

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)



Predicting Cancer Treatment Response Using Machine Learning and Tabular Deep Learning

Zahraa Naser Shahweli

Al Nahrain University, Baghdad- Iraq. Email: zahra.n@nahrainuniv.edu.iq

ARTICLE INFO

Article history:

Received: 19 /02/2026

Revised form: 04 /05/2026

Accepted : 06 /05/2026

Available online: 30 /09/2026

Keywords:

Cancer Treatment,
Treatment Response Prediction,
Machine Learning,
Tabular Deep Learning,
Clinical Data,
Classification.

ABSTRACT

Cancer is one of the most difficult and prevalent diseases around the globe. It is characterized by different signs among patients and their mode of treatments also varies. Despite the development of numerous advanced therapeutic approaches, selecting the most suitable treatment for each of the patients still remains one of the main problems. This study aims to develop a predictive framework according to machine learning as well as deep learning algorithms to determine a patient's response to the cancer treatment using the Cancer Treatment Performance Dashboard Dataset. The proposed approach involved data preprocessing, feature selection and model training using: Logistic Regression (LR), Extra Trees (ET), CatBoost (CB), and TabNet (TN). The framework was evaluated using 5-fold cross validation and multiple evaluation metrics including accuracy, precision, recall, F1-score and AUC-ROC. The experimental results indicated that TabNet achieved the highest performance in term of F1-score (0.90) and AUC-ROC, followed closely to CatBoost (F1-score=0.89) which reflects their ability to comply with complex nonlinear relationship in clinical data. In contrast, Extra trees demonstrated competitive performance (F1-score=0.86), while logistic regression demonstrated lower performance (F1-score=0.83) due its linear nature. This study suggests that predict patients' treatment response at an early stage using machine learning and deep learning algorithms in analyzing clinical and therapeutic data will assist to predicting cancer treatment response to cancer and can support clinical decision-making by enabling the early identification of patients likely to benefit from specific treatments, thereby improving treatment outcomes and reducing exposure to ineffective treatments.

MSC..

<https://doi.org/10.29304/jqcm.2026.18.32670>

1. Introduction

Cancer is one of common disease that is considered a global threat in the 21st century. It is one of the diseases that are associated with the causes of death in many countries [1], posing a significant burden on healthcare systems. There are numerous types of tumors, more than a hundred types have been identified and they vary in their biological nature, extend of spread, speed of growth and their response to different treatments [2]. Each type of cancers has distinct pathological and clinical features that need therapeutic plans of treatment and stages of progression [3].

The basic treatments used may include a broader range of medical methods which include surgery, chemotherapy, immunotherapy, radiotherapy, hormonal and targeted therapies. Though these treatments are efficient, their mode

*Corresponding author: Zahraa Naser Shahweli

Email addresses: zahra.n@nahrainuniv.edu.iq

Communicated by 'sub editor'

of response differs. The reasons for the differences may be due to numerous factors like the genetic differences among patients, state of health, biological structure of the tumor, the presence of the disease and the pharmacological interface. This variability in treatment response makes it hard for clinicians to precisely determine the optimum treatment strategy for individual patient and increase the possibility of patients being exposed to weak and strict side-effect treatments [4]. Figure (1) displays response cancer patients to therapy prepared by the author based on RECIST 1.1 criteria.

