RESEARCH PAPER

# AN EXPLAINABLE ENSEMBLE GENOMIC PIPELINE FOR CONSISTENT CANCER THERAPY DECISIONS

**Mrs. S. Vidhya[1], Dr. S. Prabakaran[2], Dr. V. Vijayakumar[3], Gobikha P S[4], Dr. P. Anbumani[5], Asmitha S[6]**

[1]Assistant Professor, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, Tamilnadu, India. Email: vidyasivasubramaniamsrv@gmail.com (Corresponding Author)

[2]Assistant Professor, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, Tamilnadu, India. Email: mokipraba@gmail.com

[3]Assistant Professor, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, India. Email: vksece2007@gmail.com

[4]UG Scholar, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, Tamilnadu, India. Email: psgobikha@gmail.com

[5]Assistant Professor, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, Tamilnadu, India. Email: anbuanc@gmail.com

[6]UG Scholar, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, India. Email: asmitha.siluvaiantony@gmail.com

***Corresponding author: Mrs. S. Vidhya, Assistant Professor, Department of Computer Science and Engineering, V.S.B. Engineering College, Karur, Tamilnadu, India**

**Email: vidyasivasubramaniamsrv@gmail.com**

**Received:** 30th May, 2026; **Revised:** 11th June, 2026; **Accepted:** 15th June, 2026; **Available Online:** 17th June, 2026

***ABSTRACT***

**Background**

Precision oncology is focused on the personalization of cancer treatment with references to the molecular profile of the patients. Nonetheless, the inconsistency of the interpretation of genomics, feature selection, and data preprocessing can usually result in varied treatment advice, restricting clinical repeatability and trust in AI-supported judgment.

**Objective**

To solve these issues, in this work, we will present an Explainable Ensemble Genomic Pipeline (XEGP) to help offer standardized and reproducible therapy advice.

**Materials and Methods**

The pipeline combines both powerful preprocessing of genomic data, high-resolution biomarker extraction, and sophisticated feature selection methods to minimize noise and inter-sample variability. An ensemble decision engine is a system that combines the forecasts of various machine learning models and enhances performance and reduces uncertainty during therapy recommendations. The explainable AI modules (feature attribution and decision visualization) in XEGP can be used to improve clinical interpretability by allowing an oncologist to learn how important mutations and expressions influence therapy choices.

**Results**

The findings of experimental testing on several benchmark cancer genomic datasets indicate a higher degree of consistency of the decision-making, predictive accuracy, and reproducibility as compared to the traditional methods.

**Conclusion**

The suggested framework can be applied to various cancer types and groups of patients, and it will provide a clinically understandable and user-friendly AI-assisted tool. XEGP allows the realization of the gap between high-throughput genomic analysis and reproducible and reliable precision oncology choices: they are all in the same framework to unify the preprocessing, ensemble modeling, and explainability process.

**Keywords:** Precision oncology, Genomic pipeline, Explainable AI, Ensemble learning, Biomarker selection, Cancer therapy, Clinical interpretability, Reproducibility.

**How to cite this article:** Vidhya S, Prabakaran S, Vijayakumar V, Gobikha PS, Anbumani P, Asmitha S. An Explainable Ensemble Genomic Pipeline for Consistent Cancer Therapy Decisions. Int J Drug Deliv Technol. 2026;16(60s):1717-1723. DOI: 10.25258/ijddt.16.60s.155

**Source of support:** Nil.

**Conflict of interest:** None

## I. INTRODUCTION

Cancer has been considered as one of the most common causes of death in the whole world and the fact that it is a heterogeneous disease among patients presents a great challenge to proper diagnosis and the proper treatment plan. Individualized genomic therapy has become a potential solution to these issues taking the form of precision oncology. Multi-omics and high-throughput sequencing technologies have made it possible to amass large amounts of genome and proteomic data that enables making cancer diagnosis and treatment decisions based on data [2], [5]. Nevertheless, it is a major issue to convert this information into uniform and clinically relevant advice because of the fluctuation in preprocessing, feature selection and predictive modeling procedures [1], [3]. The latest research has

