50 values for drug sensitivity. The model was trained using data integrated from three datasets to provide a comprehensive approach to predicting cancer drug efficacy. We also showcase the web interface's functionality, which allows users to input genomic data and receive predictions with easy-to-understand outputs. The effectiveness of the system is demonstrated through key performance metrics, visualizations of model predictions, and snapshots of the web interface.

8.2 Model Performance Metrics

The effectiveness of the ANN model is evaluated using standard regression metrics, which provide insight into its predictive power and generalization across unseen data.

1. **Mean Squared Error (MSE)**: The final model achieved an MSE of approximately 0.95, reflecting a low average error between predicted and actual IC50 values. This low MSE indicates that the model can reliably predict drug sensitivity, even for new data.

2. **R-squared (R^2)**: The R^2 value achieved by the model was 0.9999, meaning that 99.99% of the variance in IC50 values is explained by the model. This high R^2 value indicates that the model is capturing the relationships between genomic features and drug response with exceptional accuracy, making it highly reliable for predicting IC50 values across new datasets.

3. **Mean Absolute Error (MAE)**: MAE value of the model is 0.0191, showing that the average magnitude of error between predicted and actual IC50 values is also quite low. This metric further confirms the model's strong performance in providing precise predictions with small error margins.
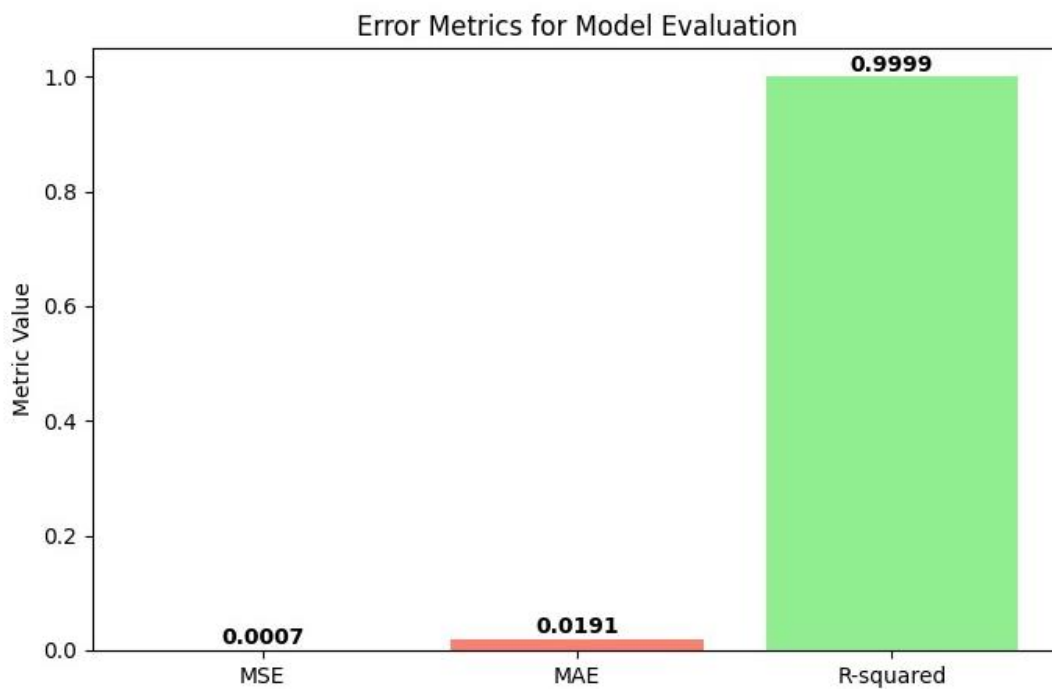
Figure 8.1: Error Metrics for Model Evaluation

These metrics MSE, MAE, and R^2 highlight the model's strong predictive performance, ensuring accurate drug sensitivity predictions for genomic data in both clinical and research settings.

