G-Tech : Jurnal Teknologi Terapan
Vol. 8, No. 2, April 2024, pp. 825-834
E-ISSN: 2623-064X | P-ISSN: 2580-8737

Ensemble Learning Approach Reveals Significant Clinical Attributes from Real-World Breast Cancer Cases

Angga Aditya Permana[1], Muhammad Fahrury Romdendine[2]✉

[1, 2] Informatics Study Program, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia

Article Information

Article History

Received : February 21, 2024
Revised : February 29, 2024
Accepted : March 01, 2024

***Keywords*:**

Breast cancer; Ensemble learning; Machine learning; Patient's outcome prediction; Clinical data analysis.

Kata Kunci:

Analisis data klinis; *Ensemble learning*; Kanker payudara; *Machine learning*; Prediksi hasil pasien.

ABSTRACT

Breast cancer has become on of the leading causes of death in Indonesia. This study contributes to global efforts to combat breast cancer by improving patient outcome prediction accuracy. This study employed ensemble learning techniques such as Random Forest, XGBoost, and LightGBM. The results of the study demonstrates LightGBM's superior performance (accuracy=85%, ROC-AUC=81%, AUPR=85%). Notably, all three algorithms identify key clinical attributes: "Relapse Free Status (Months)", "Overall Survival (Months)", "Nottingham Prognostic Index", and "Lymph Nodes Examined Positive". LightGBM uniquely highlights "pam50_LumA" as significant, suggesting reduced fatality risk for Luminal A subtype patients, while others prioritize "Tumor Size". This research lays groundwork for intelligent systems to predict breast cancer outcomes, potentially transforming patient care and clinical practice.

ABSTRAK

Kanker payudara telah menjadi salah satu penyebab kematian utama di Indonesia. Penelitian ini berupaya memberikan kontribusi dalam usaha global untuk melawan kanker payudara melalui peningkatan akurasi prediksi hasil akhir pasien. Penelitian ini menggunakan teknik *ensemble learning* seperti *Random Forest, XGBoost, dan LightGBM*. Hasil percobaan menggunakan data dari atribut klinis pasien menunjukkan *LightGBM* unggul secara performa (akurasi=85%, ROC-AUC=81%, AUPR=85%). Ketiga algoritma berhasil mengidentifikasi beberapa atribut klinis yang signifikan: *"Relapse Free Status (Months)", "Overall Survival (Months)", "Nottingham Prognostic Index", and "Lymph Nodes Examined Positive"*. *LightGBM* secara unik mengidentifikasi "*pam50_LumA*" sebagai atribut paling signifikan yang mengindikasikan terjadinya penurunan resiko fatalitas bagi pasien dengan subtipe kanker Luminal A. Penelitian ini menjadi basis bagi sistem cerdas yang dapat diaplikasikan untuk memprediksi fatalitas pasien berdasarkan atribut klinisnya dan berpotensi mentransformasi upaya perawatan dan praktek terhadap pasien.

***Corresponding Author* :**
Muhammad Fahrury Romdendine
Informatics Study Program, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Indonesia
Jl. Scientia Boulevard Gading, Curug Sangereng, Serpong, Kabupaten Tangerang, Banten 15810
E-mail: muhammad.romdendine@lecturer.umn.ac.id

This is an open access article under the CC BY license

INTRODUCTION

Cancer, a powerful foe that has ravaged civilization for millennia, continues to have a significant negative impact on global health. One of the most common cancers in the world, breast cancer, is a perfect example of the constant difficulties this illness presents (Łukasiewicz et al., 2021). Millions of people and their families' lives are persistently disrupted by breast cancer in all of its forms and stages (Lei et al., 2021). In addition to the difficult burden of treatment, the fight against breast cancer must also address the vital concern of prompt intervention to improve patient outcomes.

The cost of cancer on the world's health is tremendous. It affects people of various ages, ethnicities, and socioeconomic statuses and transcends geographic boundaries (Mustapha et al., 2020). Despite impressive advancements in cancer research, the intricacy of this illness calls for a multifaceted strategy that blends specialized medical knowledge with cutting-edge technology. It is impossible to overestimate the importance of clinical data in this situation for understanding and predicting cancer outcomes (Gray et al., 2018).