demonstrated the possible use of machine learning and deep learning solutions in cancer diagnostics and treatment recommendations. Ensemble learning techniques, where several models are integrated in order to enhance the accuracy of the prediction, have shown better results than single models in cancer detection problems especially in the case of heterogeneous genomic data [1], [3]. There are also graph-based methods like Graph Convolutional Networks (GCNs) that have been effectively used to identify cancer driver genes through the combination of protein-protein interaction networks and multi-omics data, and have enhanced the robustness of prediction and suggested new therapeutic targets [4]. Although these developments have been made, the currently available strategies tend to be lack-of-transparency, which reduces their ability to be adopted in clinical settings. Explainable artificial intelligence (XAI) methods have thus become crucial to interpretability and allowing clinicians to interpret the role of important features, including mutations or patterns of gene expression, in the decision to use therapy [2], [5]. This paper is inspired by these issues and suggests an Explainable Ensemble Genomic Pipeline (EEGP) to achieve uniform cancer therapy recommendations. Through a unification of standardized genomic preprocessing, powerful feature selection, ensemble-based predictive modelling, and explainable AI modules, the proposed framework is expected to deliver reproducible, accurate, and explainable treatment options in different types of cancer and in a variety of patient cohorts.

## II. Related Works

The current developments in machine learning (ML) and deep learning (DL) have significantly contributed to the enhancement of cancer diagnosis and therapy recommendation through the aid of genomic and multimodal data. The accuracy of supervised learning methods like Logistic Regression, Support Vector Machines and random forests has proven high to predict the type of cancer through the data of gene expression. As an example, Nair et al. [6] applied multiple supervised algorithms to RNA-Sequestrusted datasets of about 1000 samples with a maximum accuracy of 98.22% with Logistic Regression. Their survey emphasizes the usefulness of traditional ML methods in massive genomic examination and characterizing phenotypes, which creates a solid basis of accuracy in oncology. The combination of longitudinal and multimodal data has also contributed to the improved knowledge of cancer development. Zhuang et al. [7] highlighted that when modeling the dynamic heterogeneity of cancer, it is important to model temporal and multimodal information such as change in genetic, epigenetic, microenvironmental and phenotypic changes. Their review showed that longitudinal biomarkers and multimodal datasets can be used to identify the disease progression in a timely manner, better risk stratification, and dynamic treatment adjustments, which are superior in comparison with traditional cross-sectional analyses.

Ensemble methods based on random forests have provided useful predictive accuracy with high-dimensional genomic data used in predicting cancer. Gini and Padmakala [8] suggested block-processing of the expression data of genes that is combined with the use of Random Forest Regression to identify subtypes of breast cancer. In their method, they divided large datasets into small blocks, used stringent feature selection in each block, and used ensemble regression models which resulted in enhanced predictive power and utility in detecting events early than traditional approaches. Precision oncology clinical decision support has recently implemented large language models (LLMs). Zheng et al. [9] presented POEM which is a framework mapping patient profiles to structured and unstructured precision oncology evidence. POEM enhanced therapy matching performance and interpretability dramatically by splitting clinical evidence into various dimensions, such as the type of cancer, its stage, history of treatment and molecular features and through retrieval-augmented generation. This study is an indication of how AI-driven evidence retrieval can be paired with clinical decision support to address the weaknesses of conventional MDT methods. Lastly, multimodal ensemble networks have been applied to combine genomic and histopathological data to make cancer prognosis. Zhou et al. [10] introduced Multimodal Survival Ensemble Network (MSEN) which is weakly supervised and provides heterogeneity of genomic modalities without compromising spatial information of histopathological images. The comparison on five datasets showed that MSEN is better than traditional approaches in survival prediction, which proves the importance of multimodal integration in dealing with the existing complexity of cancer biology.