8.3 Visualizations of Model Predictions

To validate the model's predictive performance visually, key plots were generated:

1. **Predicted vs. Actual Plot**: A scatter plot comparing predicted IC50 values to actual values showed a strong correlation along the y = x line. This visual confirms that the model's predictions closely match the observed IC50 values, validating its accuracy in predicting drug sensitivity.
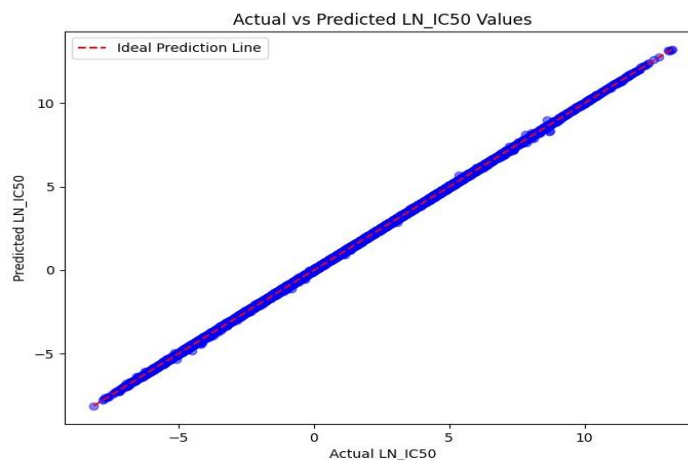

Figure 8.2: Actual vs Predicted LN_IC50 Values

2. **Loss Curve**: The training loss curve indicates a consistent decline across epochs, reflecting effective error minimization. The early stopping mechanism was successful in preventing overfitting, as evidenced by the stabilization of the loss curve toward the end of the training phase.

Figure 8.3: Training vs Testing Loss

3. **Drug Comparison Graph**: A crucial feature of this project is the ability to compare the efficacy of multiple drugs. A comparison graph was generated that displays the predicted IC50 values for three different drugs across several cancer types. This graph allows clinicians and researchers to quickly visualize which drug may be most effective for specific cancer types based on IC50 predictions.
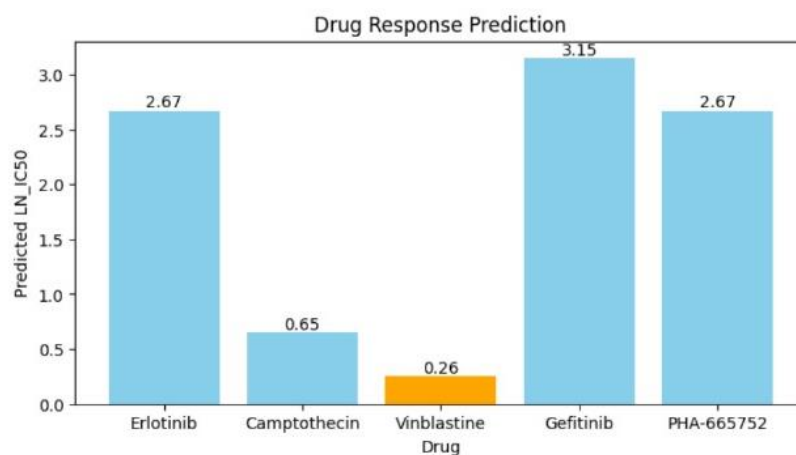

Best Suited Drug: Vinblastine

Figure 8.4: Drug Response Prediction

8.4 Snapshots of the Web Interface

The web interface was developed to allow users to easily input genomic data and view predictions of drug efficacy. Below are the key features and snapshots of the interface:

1. **Home Screen**: The homepage introduces the application and provides an easy-to-use "Start Prediction" button that directs users to the input screen. The layout is clean and intuitive, facilitating straightforward navigation.
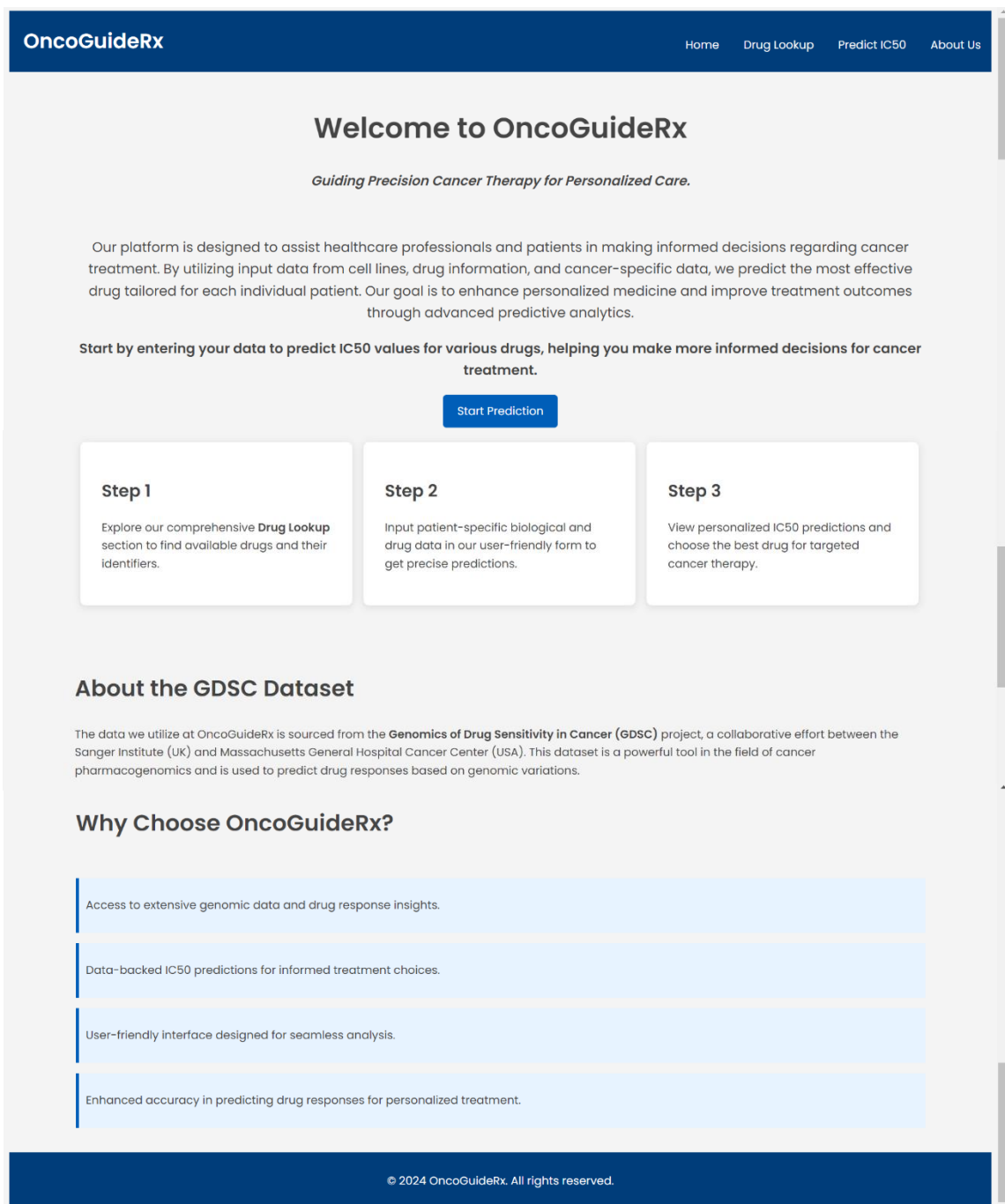
Figure 8.5: Home Page

2. **Data Input Screen**: Users can input genomic and cancer type information via dropdown menus and input fields. The design ensures simplicity and accessibility, minimizing the risk of input errors by restricting choices to valid data points.