A key strategy in the fight against cancer is the identification of important clinical attributes generated from a patient's clinical data (Garrido-Castro et al., 2019). Clinical records contain a wealth of data, including medical history, tumor features, and treatment regimens, which has the potential to reveal patterns that can guide prognostic assessments and treatment choices (Deshmukh & Phalnikar, 2021). Differentiating treatment regimens based on these important characteristics not only increases the likelihood of early intervention and higher survival rates, but also makes it possible (Gordon et al., 2018).

In order to improve patient outcomes, machine learning, especially ensemble learning, has emerged as a transformative force (Hansun, 2021; Tobing et al., 2019). The transparency and robustness of the ensemble learning techniques chosen in this project serve as a foundation. The predictions of various independent models are combined using ensemble learning techniques to provide a prediction that is more reliable and accurate. Healthcare workers may evaluate and comprehend the thinking behind predictions thanks to these models' transparency, which is an essential quality in the context of patient care. Additionally, the reliability of predictions is increased by the robustness of ensemble learning, which makes sure that they are less vulnerable to overfitting and noise (Snider et al., 2023).

Numerous prior investigations have explored the subject matter relevant to our research. In a study conducted by (Rabiei et al., 2022), machine learning techniques were applied to forecast breast cancer malignancy. Their modeling encompassed clinical attributes distinct from those employed in our study. Among the algorithms utilized, Random Forest achieved an 80% accuracy rate, although no precision measurement was reported.

Another study conducted by (González-Castro et al., 2023) demonstrated that the ensemble learning algorithm XGBoost outperformed deep learning in predicting the recurrence of breast cancer in patients, achieving a precision rate of up to 90%. Notably, this study employed clinical attributes that differed from those utilized in our research, and it did not involve the interpretation of the model to identify significant attributes.

Furthermore, in a study carried out by (Zhu et al., 2020), deep learning was employed to predict the prognosis of cancer patients. It is acknowledged that deep learning models lack interpretability as a limitation.

This research study uses an ensemble learning approach to identify major characteristics associated with breast cancer fatality in an effort to improve breast cancer outcomes. The main goal of this research is to better understand the factors influencing breast cancer fatality by utilizing the capabilities of ensemble learning. By doing this, we hope to provide healthcare professionals with insightful information that will aid in decision-making and, ultimately, result in more individualized and efficient treatment plans.

MATERIALS AND METHODS

The Dataset

The data utilized in this study were sourced from a composite of three distinct research investigations, specifically, those conducted by (Rueda et al., 2019). The amalgamated dataset is available for download through the cBioPortal open-source database. Interested readers can access the dataset directly via the following link: https://www.cbioportal.org/study/clinicalData?id=brca_metabric. This dataset comprises clinical information concerning 2,509 cancer patients, including vital information pertaining to their survival status at the culmination of the study duration.

The raw dataset obtained from the aforementioned sources contained a total of 39 attributes. However, not all of these attributes can be considered clinical in nature, necessitating a pre-processing step prior to their utilization in machine learning (ML) modeling. Each attribute underwent manual annotation to assess its relevance for inclusion in the modeling process. This meticulous annotation process involved a comprehensive examination of each attribute, followed by an extensive literature search to determine its clinical relevance.

Among the 39 attributes in the raw dataset sourced from the aforementioned studies, three attributes, namely "Study ID," "Patient ID," and "Sample ID," can be unequivocally categorized as non-clinical, as they serve as identification fields. Another non-clinical attributes identifiend from annotation process are “Number of Samples per Patient” and “Sample Type”. Consequently, these attributes were excluded from consideration for utilization in the subsequent machine learning (ML) modeling stages.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is a pivotal step that precedes advanced data processing for machine learning modeling (Chakri et al., 2023). Its primary purpose is to thoroughly investigate each of the previously selected attributes, as described earlier. Through EDA, insights are gained into the characteristics of these attributes, enabling informed decisions in subsequent preprocessing and feature engineering stages.

The target attribute, “Patient’s Vital Status,” is a categorical attribute, classifying it as a case of classification. It encompasses three distinct value types: “Living” signifies that the patient remained alive throughout the data collection period, “Died of Disease” indicates the patient’s demise due to cancer, and “Died of Other Causes” denotes that the patient passed away from reasons unrelated to cancer.

To ensure that the modeling becomes a binary classification case with a focus on cancer fatality, it is necessary to simplify the target attribute values (hereafter referred to as classes). This simplification involves reducing the classes to just two: the "Living" class and the "Died of Other Causes" class will be transformed into a negative class denoted as "0," while the "Died of Disease" class will be transformed into a positive class denoted as "1." The target attribute with these new class values will then be referred to as the "Died of Cancer" attribute, where a value of "0" signifies "no," and a value of "1" indicates otherwise.