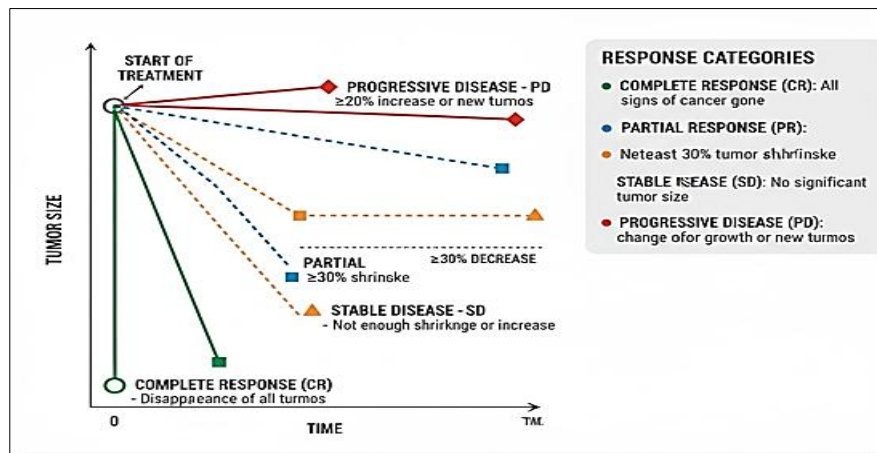


Fig. 1- Response Cancer Patients to Therapy

From these potential challenges, artificial intelligence (AI) manifested as promising mechanisms that can be used to transform cancer treatment. AI algorithms offer the capability to analyze large amounts of clinical, genetic and demographic data and unravel disease that may not be obvious to the clinicians. These latest innovations can assist in the prediction of the response of patients to various treatments and determine the most efficient treatment for each of the patients [5].

Based on this information, the study will focus on the development and evaluation of a predictive model that will incorporate machine learning (ML) and deep learning (DL) algorithms to assume patients' response to cancer treatment with the use of clinical and therapeutic information from the Cancer Treatment Performance Dashboard Dataset. This study offers a systematic comparison between tree-based models (CatBoost and Extra Trees), a linear reference model (logistic regression), as well as a deep learning model dedicated to tabular data (TabNet), within a united framework of pre-processing and assessment. The main contributions of this study are:

1. Developing predictive framework for cancer treatment response using both ML (CatBoost, Extra Trees, Logistic Regression) and tabular DL (TabNet), Within a consistent preprocessing pipeline and a structured evaluation protocol (5-fold cross-validation), ensuring the transparency and consistency of comparison between results.
2. Designing a leakage-aware modeling pipeline, where only pretreatment features are retained to ensure applying the model clinically.
3. Incorporating class imbalance pathway with cross-validation, using SMOTE, applying it to training sets only, to avoid bias and upgrade the model robustness.
4. Using multiple performance measures (specificity, recall and AUC-ROC) for clinical assessment, where specificity ensures a false positive cases are reduced, Recall reflects the model's ability to identify responding patients correctly, and, both of which are critical for making safe and effective treatment decisions. AUC-ROC measures how the effectively the model distinguish between respondents and non-responders.
5. Analyzing the model's suitability for clinical approach, comparison with some previous studies, and suggestion for future implementation.

The rest of this study is organized as follows: section (2) is the review of the related works. Section (3) offers materials and methods used. The section (4) contains the results and discussion of the performance. The conclusion and future directions of study is in section (5).

2. Related Work

In latest years, artificial intelligence (AI) algorithms have become widely used, in predicting patients' response to cancer treatment, due to the availability of clinical databases [6]. A number of studies summarized in table (1) were selected based on their focus on predicting cancer patient response to treatment across multiple data types (clinical, imaging, and genomic data). Although these studies differ in terms of data sources and cancer type.

Table 1- Comparison of the proposed model with previous studies.

Ref	Data Type	Models	Metric	Best Score
[7]	Imaging (NSCLC)	CNN + RNN	AUC	0.67
[8]	Clinical	XGBoost, RF, LR	Accuracy	0.67
[9]	Histopathology	CNN + Transformer	AUC	0.71
[10]	Clinical + fMRI	Elastic Net, SVR	AUC	0.72
[11]	Clinical + Imaging	VGG16, ResNet	Accuracy	0.82
This Work	Clinical + Therapeutic	CB, ET, LR, TabNet	F1	0.90

The comparison in table above places the proposed work to be placed within a broader research context and presents how different data sources and modeling techniques affect predictive performance. Through seeking, it was found that the cancer treatment performance dashboard dataset has not been widely used in previous studies, underscoring the novelty of this work. Therefore, comparisons performed with previous studies that have similar problem rather than different datasets.

3. Material and Methods

3.1. Data source

The dataset on the Kaggle platform (Cancer Treatment Performance Dashboard Dataset) [12] was used in this study, which included number of features such as treatment regimens, diagnoses and treatment response results. The number of sample was 450 with 15 features with multiple types of cancers such as (breast, prostate, colorectal, etc.), this diversity has been treated as part of the heterogenous nature of clinical data to enhance the realism and generalizability. To build a predictive framework that distinguishes between responsive and non-responsive cancer patients, a set of therapeutic and clinical variables was chosen. The choice of the features will focus on the early stages of the treatment strategy like the level of cancer, the treatment structure and its seriousness with the aim of building a genuine and clinically potent model.

