# VISVESVARAYA TECHNOLOGICAL UNIVERSITY

**Jnana Sangama, Belagavi - 590018**

**Project Report**

**on**

# "LEVERAGING ANN FOR TARGETED DRUG SENSITIVITY PREDICTION ON GDSC DATA"

Submitted in partial fulfillment of the requirements for the award of the degree of

## BACHELOR OF ENGINEERING

**in**

## COMPUTER SCIENCE & ENGINEERING

**by**


| | |
|---|---|
| **DHANUSH A** | **4MT21CS046** |
| **DHANUSH S SHETTY** | **4MT21CS047** |
| **K SAHANA RAO** | **4MT21CS062** |
| **KARTHIK U SHETTIGAR** | **4MT21CS064** |

**Under the Guidance of**
**Mrs. Suma K**
**Assistant Professor**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

## MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust ®, Mangalore)
**Autonomous Institute Affiliated to V.T.U, Belagavi, Approved by AICTE, New Delhi**
Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution


**2024-25**

MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust ®, Mangalore)

**Autonomous Institute Affiliated to V.T.U, Belagavi, Approved by AICTE, New Delhi**

Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

# CERTIFICATE

Certified that the project work entitled **Leveraging ANN for Targeted Drug Sensitivity Prediction on GDSC Data** carried out by **Mr. DHANUSH A (USN: 4MT21CS046), Mr. DHANUSH S SHETTY (USN: 4MT21CS047), Ms. K SAHANA RAO (USN: 4MT21CS062) and Mr. KARTHIK U SHETTIGAR (USN: 4MT21CS064)** the bonafide student of **Mangalore Institute of Technology & Engineering** in partial fulfillment for the award of **Bachelor of Engineering** in **Computer Science & Engineering** of the **Visvesvaraya Technological University, Belagavi**, during the year **2024-25**. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report deposited in the departmental library. The project report has beenapproved as it satisfies the academic requirements in respect of Project Work prescribed for the said Degree.

| Signature of the Guide | Signature of the HOD | Signature of the Principal |
|---|---|---|
| **Mrs. Suma K** | **Dr. Ravinarayana B** | **Dr. Prashanth C M** |

**External Viva**

| **Name of the Examiners** | **Signature with Date** |
|---|---|
| **1.** | |
| **2.** | |

MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust ®, Mangalore)

**Autonomous Institute Affiliated to V.T.U, Belagavi, Approved by AICTE, New Delhi**

Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

# DECLARATION

We, **DHANUSH A**(4MT21CS046)**, DHANUSH S SHETTY** (4MT21CS047)**, K SAHANA RAO** (4MT21CS062) and **KARTHIK U SHETTIGAR** (4MT21CS064) student of 7th semester BE in Computer Science & Engineering, **Mangalore Institute of Technology and Engineering, Moodabidri,** hereby declare that the project work entitled "**Leveraging ANN for Targeted Drug Sensitivity Prediction on GDSC Data",** submitted to the **Visvesvaraya Technological University, Belagavi** during the academic year **2024-25**, is a record of an originalwork done by us under the guidance of **Mrs. Suma K,** Assistant Professor, Departmentof Computer Science & Engineering, Mangalore Institute of Technology and Engineering Moodabidri. This project work is submitted in partial fulfillment of the requirements for the awardof the degree of Bachelor of Engineering in Computer Science & Engineering. The results embodied in this report have not been submitted to any other University or Institute for the awardof any degree.

Date:

Place: MOODABIDRI

**DHANUSH A**
**DHANUSH S SHETTY**
**K SAHANA RAO**
**KARTHIK U SHETTIGAR**

# ABSTRACT

Personalized medicine is transforming cancer treatment by tailoring therapies to the genetic variations found in cancer cells, significantly improving patient responses to specific drugs. This project, titled "Leveraging ANN for Targeted Drug Sensitivity Prediction on GDSC Data", utilizes a machine learning approach with Artificial Neural Networks (ANN) to predict drug sensitivity in cancer cell lines using data from the Genomics of Drug Sensitivity in Cancer (GDSC) dataset. The model focuses on predicting IC50 values, a key measure of drug effectiveness, by integrating genomic features such as gene mutations, gene expression profiles, and copy number alterations (CNAs) to identify potential biomarkers for personalized therapies. The system employs an ANN architecture that captures complex nonlinear relationships within the genomic data, optimizing predictions with techniques like normalization, feature selection, and hyperparameter tuning. The approach uses a Flask-based web interface that enables users to easily input genomic data through labeled form fields, ensuring accessibility without the need for file uploads. The interface displays IC50 predictions in real-time, making it suitable for both clinical and research use. Model performance is evaluated using key metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE) and R-squared ($R^2$). Visualization tools further enhance interpretability, displaying the comparison between predicted and actual IC50 values. This project highlights the potential of ANN models in advancing personalized oncology treatment by providing actionable drug response insights based on genomic data. Future work will focus on integrating additional datasets, enhancing multi-omics data integration, and conducting clinical validation to ensure broader applicability in real-world cancer therapies.

# ACKNOWLEDGEMENTS

The satisfaction and the successful completion of this project would be incomplete without the mention of the people who made it possible, whose constant guidance encouragement crowned our efforts with success.

This project is made under the guidance of **Mrs. Suma K**, **Assistant Professor**, in the Department of Computer Science and Engineering. I would like to express my sincere gratitudeto my guide for all the helping hand and guidance in this project.

I would like to thank my project coordinator **Mr. Shivaraj B G, Assistant Professor** in the Department of Computer Science and Engineering, for his cordial support, valuable information and guidance, which helped me in completing this project through the various stages.

