Comparative Analysis of Optimization Techniques for Enhancing Machine Learning Model Performance

Sumit Kushwaha*, Anish Kumar, Akarsh Kumar Mishra, Ansh Thakur

Department of Computer Applications, University Institute of Computing, Chandigarh University, Mohali, India

*Corresponding author: Sumit Kushwaha, sumit.kushwaha1@gmail.com

Abstract

Optimization methods are crucial in machine learning, and significantly improve predictive model performance. In the age of highly complex machine learning models, scalable optimization methods that can enhance optimal accuracy, computational speed and improve overall model efficiency, have been in great demand. In this paper, comparative study is based on the selection of the optimization techniques, which are vital in multiple phases of the machine learning workflow. The paper is mainly concerned with testing preprocessing methods, feature selection, dimension reduction, regularization and ensemble learning for supervised learning. Data Rescaling methods are compared for the potential optimization of model convergence and performance. In addition, feature selection methods such as SelectKBest and RFE are investigated for the integrity of the formant space with reduced dimensions. This paper also investigates the application of PCA as dimensionality reduction technique and presents its capability of preserving the important data variance. This library also teaches us about regularization techniques like L1 (Lasso) and hyperparameter tuning using GridSearchCV, and how they help in preventing the model from being overfitted and models performance optimisation. Last but not least, ensemble learning with approaches such as Voting Classifiers is explored to show how multiple models can be combined to obtain better predictive power and stability. Thus, the contrasting study here aims to offer useful guidance for academicians and practitioners to choose the best optimization methods for the targeted ML applications.

Keywords

Optimization Technique, Hyperparameter, Feature Selection, Dimensionality Reduction, Gridsearchcv, Ensemble Methods, Regularization, Sustainable Development Goals (SDGs)

1. Introduction

Optimization techniques can be termed as the core of machine learning and data-driven modeling, playing a major role in improving predictive accuracy, reducing computational cost, and enhancing the overall efficiency of algorithms or models [1]. As machine learning models become increasingly complex day by day, the importance of applying effective optimization techniques throughout the whole model development process—from data preprocessing to model tuning—has grown substantially. This research paper shows a comparative analysis of various optimization techniques commonly used in supervised learning tasks, with a specific focus on how they impact model performance and decision-making quality [2]. The objective of this study or research paper is to evaluate and compare the effectiveness of a diverse set of optimization approaches across several stages of the machine learning workflow [3]. The analysis covers data scaling techniques such as Standardization and Min-Max Normalization, which aim to normalize feature distributions, improve convergence during model training and increase efficiency [4]. This paper also includes the feature selection methods like SelectKBest and Recursive Feature Elimination (RFE), particularly designed to identify which is the most relevant attributes from high-dimensional datasets according to the requirement to check the effectiveness and change in accuracy on the basis of attributes. Adding to this, Principal Component Analysis (PCA) is used for dimensionality reduction [5]. It is explored for its ability to condense input space while preserving critical information. Beyond preprocessing, the study investigates regularization techniques such as L1 (Lasso) regularization, which helps decreasing overfitting by penalizing less important features. GridSearchCV for hyperparameter tuning is employed to systematically search for the best configuration of model parameters. The power of ensemble learning is also examined through a Voting Classifier that integrates predictions from Logistic Regression, Support Vector Machine (SVM) and Random Forest, showcasing the potential benefits of model combination strategies [6]. For each of these optimization techniques, there is a consistent machine learning algorithm for implementation, i.e. Logistic Regression—to provide a controlled basis for comparison. The experiments are created to isolate the effect of each optimization method on classification accuracy, highlighting their individual strengths and limitations [7]. The aim is to offer insights into which optimization strategies are most effective under typical modeling conditions, guiding practitioners in selecting appropriate techniques for different machine learning problems. This comparative study contributes to the growing field of model optimization by analyzing and benchmarking widely-used techniques [8]. In the modern time, where datasets can be large, noisy and imbalanced, optimization techniques ensures that the learning

process remains stable, scalable and explainable. As the demand growing for accurate and efficient machine learning solutions in fields like healthcare, finance and automation, optimization is no more a supporting element but it is a basic requirement. Thus, understanding the contribution of each of these techniques, both in isolation and in combination, is critical to building robust and accurate models [9]. The findings underscore the importance of a well-designed optimization pipeline in achieving robust, interpretable, and high-performing machine learning models [10].