The next EDA involves inspecting the data type of each attribute. Among the attributes, there were eight categorized as numeric, while fourteen were categorized as categorical. Notably, two attributes, namely "Neoplasm Histologic Grade" and "Tumor Stage," initially had numeric values but were, in fact, categorical because they represented ordinal measurements. These two attributes were subsequently manually converted to the categorical data type to accurately reflect their nature. The subsequent step in EDA involves examining potential correlations among the numeric attributes. Correlation map of all numerical attributes is presented in Figure 1.

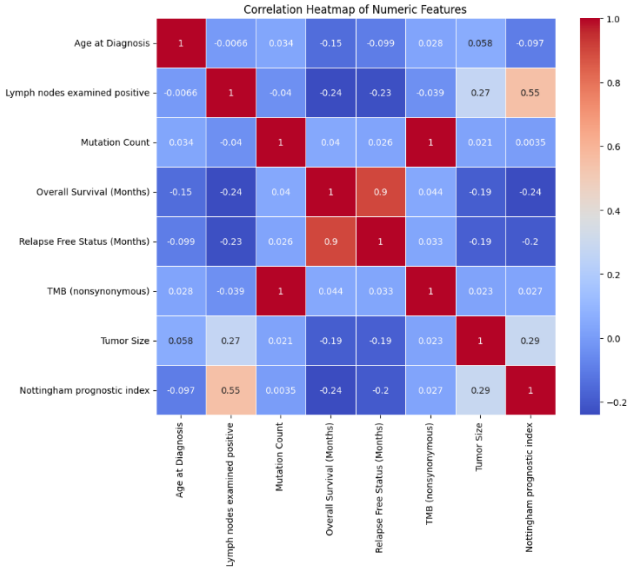


Figure 1. Correlation heat-map of numerical attributes

Feature Engineering

Feature engineering encompasses operations that can augment or diminish characteristics within the dataset, in addition to modifying the values contained therein. Within the scope of this investigation, feature engineering activities consist of one-hot encoding applied to categorical attributes and standardization applied to numerical attributes. Data splitting is conducted prior to feature engineering to ensure no data leakage happen as it would ensuring data integrity and trustworthiness (Dong, 2022).

One-hot encoding represents a technique employed for converting categorical attributes into numerical, specifically binary, form. It serves as a widely utilized approach in the realm of feature engineering. The rationale behind employing one-hot encoding lies in the enhanced learning capabilities of machine learning models when presented with data encoded in numerical format. This technique effectively transforms the categorical value associated with an attribute into a numerical representation, thus facilitating the integration of such data into machine learning models.

Subsequently, a value transformation process is executed on numerical attributes through the utilization of the z-score-based standardization method. This practice is undertaken in recognition of the fact that not all measurements pertaining to numerical attributes are acquired using consistent units. Standardization is employed to remap the values associated with each numeric attribute, ensuring they fall within the interval of -1 to 1. By doing so, standardization nullifies the influence of unit disparities, rendering the resultant model more resilient and robust in its predictive capabilities.

Machine Learning Modelling

This study employs three distinct ensemble learning algorithms, namely Random Forest (RF) (Breiman, 2001), Extreme Gradient Boosting Machine (XGBoost) (Chen & Guestrin, 2016), and Light Gradient Boosting Machine (LightGBM) (Ke et al., 2017). RF, XGBoost, and LightGBM fall under the category of ensemble learning, where RF constitutes a bagging ensemble, while XGBoost and LightGBM operate as boosting ensembles. All three algorithms exhibit proficiency in managing non-linear data, particularly clinical data, and exhibit resilience against overfitting. Additionally, they offer valuable feature importance insights, which serve to elucidate the influence of attributes on the predictive capacity of the model, or in simpler terms, the significance of an attribute in determining patient outcomes.

Transparency in the context of identifying clinically significant attributes holds paramount importance, as the primary objective of this study revolves around suggesting suitable treatments

for patients who may be at risk of experiencing adverse outcomes based on their clinical records. In this regard, a non-transparent model, such as deep learning, would prove impractical. The reason lies in the inherent opacity of deep learning models, which impedes our ability to comprehend the underlying mechanisms by which clinical attributes are taken into account when predicting patient outcomes.