3.2. Data Preprocessing

The Data Preprocessing stage is a very important step in building reliable machine framework and the data will directly affect the performance of the frameworks and the precision of their findings. Because clinical data usually contain missing values, variations in metrics imbalance in categories and categorical variables, a sequence of predetermined stages that have been used to affirm the data preparedness to be used for the modelling stage:

1. **Addressing Missing Values:** The dataset was examined for missing values in each column. Initially, three columns were deleted because they contained more than 50% of the missing values. Furthermore, defining variables, such as patient ID and post-treatment outcomes, were excluded to avoid introducing non-systematic biases into the training and evaluation process during or after the early stages of treatment.
For the remaining features, appropriate imputation strategies were applied based on the data type. Numerical features were imputed using the median to reduce the impact of outliers, while categorical features were imputed

using the most frequent category. Furthermore, only features available prior to or during the early stages of treatment were retained to ensure that the predictive models rely solely on clinically relevant and temporally valid information. After preprocessing and feature selection, a total of 7 features were retained for model development, as summarized in table (2).

Table 2- Selected Features for Predicting Treatment Response

Feature	Description	Reason of choice
Diagnosis	Type of cancer	It directly affects treatment effectiveness and the quality of response.
Stage	Stage of cancer	It is one of the most important clinical determinants of disease severity and the likelihood of response.
Drug_Regimen	Treatment regimen used	It determines the type of drug/treatment and thus directly affects the response.
Number_of_Cycles	Number of treatment cycle	It reflects the intensity and duration of treatment and is related to achieving a response.
Treatment_Cost_USD	Cost of treatment	An indirect indicator of the type and intensity of treatment used.
Time_to_Response_Days	Time to initial response	Reflects the dynamic patient interaction with the treatment.
Follow_Up_Visits	Number of follow-up visits	An indicator of monitoring the condition and the stability/change in response.

2. Feature Engineering: Date related features were processed to extract meaningful information relevant to response prediction such as diagnosis date and treatment start date were converted to numerical form to ensure compatibility to machine learning models.
3. Coding of categorical variables: The categorical features were encrypted based on the requirements of the algorithms applied. Algorithm such as the CatBoost can manage categorical inputs locally, whereas, models such as Extra Trees and Logistic Regression need obvious coding. One-Hot encoding regression encryption has been applied. Therefore, numerous derivative features have emerged to improve the predictive power of the models. As shown in the table (3) below:

Table 3- Derivative Features

Original Feature	Type	Derivative Features	Type	Derivation Notes
Diagnosis	categorical	Diagnosis	numeric	Using one-hot encoding regression
Drug Regimen	categorical	Drug Regimen	numeric	number of medications for each patient
Number_of_Cycles	numeric	Total doses (Number_of_Cycles × Number of drugs in Regimen)	numeric	actual therapeutic pregnancy
Treatment_Cost_USD	numeric	Low=0, Medium=1, High=2	numeric	To facilitate interpretation
Time_to_Response_Days	numeric	Fast=2, Medium=1, Slow=0	numeric	To facilitate interpretation
Follow_Up_Visits	numeric	Effective follow-up indicator = Follow_Up_Visits / Time_to_Response_Days	numeric	Reflects the intensity of follow-up relative to time

- Addressing class imbalances: The analysis of categories showed a relative imbalance in the distribution, with the percentage of responding patients being approximately (73%) compared to (27%) for non-responders. To address the instable outcome categories, reconfiguration procedures were applied. Options comprised were oversampling with SMOTE in line with the features of the dataset.
- Data segmentation and validation: The dataset was categorized with the use of K-Fold stratified cross-verification to affirm proportional representation of groups across different folds assisting to obtain unbiased performance estimates.

The Treatment Response Score variable was converted to binary variable to represent: Responder or Non-Responder, as target variable.