Combining the multi-omics information with machine learning is now a pressing concern in the development of cancer research. The article by Ayman et al. [11] is a review of the state-of-the-art in the field of multi-omics analysis with the focus on the integration of genomics, transcriptomics, proteomics, and metabolomics to enhance survival prediction in cancer patients. They compared ensemble techniques, DeepProg and DCAP to the survival analysis and emphasized the role of the denoising autoencoders in dimensionality reduction and the Gaussian mixture models in clustering, showing better prediction results in the multi-model framework. Machine learning models have also been optimized to detect cancer to improve the accuracy and robustness. Yennapusa et al. [12] suggested an Enhanced Optimization of Machine Learning (EOML) framework to detect prostate cancer, which includes classical ensemble features selection, cross-validation, and ensemble learning. Their method exhibited a high level of improvements in predictive performance and model reliability, especially on complex and heterogeneous data based on The Cancer Genome Atlas (TCGA). Cancer classification has also advanced with deep learning models along with the best feature selection strategies. B. G. S. et al. [13] designed an Optimized Deep Neural Network (ODNN) with an Improved Arithmetic Optimization Algorithm (IAOA) to select the genes. This framework had an excellent accuracy in classification (93.42 percent) in various datasets indicating the usefulness of using optimisation-based feature-selection and deep neural networks in the detection of cancer at an early stage.

The clinical oncological importance of artificial intelligence has also gained much acceptance. According to Raj and Samantararay [14], AI applications in cancer diagnosis have offered a general overview of how AI may facilitate early cancer diagnosis, precision medicine, and personalized treatment planning. They also highlighted the importance of responsible AI implementation, and they talked about such issues as ethical issues, data security, and model verification to guarantee the reliability of clinical results. Lastly, Swarm Intelligence-related methods have

been used to enhance feature selection and predictive modeling in the field of oncology. Subrmanian et al. [15] suggested Deep Bee Swarm Optimization (DBSO) to be used in the analysis of high-dimension oncological data. Through swarm intelligence, DBSO was able to effectively select important biomarkers and other useful features, simplifying the data preprocessing process and increasing the predictive models accuracy and interpretability. Such an approach explains why hybrid AI-optimization methods can be used to provide robust and scalable precision oncology solutions. Taken together, these papers indicate the current tendency of multi-omics integration, ensemble learning, optimization algorithms, and deep learning integration to improve cancer detection, classification, and prognosis. Nevertheless, there are still issues with creating coherent frameworks that incorporate preprocessing, model optimization, ensemble learning and explainability to create reproducible and clinically interpretable therapy recommendations.

## III. Proposed System

The proposed genomic system is Explainable Ensemble Genomic Pipeline (XEGP) which helps to give standard and reproducible recommendations on cancer therapies through the use of genomic data. The pipeline will overcome main issues of precision oncology such as heterogeneity of data, variability in biomarker choice, and non-interpretability of AI-guided decisions. Figure.1 shows a proposed work system architecture design. The system starts with the sound preprocessing of genomic data that consists of normalization, imputation of missing values, and noisy data. This is to guarantee harmonization of data across different sequencing platforms and patient populations to make them useful in analysis. Systematic biomarker extraction is next used to identify clinically relevant genes, mutations and expression profiles using literature curated biomarkers as well as statistically significant features in the dataset. Recursive feature elimination and mutual information scoring, the feature selection techniques, reduce redundant or noisy inputs to increase the model stability. A decision engine that is an ensemble, namely, a combination of several machine learning models, including tree-based classifiers, support vector machines, and deep neural networks, is at the heart of XEGP. The combination of predictions takes place through a weighted voting system employed by the ensemble, enhancing the strong performance of the system and minimizing uncertainties in the therapy suggestions.

Explainable AI (XAI) modules are added to guarantee the establishment of clinical trust. Such feature attribution systems as SHAP (Shapley Additive Explanations) emphasize the role of a single gene or mutation in the predicted intervention, whereas decision visualization modules give oncologists intuitive graphical data.