2. Literature Review

Machine learning techniques gain increasing attention because models require these techniques to improve their performance cells. interpretability, accuracy and scalability. Numerous studies have concentrated their research on evaluation activities [11]. The research has demonstrated the need for optimizing supervised learning through better implementation of optimization methods for improved integration as the growing need for efficient data-driven systems. Optimization techniques became more prevalent between 2020 and 2025 because they play essential roles to enhance machine learning model performance. The methods establish fundamental components that deliver better model interpretability together with efficiency performance and accuracy and scalability metrics [12]. The growth of machine learning implementation throughout multiple domains makes it necessary for researchers to develop better adaptive optimization approaches because the domain expansion occurs. Many research investigations focus on developing optimization methods for supervised learning because enhanced implementation produces more precise model predictions [13]. Recent studies indicate new optimization approaches represent a necessity for data-driven systems because their computational requirements keep increasing. SDG and Adam and RMSprop demonstrate optimized optimization capabilities because they accelerate convergence rate while maintaining stability [14]. Science professionals use combination approaches with metaheuristic methods that have genetic algorithms and particle swarm optimization to solve complex high-dimensional problems which standard approaches fail to handle. Research shows that adaptive optimization strategies continue development for operation with various data structures in various learning applications. Recent research has demonstrated the creation of adaptive intelligent optimization systems which satisfy current needs of artificial intelligence applications and engineering systems [15].

The importance of feature selection continues as the essential focus point because it allows researchers to retain vital information. It is the most useful optimization strategy. These filter-based methods prove effective for model optimization. The accuracy of models improves while their complexity decreases when using feature selection techniques for datasets containing high dimensions whose features increase difficulty. The computational costs of wrapper methods surpass filter-based methods to achieve better accuracy in identifying the best feature subsets with performance evaluations of model predictions. The embedded method LASSO holds widespread use for its dual function of performing feature selection and regularization activities [16]. A critical preservation of important features leads to better model understanding and operational efficiency. The recent development incorporates feature selection techniques with optimization methods to automatically modify feature sets as training occurs. Using a single strategy produces advantageous results for biomedical, text and image data analysis since it deals with situations that have too many features. Due to research activities adaptive feature selection techniques have become more popular because they generate features suitable for particular application contexts and their associated goals. Feature selection operates as a vital part that creates optimal machine learning pipelines [17].

The Dimensionality Reduction approach PCA together with other techniques provide benefitting performance in data analysis. The approach has gained popularity as a solution to overcome the issue of dimensionality. A comparative study on PCA provides double utility in speeding up training methods while maintaining vital data properties during reduction steps [18]. The technique conserves critical variations in the data that supports predictive accuracy in features with reduced dimensions and spaces. The main constraint of interpretable models exists because of their potential impact on performance. balance between complexity and performance. PCA transforms multiple correlating features into few principal components which retain most of the initial information variances while minimizing data reduction loss. The data reduction achieves faster processing as well as enhanced visualization techniques for complex structures making it suitable for data exploration. PCA delivers effective hidden pattern detection which makes it highly beneficial for image recognition and financial modeling as well as bioinformatics applications. The main disadvantage of this technique consists in the challenge of interpreting transformed data components. The interpretation of principal components remains complicated because their method that combines linear variables complicates understanding the connection between multiple variables and individual components [19]. This reduction method requires the loss of interpretability for fields requiring explainability to experience this trade-off. OR LIME or SHAP are used by researchers to generate explainable insights after applying PCA reduction together with domain expertise. When linear data patterns emerge PCA reaches its most effective state although non-linear data sets require alternative techniques such as t-SNE or UMAP to bring optimal results. PCA when used in combination with other methods enables effective data reduction at a high model performance level. The suitable dimension reduction technique requires assessing both data characteristics and analytical requirements between interpretability and computational speed and analysis clarity [20].