3.3. Machine Learning and Deep Learning

In this study, three machine learning algorithms were chosen to categorize patients according to clinical and therapeutic characteristics: Logistic Regression, Extra trees and CatBoost, to complement by TabNet for much deeper tabular learning. The selection of each model was caused by its capability to handle the specific features of the dataset as follows:

- CatBoost [13] is a gradient boosting algorithm that originally can handle categorical features without the need for the prolonged and preprocessing likewise such as rapid coding [14]. This characteristic is necessary for the Cancer Treatment Performance Dashboard dataset, which contains many categorical variables such as (diagnosis, drug regimen and stage). Furthermore, CatBoost's robust organizational innovations help control over reporting in datasets with hybrid features types. Its efficacy makes it suitable for clinical categorization tasks.
- Logistic Regression [15] is a classical linear model widely applied for binary allows clinical insights in the individual contribution features [16].
- Extra Trees [17] is a tree-based approach that creates many random choice trees and the average in their forecasts. This model is then known for its quality in the face of noise, its capability to capture the correlation of non-linear attributes and its rapid training time. It can efficiently handle digital qualities and coded categorical variables [18]

, which provide a complementary aspect to CatBoost by aligning serious randomness to decrease variance and enhance generation.

4. TabNet [19] is a deep learning plan specifically kept for tabular information, its capability to deal with the model a tricky relationship and focus on the essential features and makes it particularly appropriate datasets with a blend of numerical and categorical features [20]. TabNet also supports the detailed training with the minimum pretreatment, to enable efficient learning from primary, therapeutic variables. The overall workflow of the proposed approach is illustrated in figure (2).

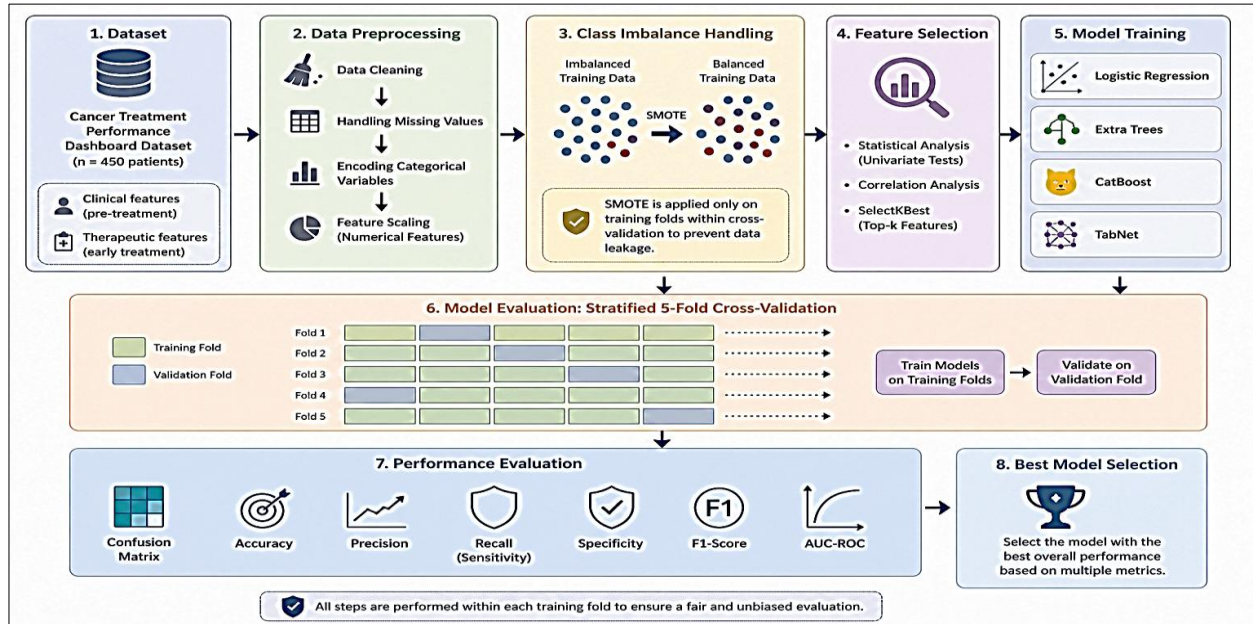


Fig. 2- Workflow of Proposed Approach.

3.4. Model Evaluation

Imbalance of the categories in the dataset and the medical tendencies of this study, numerous evaluation scales been used to verify a consistent and accurate of the proposed model. Relying on accuracy alone can be misleading in these cases. The model can achieve high accuracy with poor differentiation between clinically essential groups [21]. Thus, supplementary metrics will include Precision, Recall, F1- score, specificity and Area under the ROC Curve (AUC-ROC) were utilized to assess the models form several angles. This is critical in a clinical situation where predictive faults may produce serious therapeutic consequences.