The process of model training will be bifurcated into two distinct phases: the initial stage, referred to as "vanilla modeling," and the subsequent "fine-tuning" phase. Vanilla modeling entails the construction of models using default hyperparameters, devoid of any additional adjustments. This vanilla modeling serves a dual purpose: first, it establishes a foundational model against which the efficacy of subsequent fine-tuning efforts can be gauged, and second, it enables the assessment of whether fine-tuning endeavors yield improvements in model performance (Ghazal et al., 2022).

The fine-tuning of the model is executed through a randomized search methodology involving various combinations of hyperparameters, all encompassed within a cross-validation framework. This fine-tuning procedure was conducted using the RandomizedSearchCV function, a component of the Scikit-Learn package in the Python programming language.

Upon ascertaining the optimal hyperparameters through the aforementioned method, the model is subsequently retrained utilizing these optimized hyperparameters. Following this retraining phase, the model's performance is evaluated using a designated test dataset. The predictive outcomes are then juxtaposed with the actual class labels, allowing for a comprehensive assessment of each algorithm's model performance, encompassing various performance metrics.

Model Evaluation

The model evaluation process in this study encompasses six widely-used classification metrics: accuracy, precision, recall, F1-Score, ROC-AUC Score, and AUPR-Score. These metrics collectively provide a comprehensive assessment of the model's performance. Accuracy measures the overall correctness of predictions, while precision quantifies the proportion of true positive predictions among positive predictions. Recall calculates the proportion of true positives among all actual positive instances. F1-Score balances precision and recall into a single metric. ROC-AUC Score evaluates the model's ability to differentiate between classes using the ROC curve, and AUPR-Score assesses performance across various classification thresholds through the Precision-Recall curve. Additionally, visual aids such as the ROC curve and PR Curve are employed to depict model performance at different thresholds, enhancing the interpretability of results.

The evaluation process hinges on the utilization of model-generated predictions when confronted with the test dataset, facilitating the computation of the aforementioned metrics. Notably, a pivotal focal point is placed on precision. In the context of predicting patient outcomes, the paramount objective is to mitigate the occurrence of false positives, which could potentially incur substantial costs (Roberts et al., 2023). Therefore, our primary criterion for discerning the best model lies in achieving the highest precision value. This strategic emphasis on precision ensures that the model selected not only exhibits robust overall performance but also excels in minimizing the false positive rate, aligning with the economic considerations inherent to patient outcome prediction.

Model Interpretation

The objective of model interpretation is to uncover the specific clinical attributes that exert an impact on the outcomes of breast cancer patients. This interpretation process leverages the feature importance scores generated by each model, which quantifies the influence of each attribute on the model's predictive capacity. It's important to note that these scores don't provide a definitive indication of whether an attribute has a positive or negative effect on the outcome class. To address this uncertainty, an additional analysis, known as the Shapley additive explanation (SHAP) analysis, was conducted.

SHAP analysis is a technique employed in machine learning to interpret the predictions generated by models. This approach is model-agnostic, meaning it can be applied to various machine learning models, and it offers a systematic means of comprehending the individual

contribution of each attribute to the ultimate prediction made by the model (Silva et al., 2022). SHAP analysis provides valuable insights into the factors influencing model predictions, aiding in their understanding and trustworthiness. SHAP values are computed for each attribute within a dataset, and they quantify the influence of that particular attribute on the model's prediction relative to the average prediction made by the model. These values help illuminate how individual attributes contribute to the deviation from the model's mean prediction, facilitating a deeper understanding of the model's behavior and predictions (Stienstra et al., 2023).

RESULTS AND DISCUSSION

Model Performances

The outcomes of the baseline model evaluation during the vanilla modeling phase are detailed in Table 1. The results indicate that all algorithms achieved respectable performance, although none reached exceptional levels. Notably, XGBoost outperformed the other algorithms across all metrics, showcasing its superior performance. Furthermore, to provide a visual representation of the model's discriminatory abilities, the ROC-AUC curve and PR curve are depicted in Figure 2, offering a graphical perspective of the model's performance characteristics.