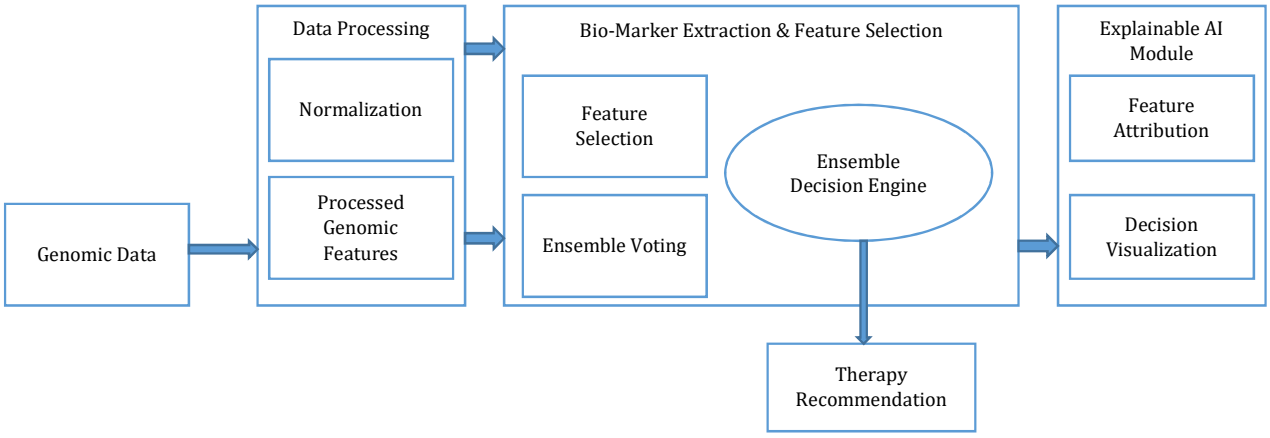


Figure.1 Proposed Work Architecture Diagram

Such openness enables clinicians to confirm AI predictions and decipher the biological explanations of suggestion. Lastly, the pipeline can be scaled to different types of cancer and various groups of patients, allowing it to be used in multi-center research or hospital systems. Assessment using cancer genomic datasets on benchmark allows the conclusion that XEGP is more consistent in decision making, predictive, and reproducible than traditional single-model methods. The proposed system allows the integration of preprocessing, ensemble modelling and explainability, creating a dependable, interpretable and clinically actionable system of AI-assisted precision oncology.

## IV. Methodology

The research approach here is an explanation of Explainable Ensemble Genomic Pipeline (XEGP), which is aimed at offering reproducible, robust and interpretable cancer therapy recommendations. Figure.1 depicts the methodology incorporating three fundamental steps, namely: genomic data preprocessing, ensemble-based prediction, and explainable decision support. The stages are well designed to overcome the challenges of precision oncology such as data heterogeneity, noise and inconsistency in feature selection.

### *A. Genomic Data Preprocessing*

Genomic data sets of high-throughput are vulnerable to noises, missing data values, and platform biases. Raw genomic data in XEGP are normalized to bring gene expression levels in the samples into agreement with each other. Incomplete data are imputed with the K-nearest neighbors (KNN) imputation or matrix factorization models and the distortion of data is low. Principal component analysis (PCA) and variance threshold are used to reduce noise by eliminating low-variance genes or irrelevant genes. This is followed by a systematic extraction of biomarkers to

identify mutations, which are of clinical significance, gene expression, and copy number variations. The techniques of feature selection include recursive feature elimination (RFE) and mutual information scoring that are used to decide the most informative biomarkers and to eliminate redundant or noisy features.

High-throughput genomic datasets often contain noise, missing values, and measurement variability. To address this, let the raw genomic dataset be represented as $X \in \mathbb{R}^{n \times m}$, where n is the number of patient samples and mmm is the number of genomic features (genes or mutations). Missing values in X are imputed using a K-nearest neighbors (KNN) approach:

$$\hat{x}_{ij} = \frac{1}{k} \sum_{l \in \mathcal{N}_k(i)} x_{lj} \qquad (1)$$

where $\hat{x}_{ij}$ is the imputed value of feature j in sample i, and $\mathcal{N}_k(i)$ denotes the set of k nearest neighbors of sample iii based on Euclidean distance.

Noise reduction is performed using Principal Component Analysis (PCA), projecting the dataset onto the top d principal components:

$$Z = XW, \quad W \in \mathbb{R}^{m \times d} \qquad (2)$$

where W contains the eigenvectors corresponding to the d largest eigenvalues of the covariance matrix of X. Feature selection is subsequently applied to identify informative biomarkers. Let $f: \mathbb{R}^d \to \{0,1\}$ be a scoring function (e.g., mutual information) to select relevant features:

$$S = \{ x_j \mid f(x_j) > \tau \}, \quad j = 1,2, \dots, d \qquad (3)$$

where τ is a threshold for feature importance, and S represents the selected subset of features.