1. Accuracy: gives percentage of samples which correctly classified:

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} * 100\% \quad (1)$$

2. Precision: evaluates positive prediction from all prediction:

$$Precision = \frac{TP}{TP+FP} * 100\% \quad (2)$$

3. Recall: Measures the extent to which positive cases are correctly identified among relevant cases:

$$Recall = \frac{TP}{TP+FN} * 100\% \quad (3)$$

4. specificity: Measures the extent to which negative cases are correctly identified among relevant cases:

$$Specificity = \frac{TN}{TN+FP} * 100\% \quad (4)$$

5. F1 score: is the harmonic mean of precision and recall:

$$F1 - score = \frac{2*(precision*Recal)}{(Precision+Recal)} \quad (5)$$

6. Area under the ROC Curve (AUC-ROC): plots the True Positive Rate (Recall) against the False Positive Rate.

Where TP, TN, FP, and FN denote the number of true positives, true negatives, false positives, and false negatives, respectively [22].

4. Result, Discussion and Comparison with Previous Studies

As previously mentioned, the primary objective of this study is to develop a precise AI model to distinguish between cancer patients' response to treatment and their lack of response. The models based on apply 5-fold cross validation, with in each fold:

1. The data was divided into two set, training set and validation set.
2. SMOTE technology applied to the training set with in each fold to prevent leakage. Specifically, SMOTE was configured using a sampling strategy of auto, which balances the minority class to match the majority class, and the number of nearest neighbors was set to k = 5, as commonly adopted in the literature.
3. The implementation process carried out by following parameters:
number of iterations=300, learning rate=0.05, and tree depth=6 for CatBoost algorithm,
The trees = 200 , maximum depth = 10 for ET.
L2 regularization and a standard solver (lbfgs) for LR. For the TabNet model, the number of decision steps=3, feature dimensions=8, and learning rate=0.02 for TabNet model. A fixed random seed was applied in all models equal to 42.

In general, these parameters were chosen to complete evaluation of the several factors through balance predictive performance, interpretability and robustness of the dataset to allow influence treatment explanation.

To enhance the reliability of the results, the 5-fold cross validation was repeated several times.

The training and testing accuracy were computed; also training and testing loss values were calculated for the models. As shown in table (4), lower loss values indicate a better fit between the model and the data, as seen with TabNet and CatBoost. Less training and testing loss compared to logistic regression suggests more efficient learning of complex patterns. Furthermore, the small gap between training and testing loss indicates that the models did not suffer from over-allocation.

Furthermore, the standard deviation (std.) values of the 5-fold cross validation confirmed model stability. This enhances reliability of the results and confidences the reliability of differences between models.

Table 4- The Training and Testing Accuracy for All Models

Algorithm	Training (Mean ± Std)	Acc.	Testing (Mean ± Std)	Acc.	Training (Mean ± Std)	Loss	Testing Loss (Mean ± Std)
Cat Boost	0.9023		0.8951		0.19		0.23
Logistic Regression	0.8744		0.8574		0.31		0.38
Extra Trees	0.8981		0.8841		0.23		0.25
Tab Net	0.9004		0.902		0.15		0.19

Where table (5) shows a performance's comparison of the proposed models.

Table 5- Comparison between Models

Model	Precision	Recall	F1-score	Specificity	AUC
CatBoost	0.9112	0.9074	0.89	0.9060	0.85
Logistic Regression	0.8678	0.8801	0.83	0.8722	0.81
Extra Trees	0.8905	0.893	0.86	0.8994	0.86
Tab Net	0.9054	0.9133	0.90	0.9149	0.89

Table 4 provides a comprehensive assessment of the models' generalizability through accuracy and loss of the training and testing performance of four models. In contrast, Table 5, offers a more detailed comparison using accuracy, recall, F1 score, specificity, and AUC, which are significantly important in the medical approach. Together, these two tables summarize a complete evaluation of the models' performance.

The TabNet model achieved best performance in terms of F1 score and AUC-ROC value with a slight advancement compared to CatBoost. This reflects its strong ability to handle variables that are peculiar to clinical data and complex relationships between them. On the other hand, this requires increase computational complexity and more control of superior terms and time in comparison with the traditional tree models. From another perspective, LR showed the lowest computational cost with minimal training and inference time, while the ET showed fast training speed. In contrast, the CatBoost model required a longer training time due to its gradient optimization mechanism, but the inference time remained efficient.