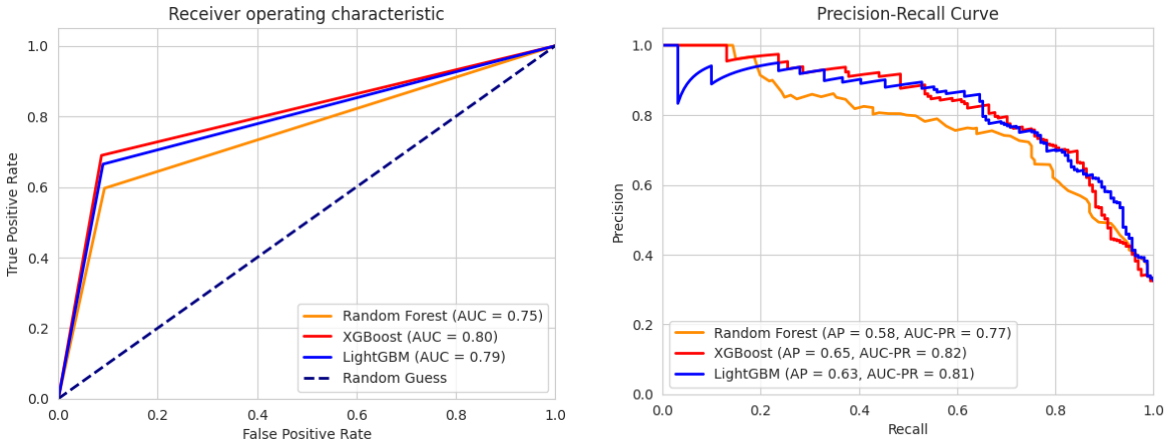
The fine-tuning process led to the identification of the optimal hyperparameters for each algorithm. Subsequently, utilizing these best-fit hyperparameters, the models for each algorithm underwent retraining and subsequent re-evaluation. The evaluation outcomes for the fine-tuned models are meticulously documented in Table 2. To complement these quantitative assessments and provide a visual representation of the improved model performance, the ROC-AUC curve and PR curve for the fine-tuned models are elegantly illustrated in Figure 3, offering a graphical portrayal of their enhanced discriminative capabilities.

Table 1. Baseline model results

Metrics	RF	XGBoost	LightGBM
Accuracy	0.81	0.84	0.83
Precision	0.76	0.79	0.78
Recall	0.60	0.69	0.66
F1-Score	0.67	0.74	0.72
ROC-AUC Score	0.75	0.80	0.79
AUPR Score	0.77	0.82	0.81

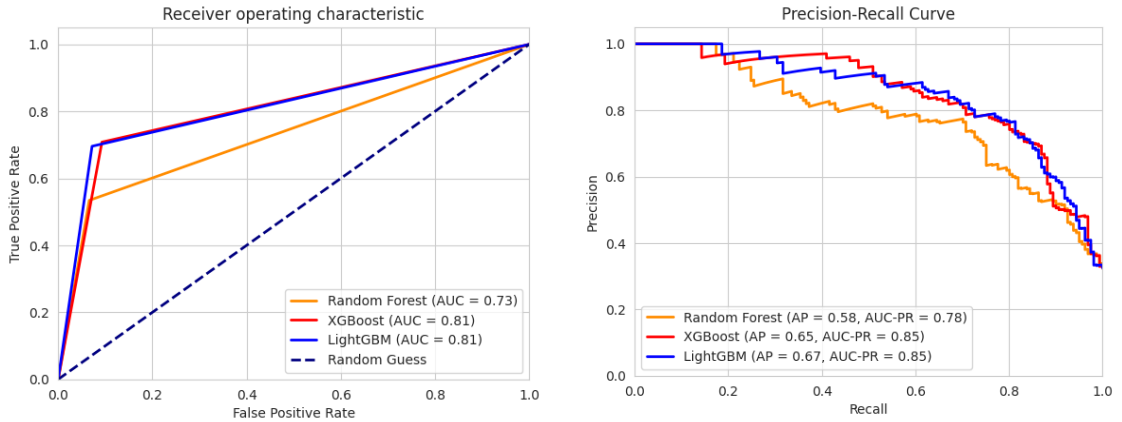
Table 2. Fine tuned model results (↓ indicating decreased performance and ↑ indicating increased performance)

Metrics	RF	XGBoost	LightGBM
Accuracy	0.80 ↓	0.84	0.85 ↑
Precision	0.80 ↑	0.79	0.82 ↑
Recall	0.53 ↓	0.71 ↑	0.70 ↑
F1-Score	0.64 ↓	0.75 ↑	0.75 ↑
ROC-AUC Score	0.73 ↓	0.81 ↑	0.81 ↑
AUPR Score	0.78 ↑	0.85 ↑	0.85 ↑



(a) (b)

Figure 2. ROC-AUC Curve (a) and PR Curve (b) resulted from baseline model evaluation



(a) (b)

Figure 3. ROC-AUC curve (a) and PR curve (b) resulted from fine tuned model evaluation