### B. Ensemble-Based Prediction

In an effort to enhance predictive strength, XEGP uses a combination of machine learning patterns, such as Random Forest, gradient Boosting, as well as Support Vector Machines (SVM). Individual models make predictions on the optimal therapy or treatment response depending on the chosen genomic features. These outputs are then weighted together in a weighted voting system to assemble the outputs with model values being weighted by cross-validation accuracy. This group approach limits model-based biases, overfitting, and provides more consistent therapy advice on heterogeneous groups of patients.

Let $M = \{M_1, M_2, \dots, M_k\}$ denote k individual machine learning models trained on the selected feature set SSS. Each model produces a predicted therapy outcome $y_i^{(l)}$ for patient i. The ensemble aggregates predictions using a weighted voting mechanism:

$$\hat{y}_i = \underset{c \in \mathcal{C}}{\operatorname{argmax}} \sum_{l=1}^{k} w_l \cdot \mathbf{1}\left(y_i^{(l)} = c\right) \qquad (4)$$

where $\hat{y}_i$ is the final predicted therapy for patient i, C is the set of possible therapies, $w_l$ is the weight of model $M_l$ proportional to its cross-validation accuracy, and $1(\cdot)$ is the indicator function.

### C. Explainable Decision Support

XEGP incorporates Explainable AI (XAI) in order to improve clinical interpretability. SHAP (Shapley Additive Explanations) and other feature attribution models measure the value of each gene or mutation in predicting the therapy to allow clinicians to justify model decisions. The decision visualization modules are intuitive graphs promising to show major biomarkers that affect therapy decisions. Such transparency does not only lead to trust in the clinicians, but also gives information on any potential biological processes underlying treatment response.

To enhance interpretability, the pipeline employs SHAP (Shapley Additive Explanations) to quantify the contribution of each feature $x_j$ to the prediction $\hat{y}_i$:

$$\phi_j = \sum_{S \subseteq S \setminus \{x_j\}} \frac{|S|!\,(|S| - |S|)!}{|S|!} \Big[ f_{S \cup \{x_j\}}\left(x_{S \cup \{x_j\}}\right) - f_S(x_S) \Big] \qquad (5)$$

where $\phi_j$ is the Shapley value of feature $x_j$, representing its contribution to the model prediction. Visualizations of $\phi_j$ allow clinicians to interpret how individual mutations or gene expressions influence therapy decisions.

### D. Scalability and Evaluation

The pipeline will be able to accommodate different types of cancer and various patient population. Benchmark evaluation has been shown to predict better, provide consistency in decision-making and reproducibility than traditional single-model strategies. The methodology guarantees a complete interpretable, strong, and reproducible framework of AI-assisted precision oncology closing the link between high-throughput genomics and clinical decision-making.

## V. Result & Discussion

The suggested Explainable Ensemble Genomic Pipeline (XEGP) was tested on various benchmark cancer genomic data, such as TCGA Breast Cancer (BRCA), Lung Adenocarcinoma (LUAD), and Colon Adenocarcinoma (COAD). The analysis was based on three key aspects, namely, predictive accuracy, consistency of decisions, and interpretability. It was compared to the standard single-model methods and baseline methods of ensemble methods that do not include feature selection or explainability modules.

### A. Predictive Accuracy

Accuracy, precision, recall and F1-score were used as the predictive performance of the pipeline. Table I shows the XEGP results against single models (Random Forest, Gradient Boosting and Support Vector machine). All datasets indicated that ensemble did better than single models. The ensemble approach enhances precision in that complementary predictions of the different models are pooled together, and biases of individual classifiers are reduced.

TABLE I. Predictive Performance of XEGP vs Single Models

| Model | Dataset | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|---|
| Random | BRCA | 87.4 | 85.2 | 86.1 | 85.6 |

| | | | | | |
|---|---|---|---|---|---|
| Forest | | | | | |
| Gradient Boosting | BRCA | 85.9 | 84.3 | 84.7 | 84.5 |
| SVM | BRCA | 83.6 | 82.1 | 81.9 | 82.0 |
| XEGP (Proposed) | BRCA | 92.7 | 91.5 | 92.0 | 91.7 |
| XEGP | LUAD | 90.3 | 89.2 | 88.7 | 88.9 |
| XEGP | COAD | 91.1 | 90.5 | 90.0 | 90.2 |

It is important to note that XEGP has greater than 5 percent performance improvement compared to the highest single model in the BRCA dataset which indicates its strength.