An F1 value close to 0.90 for TabNet indicates a clear balance between accuracy and recall, illustrating the model's efficiency in identifying both treatment responders and non-responders. Practically, approximately 10% of treatment responders being misclassified as false negatives. This is critical in clinical approaches, as misclassification may deprive patients who could benefit from effective treatments.

Similarly, accuracy reflects the positive predictions. The high-rate accuracy achieved by CatBoost indicates a lower false positive. This highlights an essential trade-off between false negative rate (measured by recall) and false positive rate (measured by accuracy), both of which have important clinical effects.

Specificity also accurately identifies patients who are not responding to treatment. High specificity reduces the prescribing of ineffective treatments. In this study, both TabNet and CatBoost demonstrated high specificity, reinforcing their role in supporting clinical decision-making. The comparison between the results is clearly shown by figure (3)

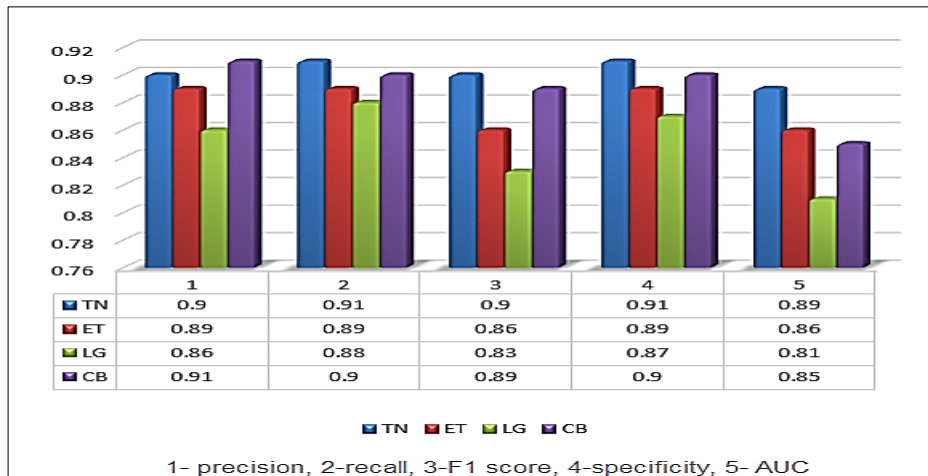


Fig. 3- Comparison between Models.

Furthermore, the ROC curve analysis further supports the assessment, with high value of AUC ranging (0.81 - 0.89), which supports the ability to differentiate between responders and non-responders across different thresholds. On other hand, the results showed that using SMOTE improved the models' ability compared to model trained on the original data. This improvement is critical in the medical approach.

The above results showed that clinical and therapeutic features, such as (cancer type, treatment regimen, and intensity) play an essential role in determining patient response, these results align with established medical information. For example, cancer's type and stage affect patient response to different treatment. These results not only enhance the model's transparency but also support its clinical applicability by highlighting key factors influencing treatment decisions.

All of the proposed systems were implemented using the Python programming language (version 3.12), utilizing the PyTorch deep learning framework, an Intel® i7 processor, and 32 GB of RAM.

Table (6) presents a comparison with previous studies that used different datasets, as the database used in this study is being used for the first time. This comparison is essential to highlight the contribution of this study, not only through the use of a new dataset, but also through achieving competitive results compared to those found in the relevant studies.

Table 6- Comparison with Previous Studies

Ref.	Technique	High Accuracy	Precision	Recall	F1 score	AUC	Data type
[23]	ANN	77.78	0.816	0.882	0.816	0.75	clinical and genomic data
[24]	L1-regularized logistic regression ,ANNs	-	82	79	80	0.78	clinical
Proposed Work	LR, ET, CatBoost ,TabNet	0.902	0.9054	0.9133	0.90	0.89	Clinical and therapeutic data

The previous studies used an artificial neural network (ANNs) and regularized logistic regression, achieving levels of performance within the average range, such as accuracy and other evaluation metrics.

In contrast, the proposed work achieved relatively superior performance, with AUC of 0.89. This reflects the ability of the models to effectively handle clinical and therapeutic data and extract complex patterns from them.