Standardization served with normalization as the core operations for pre-processing model convergence. The authors pointed to standard scaling as a robust approach for model applications Standard scaling contributed to better gradient-based system stability when used with logistic regression and neural networks models [21]. The main successes of Min-

Max scaling emerged from dealing with systems using defined input constraint boundaries input features, such as distance-based algorithms. Model performance alongside convergence speed gets improved effectively when machine learning preprocessing implements standardization and normalization approaches. Standardization becomes the most suitable pre-processing method when dealing with algorithms that depend on Gaussian distributions or gradient descent-based algorithms like logistic regression and neural networks. Standardization makes data follow a mean of zero and standard deviation of one. The transformation maintains equal contribution of each feature to the model so no single feature can dominate the learning process [22]. The transformation achieved through Min-Max scaling creates data compression within the [0, 1] range which suits machine learning algorithms like KNN and SVM because they depend on the input feature scale. Min-Max scaling achieves success by efficiently managing input constraints during distance calculation models since the directly affected output depends on scaled features. The normalization technique shows sensitivity to extreme data points since they modify the measurement scale of data sets. The selection between standardization and normalization needs to consider both the nature of the data and the specifications of the algorithm used in the process. The combination of these two techniques in selective applications results in the best model performance levels and more predictable training procedures [23].

Regularization techniques, the process of development continues using L2 (Ridge) along with L1 (Lasso) regularized approaches Research. The Automatic feature selection occurs in MLRL systems since they achieve optimal outcomes when operating with limited available data. Deep learning algorithms activate regularization benefits when implemented during the implementation phase.The automatic control system integrates a weight decay mechanism which functions with dropout features to serve as built-in optimization features [24]. Machine learning utilizes Regularization as its core method to decrease overfitting problems particularly when handling big data or reduced available samples. L1 (Lasso) regularization successfully creates feature sparsity by making trivial coefficients zero but L2 (Ridge) regularization imposes even weight penalties without eliminating features. Elastic Net combines the advantages of L1 and L2 to provide a solution that works best for situations containing feature correlations. Weight decay and dropout serve as deep learning regularization methods which provide additional strength to these systems. The two regularization strategies combine during backpropagation to reduce model weights (weight decay) while dropout disables random neuron groups to avoid network dependency on specific training paths. These optimization features embedded within the model enhance its robustness when implemented. Multiple regularization strategies must be integrated into modern data environments because the rising complexity drives the need for maintaining prediction accuracy on multiple machine learning frameworks [25].

Machine learning model optimization requires Hyperparameter Tuning for adjusting parameters which training does not learn but directly affect performance levels. Different automated techniques namely Grid Search, Random Search, and Bayesian Optimization nowadays help streamline the optimization process in machine learning pipelines. The preferred tool GridSearchCV emerged because it evaluates every potential combination of predefined parameter values through cross-validation that leads to consistent performance results. The computation needs for this method become extensive when handling both extensive datasets and sophisticated models [26]. The method proves effective for high-performance computing needs together with applications requiring small-scale problems. Random Search provides speed through random parameter selection however it requires proper construction of search space to avoid improper configuration identification. The combination of probabilistic models in Bayesian Optimization enables it to identify promising parameter values based on previous outcomes which decreases the required number of necessary trials. Stratified sampling techniques used on small to medium datasets ensure better reliability when they distribute training dataset samples evenly throughout all folds. An ideal hyperparameter tuning approach emerges from the combination between dataset size and model complexity and available computational resources which requires proper balance between thoroughness and efficiency [27]. The entire process of hyperparameter optimization now runs automatically as automated systems have taken over human tasks during optimization operations techniques like Grid Search, Random Search, and Bayesian Optimization. And particularly, The virtual optimization tool GridSearchCV functions as the most significant method which scientists and professionals use for their work.

There has been a rapid development of Ensemble Methods during the recent several years. Multiple studies demonstrate the effectiveness of stacking multiple classifiers such as Random Forest, SVM and Logistic Regression together in voting-based and stacking model systems. The utmost significance of hybrid optimization strategies which unite various models and tuning platforms becomes more apparent in current times [28]. The combination of multiple base learners for their individual strength characteristics makes ensemble learning one of the dominant performance enhancement approaches. These methods namely bagging boosting and stacking were developed to resolve how single models face issues with variance and bias and generalization [29]. The training process of AdaBoost and Gradient Boosting follows a weak learner sequence to produce predictive models with stronger performance. Through stacking algorithm various mathematical models which include SVMs and neural networks and logistic regression operate under a single meta-classifier. The use of layered design structure enables experts to detect complex data connections that produces enhanced outcomes. Soft-voting ensemble methods in research prove effective particularly in noisy and imbalanced data thus making them suitable for medical and sentiment recognition systems along with fraud identification systems [30].

3. Methodology

This section explains the detailed experiment design and optimization method that would be included in the research to enhance the performance of the machine learning models. The discussed model will aim to assess and examine some of the most crucial techniques such as preprocessing methods, feature selection, dimensional reduction, regularization, hyperparameter tuning, and ensemble methods to establish the best method to develop the best results, as in figure 1. These methods are scientifically compared under controlled aspects to make the results reliable and constant.