Model Interpretation and Clinal Attributes Analysis

The identification of significant attributes was accomplished by scrutinizing the outcomes of the SHAP analysis, a procedure systematically conducted for all models. In Figure 4a, we depict the results of the SHAP analysis specific to the fine-tuned RF model. The interpretation of this model through SHAP analysis reveals that the attribute "Relapse Free Status (Months)" attains the highest SHAP value, signifying its preeminent influence on the target attribute. Following closely in terms of importance, the attributes "Overall Survival (Months)," "Nottingham prognostic index," "Lymph nodes examined positive," and "Tumor Size" occupy the second, third, fourth, and fifth positions, respectively, in their impact on the predictive outcome.

The SHAP analysis outcomes for the fine-tuned XGBoost model are showcased in Figure 4b. Notably, the fine-tuned XGBoost model mirrors the RF model by ranking the first five most significant attributes in precisely the same order. Intriguingly, the fine-tuned XGBoost model demonstrates an overall improvement in performance compared to its pre-tuned counterpart, whereas the RF model exhibits a divergent trend, indicating a potential sensitivity to hyperparameter adjustments.

The SHAP analysis findings for the fine-tuned LightGBM model are showcased in Figure 4c. LightGBM aligns with both XGBoost and RF by ranking the first four most significant clinical attributes in the same order. However, LightGBM diverges in its fifth-ranking attribute, designating "pam50_LumB" as the fifth most influential feature. It's noteworthy that LightGBM achieved the highest precision, reaching an impressive 82%, after undergoing the fine-tuning process. This achievement underscores the effectiveness of hyperparameter optimization in

enhancing model performance and the unique strengths of the LightGBM algorithm in this specific predictive task.