This improvement can be attributed to the adoption of a precise preprocessing steps and addressing data imbalances, which contributed to improved prediction quality compared to previous studies.

5. Conclusion and Future Work

This study offered an integrative predictive framework for the prediction of patients' response to cancer treatment based on a mixture of machine learning and deep learning algorithms, in consonance with CatBoost, Extra Trees, Logistic Regression, and TabNet, applying clinical and therapeutic data from the Cancer Treatment Performance Dashboard Dataset. The study engrossed on the levels of pre-processing data, which include data cleaning, values, processing missing values, dealing with category imbalances using SMOTE because of its effectiveness in generating synthetic samples while preserving valuable data, which is particularly important given the relatively small size of the dataset, then selecting features most essential prediction task, and evaluation models using stratified 5- fold cross validation and a set of standard performance measures.

The results indicated that nonlinear models, particularly CatBoost and TabNet, outperformed the basic linear model in predicting patient response to treatment. They achieved the best overall performance in terms of accuracy, F1 score, and AUC-ROC value, highlighting their suitability for handling tabular clinical data that combines quantitative and qualitative variables. In contrast, while logistic regression is simple and easy to interpret, its predictive power was relatively good, and the Extra Trees model also yielded satisfactory and stable results.

From clinical perspective, the proposed model support decision making, by assisting to predict early patient response to treatment, thus contributing to improve treatment planning, reduce exposure to ineffective therapies, and advance personalized medicine. However, this study aimed to assist clinicians, rather than replace clinical judgment.

Despite these impressive results, several limitations hampered the work; most significant is the small size of the dataset, which may limit the generalizability of the models. Furthermore, a single dataset does not practically reflect the diversity of clinical categories. Finally, there is a lack of external validation to ensure the robustness of the proposed framework.

Future suggestions include, validating the models using larger and more diverse clinical datasets containing genomic or imaging features; and expanding the work methodology to support multi category classification, which would enhance its clinical applicability.

References

- [1] G. A. C. Karen M. Emmons, "Realizing the Potential of Cancer Prevention — The Role of Implementation Science," *The New England Journal of Medicine*, vol. 376, no. 10, 2017. DOI: 10.1056/NEJMs1609101.
- [2] L. e. a. Galluzzi, "Cancer cell death: Cell-autonomous and immunogenic dimensions," *Cancer Cell*, vol. 44, no. 2, pp. 281-305, 2025. DOI: 10.1016/j.ccell.2025.12.005
- [3] S. S. R. F. a. e. a. Keene L. Abbott, "Site of breast cancer metastasis is independent of single nutrient levels," *bioRxiv*, vol. 10, no. 24, 2024. DOI: 10.1101/2024.10.24.616714
- [4] S. M. P. A. Eric K. A. Atsu, "An Exploratory Study of ResNet and Capsule Neural Networks for Brain Tumor Detection in MRI," *medRxiv*, vol. 2, no. 5, 2026.
- [5] M. S.-T. R. B.-R. Claudia Mendoza-Pinto, "Machine learning in the prediction of treatment response in rheumatoid arthritis: A systematic review," *Seminars in Arthritis and Rheumatism*, vol. 68, 2024. DOI: 10.1016/j.semarthrit.2024.152501
- [6] C. D. Joshua Curtiss, "Machine learning in the prediction of treatment response for emotional disorders: A systematic review and meta-analysis," *Clinical Psychology Review*, vol. 120, 2025. DOI: 10.1016/j.cpr.2025.102593
- [7] A. H. R. Z. C. P. T. C. I. F. R. H. M. a. H. J. A. Yiwen Xu, "Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging," *Clin Cancer Res.*, vol. 25, no. 11, 2019. DOI: 10.1158/1078-0432.CCR-18-2495
- [8] C.-Y. C. T.-H. C. Chun-Chi Lai, "Predicting Pathological Complete Response Following Neoadjuvant Therapy in Patients With Breast Cancer: Development of Machine Learning–Based Prediction Models in a Retrospective Study," *JMIR CANCER*, vol. 1, no. 11, 2025. DOI: <https://doi.org/10.1038/s41598-025-99496-6>
- [9] W. C. R. R. Z. Z. a. e. a. Yifan Liu1, "Deep learning based digital pathology for predicting treatment response to first-line PD-1 blockade in advanced gastric cancer," *Journal of Translational Medicine*, vol. 22, no. 438, 2024. DOI: 10.1186/s12967-024-05262-z
- [10] M. H. T. M. W. a. e. a. Peter Zhukovsky, "Generalizability of Treatment Outcome Prediction Across Antidepressant Treatment Trials in Depression," *JAMA Network Open*, vol. 8, no. 3, 2025. DOI: 10.1001/jamanetworkopen.2025.1310