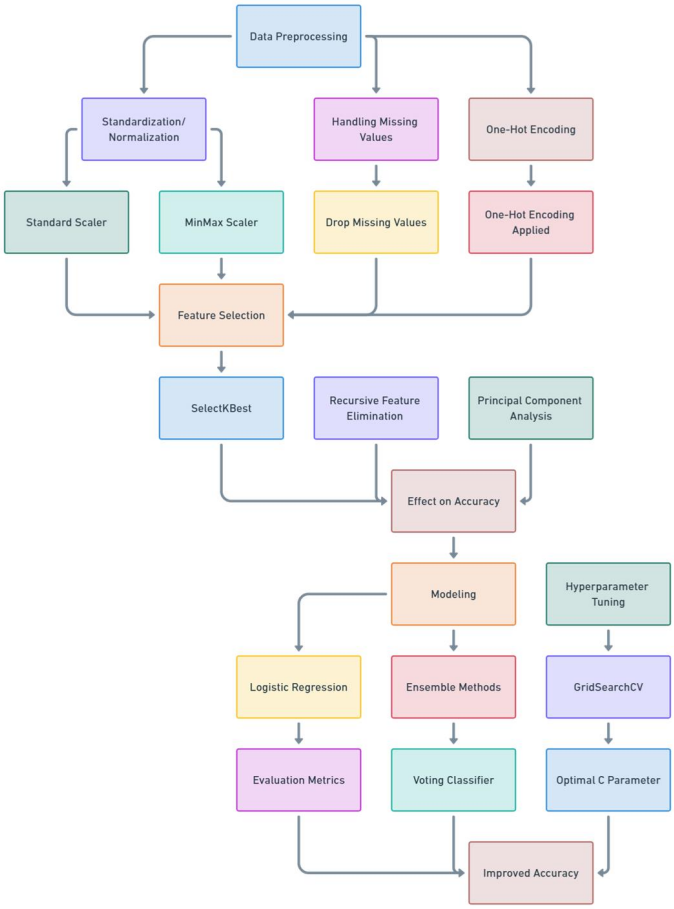

Figure 1. Optimizing machine learning models, including data preprocessing, feature selection, dimensionality reduction, model training, evaluation, ensemble methods, and hyperparameter tuning.

To train the machine learning, data is preprocessed using different kinds of methods. The Heart Disease Train-Test data, which is obtained via Kaggle, has 13 numeric attributes, and one classification label. Then using Pandas, missing values are dropped by using dropna(), and categorical data are coded into numerical data using one-hot encoding method throughget_dummies(). A certain standardization and normalization steps have been used e.g. by applying standard scaler and Min Max scaler because we do not want one of the attributes to overrule the other ones during the training process. The data is subsequently divided into training and test options, these are often in the ratio of 80-20.

In the baseline model, a general Logistic Regression model has been trained to generate reference performance on the clean and scaled dataset. The initial model is evaluated with the use of precision, recall, F1-score, and accuracy metrics. Such metrics can be used to infer any model bias and provide a basis to compare the results of the optimization techniques that will be provided below.

Depending on which step to the model development process, deoptimization techniques are applied. To see the effect that feature scaling has on model performance, scaling methods such as StandardScaler and MinMaxScaler are applied. Scaling balances out all the features to assure equal weighting in training process that reduces the possibility of biased models such as logistic regression. The method (feature selection techniques) such as SelectKBest and Recursive Feature Elimination (RFE) are then applied to select the most important features thus, making the dataset less dimensional. After removing redundant features, the model will be easier to understand and will work more effectively.

Reduction of dimensionality is also done using Principal Component Analysis (PCA) which fits the features by keeping the component which accounts the highest variation. Noise and duplication are minimized by this process and the model becomes able to work more efficiently and maintain essential information. Also, regularisation methods, e.g. L1 regularization, are utilised to avoid overfitting. The issue with this L1 penalty is that it enforces sparsity in the model; i.e., features that are of lesser importance are assigned a zero weight which enhances interpretability in the model as well as its generalization behavior.