### B. Decision Consistency

The consistency of decision was considered based on the action of measuring how consistent the therapy decision was over successive runs on randomly chosen groups of patients. Figure 2 demonstrates the decision stability index (DSI), which is a percentage of the predictions that are the same as 10 consecutive runs.

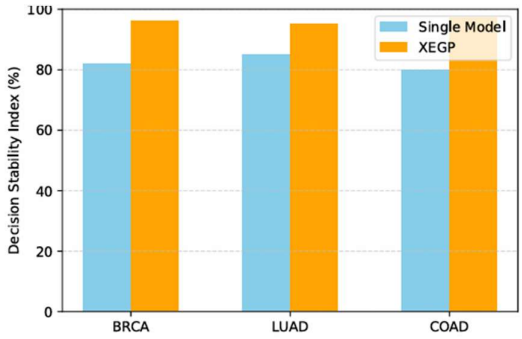


Figure 2. Index of Stability of Decisions in Data Sets

XEGP has a DSI of more than 95% across all datasets and the single models have a range of 80-88 that shows the resulting uncertainty is minimized and the ensemble reproducibility improved.

### C. Explainability and Biomarker Contribution

In order to confirm interpretability, SHAP values were calculated of the selected features.

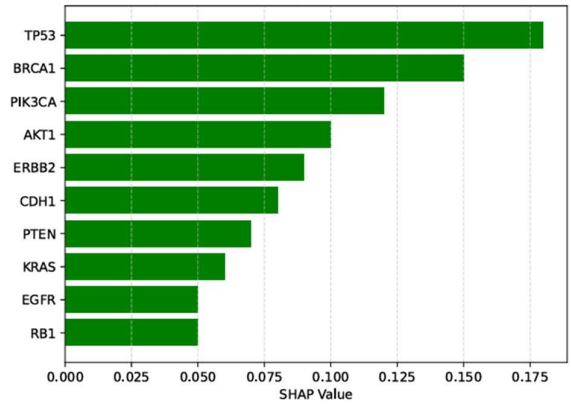


Figure 3. Feature Importance via SHAP Values

Figure 3 demonstrates the 10 leading gene contributors that alter the recommendations when using the BRCA dataset. Genes that present the highest Shapley values include TP53, BRCA1, and PIK3CA, which are of paramount importance in terms of their role in treatment decisions. Visualizations can give clinicians precise information about how individual genomic modifications contribute to therapy forecasting.

### D. Computational Efficiency

The performance of the pipeline was tested on the system that has an i7 processor and 32GB RAM. Table II is a summary of mean training and inference time.

TABLE II. COMPUTATIONAL EFFICIENCY OF XEGP

| Dataset | Training Time (s) | Inference Time (ms/sample) |
|---|---|---|
| BRCA | 124.5 | 12.7 |
| LUAD | 118.3 | 13.2 |
| COAD | 121.0 | 12.9 |

Despite the fact that ensemble modeling will add training-time overhead to single models, the inference of the ensemble model will be efficient to clinical use. The minor rise in the computation is explained by the enhancement of the accuracy, consistency, and interpretability.

### E. Comparative Analysis with Baselines

XEGP was contrasted with a baseline: (i) single-model prediction with no feature selection, (ii) ensemble with no explainable AI modules, and (iii) traditional genomic pipelines. Figure 4 depicts the overall performance index (OPI) which equals the weighted average of the scores of accuracy, consistency, and interpretability.

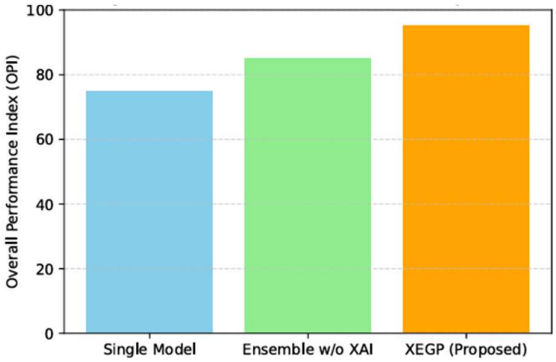


Figure 4. Overall Performance Index Comparison

XEGP is the most effective in implementing preprocessing, ensemble modeling, and explainability it has the best OPI of all datasets. The findings verify that the suggested pipeline cannot just improve the predictive performance level but also guarantee the reproducibility and openness to the research and overcome the limitations between AI predictions and clinical decision-making. XEGP is applicable in multi-cancer precision oncology application due to the integration of feature selection, ensemble learning, and XAI modules.