-
- [11] F. A. M. M. T. e. a. Butt, "Deep learning-driven prediction of chemotherapy response in breast cancer: a pathway toward precision medicine," *Soft Comput*, 2026. DOI: <https://doi.org/10.1007/s00500-025-11048-2>
 - [12] Kaggle, "Cancer Treatment Performance Dashboard Dataset," 2023. [Online]. Available: <https://www.kaggle.com/datasets/olagokeblissman/cancer-treatment-performance-dashboard-dataset>. URL: <https://www.kaggle.com/datasets/olagokeblissman/cancer-treatment-performance-dashboard-dataset>
 - [13] G. G. A. V. A. V. D. a. A. G. L. Prokhorenkova, "CatBoost: unbiased boosting with categorical features," in *32nd Conference on Neural Information Processing Systems (NeurIPS 2018)*, Montréal, Canada, 2018.
 - [14] G. N. Z. H. Li C, "Anticancer drug synergy prediction based on CatBoost," *PeerJ Computer Science*, vol. 11, no. e2829, 2025. DOI: 10.7717/peerj-cs.2829
 - [15] S. L. a. R. X. S. D. W. Hosmer, *Applied Logistic Regression*, W. S. i. P. a. statistics, Ed., 3rd ed. Wiley, 2013. DOI: 10.1002/9781118548387
 - [16] C. C. J. F. M. F. R. H. K. Goto T, "Machine Learning-Based Prediction of Clinical Outcomes for Children During Emergency Department Triage," *JAMA Netw Open*, vol. 2, no. 1, 2019. DOI: 10.1001/jamanetworkopen.2018.6937
 - [17] D. E. a. L. W. P. Geurts, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3-42, 2006. DOI: <https://doi.org/10.1007/s10994-006-6226-1>
 - [18] S. G. H. M. A. K. S. Z. A. Arya M, "A Novel Extra Tree Ensemble Optimized DL Framework (ETEODL) for Early Detection of Diabetes," *Front Public Health*, vol. 9, 2022. DOI: <https://doi.org/10.3389/fpubh.2021.797877>
 - [19] S. Ö. A. a. T. Pfister, "TabNet: Attentive interpretable tabular learning," in *in Proceedings of the AAAI Conference on Artificial Intelligence*, 2021. DOI: <https://doi.org/10.48550/arXiv.1908.07442>
 - [20] Z. W. G. X. Z. Z. & C. Z. Maiqiang Zhou, "Deep learning with TabNet: rapid coal ash content estimation via X-ray fluorescence," *International Journal of Coal Preparation and Utilization*, vol. 45, pp. 523-547, 2024. DOI: 10.1080/19392699.2024.2341952
 - [21] S. S. I. T. V. e. a. Hicks, "On evaluation metrics for medical applications of artificial intelligence," *Sci Rep*, vol. 12, no. 5979, 2022. DOI: <https://doi.org/10.1038/s41598-022-09954-8>
 - [22] z. n. s. weli, "Deep Belief Network for Predicting the Predisposition to Lung Cancer in TP53 Gene," *Iraqi Jornal of Science*, vol. 61, no. 1, pp. 171-177, 2020. DOI: <https://doi.org/10.24996/ijjs.2020.61.1.19>
 - [23] S. B. Sri Chandana K, "Cancer Treatment Recommendation using Machine Learning Technologies," *International Journal of Precious Engineering Research and Applications (IJPERA)*, vol. 7, no. 1, pp. 83-88, 2022.
 - [24] M. Cruz-Fernandez, F. Castillo-Velásquez, C. Fuentes-Silva, O. Rodríguez-Abreo, R. Rojas-Galván, M. Avilés and J. Rodríguez-Reséndiz, "Predicting Breast Cancer Mortality Using SEER Data: A Comparative Analysis of L1-Logistic Regression and Neural Networks," *Technologies*, vol. 14, no. 66, 2026. DOI: <https://doi.org/10.3390/technologies14010066>
-