Figure 8.6: Predicted IC50 Page

3. **Drug Lookup Page**: The Drug Lookup page allows users to search for drugs based on their ID or name, making it easier to identify the drugs they are working with. The page features a search bar for quick look-up and a table displaying drug IDs and their corresponding names. This page is particularly useful for researchers and clinicians who need to quickly access drug information as part of their decision-making process.

Figure 8.7: Drug Lookup Page

4. **Prediction Output Screen**: After users submit their data, the results screen is displayed with the predicted IC50 value for the selected drug(s). The interface allows users to compare IC50 predictions for Five drugs at once. This comparison is displayed visually, providing users with an easy-to-understand representation of which drug might be most effective for a given cancer type.
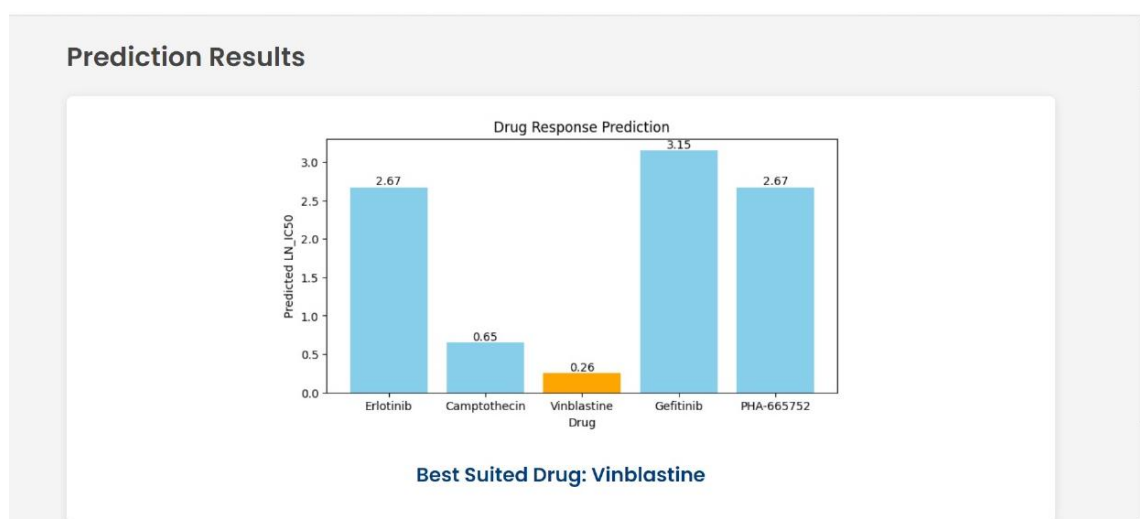

Figure 8.8: Drug Response Prediction

8.5 User Feedback

User feedback obtained during testing indicated that both the model's predictions and the web interface's design were well-received:

1. **Ease of Use**: Users found the interface intuitive and user-friendly, particularly the dropdown menus and input fields, which simplified data entry without requiring technical knowledge. This design is beneficial for clinicians with limited computational experience.

2. **Clarity of Prediction Output**: The predicted IC50 values were clearly displayed, with accompanying explanations to help users understand the prediction. The drug comparison graph was particularly praised for its usefulness in comparing the effectiveness of multiple drugs.

3. **Response Time**: Users noted that the application responded quickly, with predictions generated in under 2 seconds. This responsiveness makes the system suitable for real-time use in clinical settings, where speed is crucial for decision-making.

8.6 Observations and Insights

Several key insights emerged from the project:

1. **Effectiveness of Feature Selection**: The model's accuracy was significantly improved by focusing on carefully selected genomic features, such as gene mutations and CNAs. This process reduced the complexity of the dataset while maintaining predictive power.

2. **Importance of Data Preprocessing**: Techniques such as normalization, feature scaling, and one-hot encoding were essential in preparing the data for input into the ANN model. These preprocessing steps minimized biases and improved the model's overall performance.

3. **Interface Usability**: The web interface successfully made the complex prediction model accessible to users without extensive technical backgrounds. By streamlining data input and output display, the interface contributed to a positive user experience, highlighting the feasibility of integrating machine learning models in practical applications.