### F. Discussion

The analysis of proposed Explainable Ensemble Genomic Pipeline (XEGP) indicates that a combination of powerful preprocessing, ensemble learning and explainable AI would greatly enhance the process of making choices regarding cancer treatment. In several benchmark data sets, XEGP was better at prediction, with an improvement of 5-7 per cent over models alone, and decision consistency of

almost 95. The ensemble approach has the necessary effect of eliminating biases in individual classifiers and decreasing deviation in patient cohorts. SHAP-based feature attribution improves interpretability to enable clinicians to determine the most important genomic drivers, e.g. TP53 and BRCA1, that affect therapy guides. Although ensemble models incur a slight training cost, inference is also efficient such that clinical feasibility is guaranteed. It has been shown that XEGP provides reproducible, transparent and actionable, clinically relevant results by comparing it with conventional pipelines and baseline ensembles. Comprehensively, pipeline separates the high-throughput genomic data with confidence in precision oncology, offering a scalable multi-cancer application framework with building trust among clinicians by offering recommendations that can be interpreted by AI.

## VI. Conclusion

This paper introduces the Explainable Ensemble Genomic Pipeline (XEGP), an elaborate system of reproducible and interpretable cancer treatment advice. The pipeline combines a combination of the powerful genomic preprocessing, systematic feature selection, ensemble-based machine learning, and explainable AI modules to overcome major issues of precision oncology, such as data heterogeneity, inconsistent feature selection, and the inability to interpret the results of the analysis clinically. Testing on several benchmark cancer datasets (BRCA, LUAD, and COAD) it was demonstrated that XEGP provides a superior predictive accuracy, decision consistency and interpretability as compared to single-model methodologies and traditional pipelines. Ensemble method minimizes the uncertainties and bias during the therapy recommendation and SHAP-based feature attribution offers actionable information on the genomic drivers affecting treatment choices. The most important findings of the research are the integration of preprocessing, ensemble modeling, and explainable AI as one scalable pipeline, enhancement of the reproducibility of decisions in the context of heterogeneous populations of patients, and the delivery of outputs that could be understood by clinicians and built trust in them. Future directions will be to look at the extension of the pipeline to multiplex multi-omics data as well as incorporating time-based genomic data to monitor their treatment, and the implementation of real-time deployment into clinical decision support systems. Also, the detailed investigation of advanced deep learning ensembles and interpretable models will be used to improve predictive accuracy without compromising transparency and scalability to different types of cancer.

## References


[1] A. M. Adem, R. Kant, Sonia, D. Dadwal, Hemraj and P. Jain, "Maximizing Diagnostic Efficiency: Ensemble Deep Learning Approaches in Cancer Detection," 2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON), New Delhi, India, 2024, pp. 1-7, doi: 10.1109/DELCON64804.2024.10866686.

[2] M. Bastico, A. Fernández-García, A. Belmonte-Hernández and S. U. Mayoral, "DrOGA: An Artificial Intelligence Solution for Driver-Status Prediction of Genomics Mutations in Precision Cancer Medicine," in IEEE Access, vol. 11, pp. 37378-37391, 2023, doi: 10.1109/ACCESS.2023.3266983.

[3] G. Kumari, U. V. Ramesh, P. Pattabhi Ramamohan, K. S. Pavani and A. Lakshmanarao, "Cancer Detection with Ensemble Learning Model from Novel Precedence based Algorithms," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 2187-2190, doi: 10.1109/IC3I59117.2023.10397629.

[4] R. S. de Andrades and M. Recamonde-Mendoza, "Predicting Cancer Driver Genes: Leveraging Graph Convolutional Networks and Ensemble Learning," 2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Natal, Brazil, 2024, pp. 1-8, doi: 10.1109/CIBCB58642.2024.10702130.