I would like to express appreciation to **Dr. Ravinarayana B, Professor and Head**, Department of Computer Science and Engineering, for his support and guidance.

I would like to thank our Principal **Dr. Prashanth C M**, for encouraging me and giving me an opportunity to accomplish the project.

I also thank our management who helped me directly or indirectly in the completion of this project.

My special thanks to faculty members and others for their constant help and support.

Above all, I extend my sincere gratitude to my parents and friends for their constant encouragement with moral support.

**DHANUSH A**
**DHANUSH S SHETTY**
**K SAHANA RAO**
**KARTHIK U SHETTIGAR**

# TABLE OF CONTENTS


| Contents | | | Page No |
|---|---|---|---|
| **ABSTRACT** | | | i |
| **ACKNOWLEDGEMENT** | | | ii |
| **TABLE OF CONTENTS** | | | iii |
| **LIST OF FIGURES** | | | Vi |
| **LIST OF TABLES** | | | vii |
| **Chapter no** | | **TITLE** | |
| **1.** | | **INTRODUCTION** | 1-4 |
| | 1.1 | Introduction | 1 |
| | 1.2 | Problem Statement | 2 |
| | 1.3 | Objectives | 3 |
| | 1.4 | Scope of the project | 3 |
| | 1.5 | Organization of the report | 4 |
| **2** | | **LITERATURE SURVEY** | 5-8 |
| | 2.1 | Existing System | 5 |
| | 2.2 | Limitations of existing systems | 7 |
| | 2.3 | Proposed System | 8 |
| **3** | | **SYSTEM REQUIREMENTS SPECIFICATION** | 9-13 |
| | 3.1 | Overall description | 9 |
| | | 3.1.1 Product Perspective | 9 |
| | | 3.1.2 Product Functions | 10 |
| | | 3.1.3 User classes and Characteristics | 10 |
| | | 3.1.4 Design and Implementation constraints | 10 |
| | | 3.1.5 Assumptions and Dependencies | 10 |
| | 3.2 | Specific Requirements | 11 |
| | | 3.2.1 Hardware Requirements | 11 |
| | | 3.2.1 Software Requirements | 11 |
| | 3.3 | Functional Requirements | 12 |
| | | 3.3.1 Model Development and Prediction | 12 |
| | | 3.3.1 User Interface and Accessibility | 12 |
| | | 3.3.3 Model Validation and Evaluation | 12 |

| | | | |
|---|---|---|---|
| | | 3.3.4 Data Security | 13 |
| | 3.4 | Non-Functional Requirements | 13 |
| | | 3.4.1 Usability | 11 |
| | | 3.4.1 Reliability | 11 |
| | | 3.4.3 Performance | 13 |
| | | 3.4.4 Maintainability | 13 |
| **4** | | **GANTT CHART** | 15-16 |
| **5** | | **SYSTEM DESIGN** | 17-22 |
| | 5.1 | Architecture Diagram | 17 |
| | 5.2 | Use Case Diagram and Description | 19 |
| | 5.3 | Sequence Diagram | 20 |
| | 5.4 | Activity Diagram | 21 |
| | 5.5 | Data flow Diagram | 22 |
| **6** | | **IMPLEMENTATION** | 23-28 |
| | 6.1 | Module Implementation | 23 |
| | 6.2 | Data Preprocessing | 23 |
| | 6.3 | Model Development | 24 |
| | 6.4 | Web Interface Development | 27 |
| | 6.5 | Deployment | 28 |
| **7** | | **TESTING** | 29-32 |
| | 7.1 | System Testing | 29 |
| | 7.2 | Testing Methodology | 29 |
| | 7.3 | Test cases | 29 |
| | 7.4 | Evaluation Metrics | 31 |
| | 7.5 | Test Results and Analysis | 31 |
| | 7.6 | Challenges and Resolutions | 32 |
| **8** | | **RESULTS AND SNAPSHOTS** | 33-39 |
| | 8.1 | Introduction | 33 |
| | 8.2 | Model Performance Metrics | 33 |
| | 8.3 | Visualizations of Model Predictions | 34 |
| | 8.4 | Snapshots of Web Interface | 36 |
| | 8.5 | User Feedback | 38 |
| | 8.6 | Observations and Insights | 39 |

| | | | |
|---|---|---|---|
| **9** | | **CONCLUSION AND FUTURE WORK** | 40 |
| | 9.1 | Conclusion | 40 |
| | 9.2 | Future Work | 41 |
| | | **REFERENCES** | 43 |
| | | **PAPERS PRESENTED** | 44 |

# LIST OF FIGURES


| Figure No. | Figure Name | Page No. |
|---|---|---|
| Figure 4.1 | Gantt Chart | 16 |
| Figure 5.1 | Architectural Diagram | 18 |
| Figure 5.2 | Use Case Diagram | 19 |
| Figure 5.3 | Sequence Diagram | 20 |
| Figure 5.4 | Activity Diagram | 21 |
| Figure 5.5 | Data Flow Diagram | 22 |
| Figure 8.1 | Error Metrics for Model Evaluation | 34 |
| Figure 8.2 | Actual vs Predicted LN_IC50 Values | 34 |
| Figure 8.3 | Training vs Testing Loss | 35 |
| Figure 8.4 | Drug Response Prediction | 35 |
| Figure 8.5 | Home Page | 36 |
| Figure 8.6 | Predicted IC50 Page | 37 |
| Figure 8.7 | Drug Lookup Page | 38 |
| Figure 8.8 | Drug Response Prediction | 38 |

# LIST OF TABLES


| Table No. | Table Name | Page No. |
|---|---|---|
| Table 4.1 | Gantt Chart of Planning and Scheduling of Project | 15 |
| Table 7.1 | Test Cases | 30 |