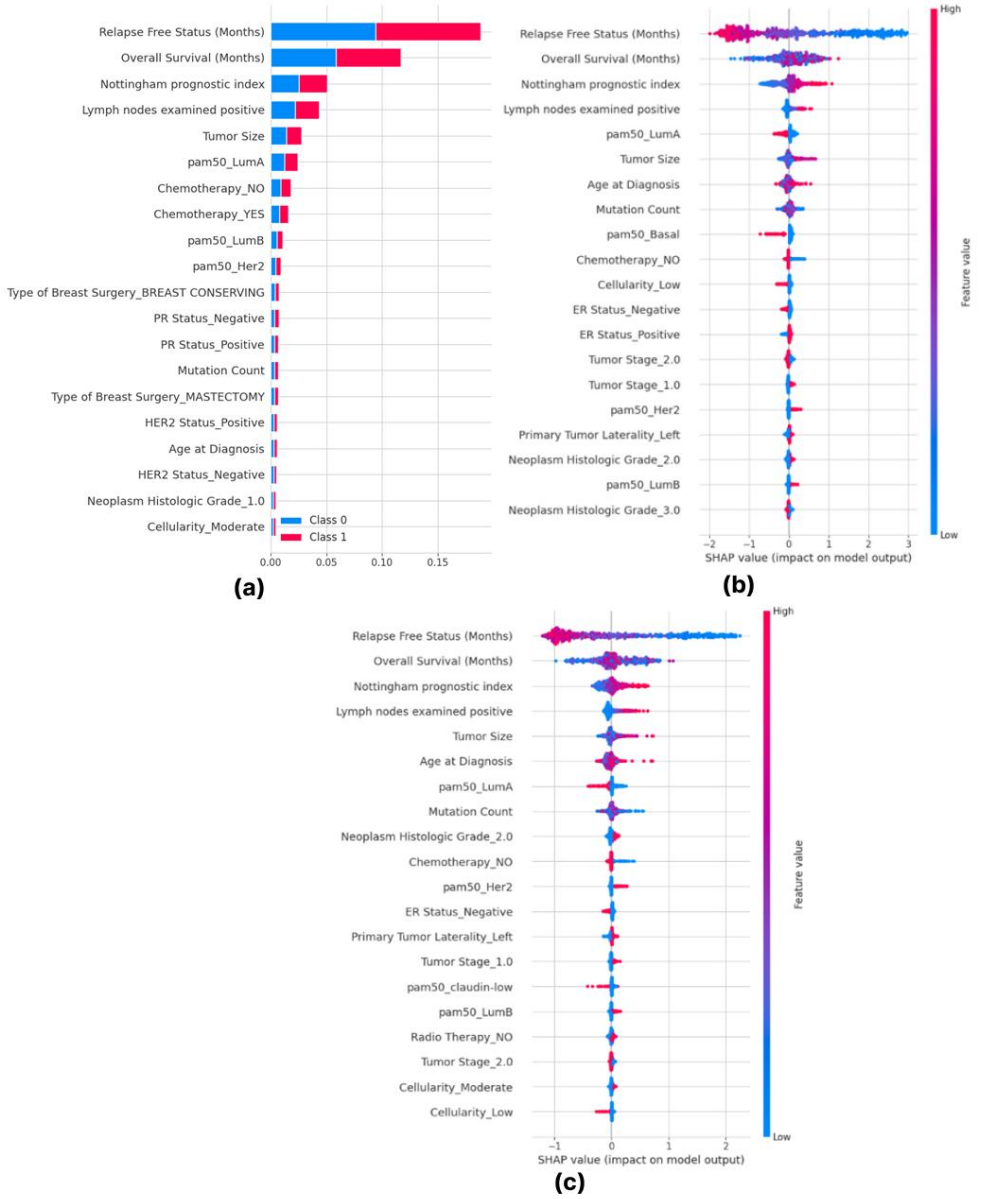


Figure 4. SHAP summary plot from RF (a), XGBoost (b), and LightGBM (c).

CONCLUSION

The results of this study underscore the pivotal role of ensemble learning in uncovering critical clinical attributes influencing breast cancer patient outcomes, with the potential to revolutionize patient care and reduce fatality rates. Notably, when examining the top five influential attributes, all three algorithms concurred on the same order, with the exception of LightGBM, which placed "pam50_LumA" in fifth position instead of "Tumor Size". Significantly, LightGBM, boasting the highest precision, highlighted that patients exhibiting the Luminal A phenotype or subtype tend to experience a lower fatality rate. In conclusion, the findings of this study strongly advocate for the effectiveness of ensemble learning, particularly in the form of boosting ensemble learning, as a powerful approach for identifying clinically significant attributes in the realm of breast cancer patient prognosis, offering valuable insights for improved healthcare management.

ACKNOWLEDGEMENTS

The authors would like to thank Universitas Multimedia Nusantara for the support of this research work.

REFERENCES

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Chakri, P., Pratap, S., Lakshay, & Gouda, S. K. (2023). An exploratory data analysis approach for analyzing financial accounting data using machine learning. *Decision Analytics Journal*, *7*, 100212. https://doi.org/https://doi.org/10.1016/j.dajour.2023.100212

Chen, T., & Guestrin, C. (2016). XGBoost. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Deshmukh, P. R., & Phalnikar, R. (2021). Information extraction for prognostic stage prediction from breast cancer medical records using NLP and ML. *Medical & Biological Engineering & Computing*, *59*(9), 1751–1772. https://doi.org/10.1007/s11517-021-02399-7

Dong, Q. (2022). Leakage Prediction in Machine Learning Models When Using Data from Sports Wearable Sensors. *Computational Intelligence and Neuroscience*, *2022*, 1–9. https://doi.org/10.1155/2022/5314671

Garrido-Castro, A. C., Lin, N. U., & Polyak, K. (2019). Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discovery*, *9*(2), 176–198. https://doi.org/10.1158/2159-8290.CD-18-1177

Ghazal, T. M., Rehman, A. U., Saleem, M., Ahmad, M., Ahmad, S., & Mehmood, F. (2022). Intelligent Model to Predict Early Liver Disease using Machine Learning Technique. *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, 1–5. https://doi.org/10.1109/ICBATS54253.2022.9758929

González-Castro, L., Chávez, M., Duflot, P., Bleret, V., Martin, A. G., Zobel, M., Nateqi, J., Lin, S., Pazos-Arias, J. J., Del Fiol, G., & López-Nores, M. (2023). Machine Learning Algorithms to Predict Breast Cancer Recurrence Using Structured and Unstructured Sources from Electronic Health Records. *Cancers*, *15*(10), 2741. https://doi.org/10.3390/cancers15102741

Gordon, M. N., Cha, K. H., Chan, H.-P., Cohan, R. H., Caoili, E. M., Paramagul, C., Alva, A., Weizer, A. Z., & Hadjiiski, L. M. (2018). Bladder cancer treatment response assessment with radiomic, clinical, and radiologist semantic features. In K. Mori & N. Petrick (Eds.), *Medical Imaging 2018: Computer-Aided Diagnosis* (p. 69). SPIE. https://doi.org/10.1117/12.2294951