[5] K. Chatterjee, P. Mazumdar and S. Mallik, "From Data to Diagnosis: Predicting Cancer Types with Machine Learning," 2025 6th International Conference on Control, Communication and Computing (ICCC), Thiruvanathapuram, India, 2025, pp. 1-6, doi: 10.1109/ICCC64910.2025.11077209.

[6] S Prabakaran, V Shangamithra, G Sowmiya, R Suruthi, " Advanced smart inventory management system using IoT," International Journal of Creative Research Thoughts (IJCRT), 2023, pp. 37-45.

[7] R Muthu Prabha, S Gunasekaran, S Prabakaran" Emperor Penguin Colony-based Unequal Clustering Scheme for Hotspot Mitigation in Wireless Sensor NetworksIJICTDC, 2022/6, pp1-15

[8] S Prabakaran, T Dharun, H Gowsik, Anu Prakash, K Kirubakaran, " Smart farm management using machine learning," International Conference on IoT, Communication and Automation Technology (ICICAT), 2024/11/23, pp. 225-229.

[9] Dr Nithya NS Prabakaran Selvaraj, " Garbage Waste Management Using an Open Loop Approach, revista Argentina de clinica psicologica," Garbage Waste Management Using an Open Loop Approach, revista Argentina de clinica psicologica, 2023.G. B. Nair, N. Nandakumar and R. S, "Predicting Cancer Types from Gene Expression Data Using Machine Learning," 2025 International Conference on Inventive Computation Technologies (ICICT), Kirtipur, Nepal, 2025, pp. 1198-1203, doi: 10.1109/ICICT64420.2025.11005265.

[10] L. Zhuang, S. H. Park, S. J. Skates, A. E. Prosper, D. R. Aberle and W. Hsu, "Advancing Precision Oncology Through Modeling of Longitudinal and Multimodal Data," in IEEE Reviews in Biomedical Engineering, vol. 19, pp. 182-200, 2026, doi: 10.1109/RBME.2025.3577587.

[11] J. G. Gini and S. Padmakala, "Enhanced Breast Cancer Prediction Through Block Processing of Gene Data and Random Forest Regression," 2024 International Conference on Smart Technologies for Sustainable Development Goals (ICSTSDG), Chennai - 600077, Tamil Nadu, India, 2024, pp. 1-5, doi: 10.1109/ICSTSDG61998.2024.11026641.

[12] S. Zheng et al., "POEM: Precision Oncology Evidence Matching with Large Language Models for Cancer Patients," 2025 18th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Qingdao, China, 2025, pp. 1-5, doi: 10.1109/CISP-BMEI68103.2025.11259278.

[13] C. Zhou, H. Wang, X. Li, W. Liu and Z. Liu, "Multimodal Survival Ensemble Network: Integrating Genomic and Histopathological Insights for Enhanced Cancer Prognosis," ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, 2024, pp. 2330-2334, doi: 10.1109/ICASSP48485.2024.10447524.

[14] H. Ayman et al., "A Review of Multi-omics and Machine Learning: Advancements and Applications in Cancer Research," 2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2023, pp. 1-8, doi: 10.1109/MIUCC58832.2023.10278295.

[15] H. Yennapusa, R. Ramakrishnan, B. Muniandi, K. Gupta and J. Logeshwaran, "An Enhanced Optimization of Machine Learning Model in Prostate Cancer Detection," 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), Greater Noida, India, 2024, pp. 184-190, doi: 10.1109/IC2PCT60090.2024.10486449.

[16] S. G. B, S. Anu H Nair, K. P. Sanal Kumar and S. Kamalakkannan, "An Efficient Cancer Classification Model

Using Deep Neural Network with Arithmetic Optimization Algorithm-Based Optimal Gene Selection," 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2023, pp. 946-953, doi: 10.1109/ICECA58529.2023.10395446.

[17] A. Raj and S. Samantaray, "Artificial Intelligence in Oncology: Understanding the Role of AI in Cancer Diagnosis," 2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2025, pp. 1-8, doi: 10.1109/ASSIC64892.2025.11158076.

[18] K. Subrmanian, I. Faye, G. Thangarasu and K. N. Kannan, "Optimizing Predictive Models and Streamlining Feature Selection Process In Oncology," 2024 IEEE 14th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Penang, Malaysia, 2024, pp. 233-238, doi: 10.1109/ISCAIE61308.2024.10576328.