Gray, E., Marti, J., Brewster, D. H., Wyatt, J. C., & Hall, P. S. (2018). Independent validation of the PREDICT breast cancer prognosis prediction tool in 45,789 patients using Scottish Cancer Registry data. *British Journal of Cancer*, *119*(7), 808–814. https://doi.org/10.1038/s41416-018-0256-x

Hansun, S. (2021). Medical Informatics in Indonesia: Importance, Development, and Future Directions. *TEM Journal*, *10*(1), 78–81. https://doi.org/10.18421/TEM101-10

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9

eb6b76fa-Paper.pdf

Lei, S., Zheng, R., Zhang, S., Chen, R., Wang, S., Sun, K., Zeng, H., Wei, W., & He, J. (2021). Breast cancer incidence and mortality in women in China: temporal trends and projections to 2030. *Cancer Biology and Medicine*, *18*(3), 900–909. https://doi.org/10.20892/j.issn.2095-3941.2020.0523

Łukasiewicz, S., Czeczelewski, M., Forma, A., Baj, J., Sitarz, R., & Stanisławek, A. (2021). Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review. *Cancers*, *13*(17), 4287. https://doi.org/10.3390/cancers13174287

Mustapha, M., Ali-Gombe, M., Abdullahi, A., Adenipekun, A., & Campbell, O. (2020). Financial burden of cancer on patients treated at a tertiary health facility in South West Nigeria. *Journal of West African College of Surgeons*, *10*(4), 23. https://doi.org/10.4103/jwas.jwas_4_22

Rabiei, R., Ayyoubzadeh, S. M., Sohrabei, S., Esmaeili, M., & Atashi, A. (2022). Prediction of Breast Cancer using Machine Learning Approaches. *Journal of Biomedical Physics and Engineering*, *12*(3), 297–308. https://doi.org/10.31661/jbpe.v0i0.2109-1403

Roberts, T. J., McGuire, J., Temel, J. S., Kamdar, M., Greer, J. A., & Mulvey, T. M. (2023). Developing machine learning algorithms incorporating patient reported outcome measures to predict acute care events among patients with cancer. *Journal of Clinical Oncology*, *41*(16_suppl), 1516–1516. https://doi.org/10.1200/JCO.2023.41.16_suppl.1516

Rueda, O. M., Sammut, S.-J., Seoane, J. A., Chin, S.-F., Caswell-Jin, J. L., Callari, M., Batra, R., Pereira, B., Bruna, A., Ali, H. R., Provenzano, E., Liu, B., Parisien, M., Gillett, C., McKinney, S., Green, A. R., Murphy, L., Purushotham, A., Ellis, I. O., … Curtis, C. (2019). Dynamics of breast-cancer relapse reveal late-recurring ER-positive genomic subgroups. *Nature*, *567*(7748), 399–404. https://doi.org/10.1038/s41586-019-1007-8

Silva, S. J., Keller, C. A., & Hardin, J. (2022). Using an Explainable Machine Learning Approach to Characterize Earth System Model Errors: Application of SHAP Analysis to Modeling Lightning Flash Occurrence. *Journal of Advances in Modeling Earth Systems*, *14*(4). https://doi.org/10.1029/2021MS002881

Snider, E. J., Hernandez-Torres, S. I., & Hennessey, R. (2023). Using Ultrasound Image Augmentation and Ensemble Predictions to Prevent Machine-Learning Model Overfitting. *Diagnostics*, *13*(3), 417. https://doi.org/10.3390/diagnostics13030417

Stienstra, C. M. K., Ieritano, C., Haack, A., & Hopkins, W. S. (2023). Bridging the Gap between Differential Mobility, Log S , and Log P Using Machine Learning and SHAP Analysis. *Analytical Chemistry*, *95*(27), 10309–10321. https://doi.org/10.1021/acs.analchem.3c00921

Tobing, F. A. T., Dzulhaq, M. I., Overbeek, M. V., & Setiawan, E. A. (2019). Design of Decision Support System with C4.5 Algorithm Methods in Determining Results Vaname Shrimp Cultivation. *2019 5th International Conference on New Media Studies (CONMEDIA)*, 195–200. https://doi.org/10.1109/CONMEDIA46929.2019.8981848

Zhu, W., Xie, L., Han, J., & Guo, X. (2020). The Application of Deep Learning in Cancer Prognosis Prediction. *Cancers*, *12*(3), 603. https://doi.org/10.3390/cancers12030603