The optimal strength of regularization (C parameter) of the Logistic Regression is determined with the use of GridSearchCV. This procedure is complemented by 5-fold cross-validation, so that the model fits well in various data subsamples. An ensemble learning is conducted as well, using a VotingClassifier that uses soft voting-based combination of Logistic Regression, Random Forest, and SVM classifiers. This combination method contributes to model robustness when pooling the strengths of several classifiers that ultimately can lead to a more accurate prediction of the model and an increase in the overall performance of the model. The quality of the models is measured in terms of accuracy, precision, recall, F1- score, as well as AUC-ROC to get a clear picture of the relative performance of each of the methods employed.

4. Result

The findings of the application of some optimization methods to the Logistic Regression model prove that there is a significant modification of the performance of the models in the enhancement stages. The model pretrained on the raw data scored an accuracy of 80%. This was the basis on which comparison of efficiency of different methods applied to optimize the model was made. When the fit method StandardScaler was used to standardize the features, the accuracy rose by a small level to 80.4 %, indicating that scaling contributed to the enhancement of the model performance at a slight level.

A more appreciable change was noticed when we used MinMaxScalerANb E quot d and compared it to Mininea prob contemporaneous to the rise of scaler in the US abolition case law. As in table 1, the feature scaling technique converted the data into an interval of between 0-1 and it produced an 83.9 % accuracy. This revealed that MinMaxScaler was more practical compared to the StandardScaler as it is effective when dealing with data with divergent ranges. The SelectKBest method with the f_classif score function was also an important factor in feature selection in order to focus on the top 10 features. Nevertheless, it caused a drop in the accuracy to 79%, which illustrated that feature selection might cause a loss of important information at times, as in figure 2.

The Recursive Feature Elimination (RFE) algorithm that successively drops less significant features demonstrated a good improvement. The results show that RFE reached 84% accuracy whereas SelectKBest performed at a lower level of 78%, this shows that model specific features selection is crucial. Multidimensionality reduction using Principal Component Analysis (PCA) was another area that helped in improving a model. PCA succeeded with an accuracy of 80.9% by reducing the data to only five principal components which demonstrates that dimensionality reduction can also be used to simplify the model without compromising much of the accuracy. The PCA was effective in explaining the variability of data, as opposed to the direct improvements in predictive ability which, it accomplished.

Table 1. The accuracy, precision, F1-score, and sensitivity of different optimization techniques.

Technique	Accuracy (%)	Precision (%)	F1-Score (%)	Sensitivity (%)
Baseline Model	80	78.5	79.2	77
Standard Scaler	80.4	80.2	80.5	79.1
MinMax Scaler	83.9	82.1	83.3	82
SelectKBest (Feature Selection)	79	75.3	76	74.5
RFE (Feature Selection)	84	84.3	84.5	83
PCA	80.9	80.1	81	79.5
GridSearchCV	81.9	81.2	81.7	80.5
Voting Classifier (Ensemble)	87.8	88.1	88.3	87.5

Further improvement of the model in terms of accuracy was achieved through the use of hyperparameter tuning via GridSearchCV. After tuning the regularization parameter (C value) using the GridSearchCV with Logistic Regression, the accuracy was improved to 81.9%. This demonstrated the challenge of hyperparameter tuning in the improved performance of the model. Finally, the ensemble model that used the VotingClassifier combined Logistic Regression, Random Forest and SVM and produced the highest degree of accuracy at 87.8 %. This combination of classifiers showed superior performance as it build strength out of numerous classifiers, which translated to better generalization and resiliency as compared to the individual classifiers and on the whole, the outcomes depict the important role of scaling, feature selection, dimensionality reduction, hyperparameter tuning, and ensemble learning in the increased accuracy and generalization of Logistic Regression models.

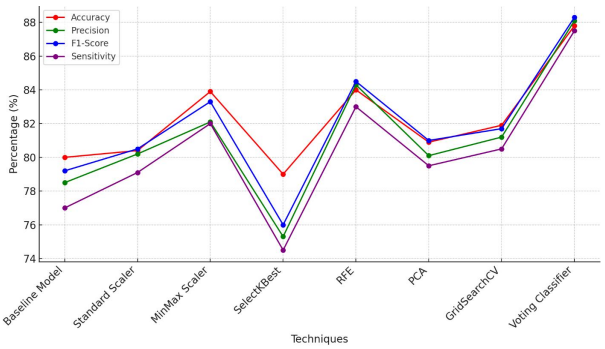

Figure 2. Comparison of the accuracy, precision, F1-score, and sensitivity across different optimization techniques.

5. Discussion

The findings of the optimization procedures used with the models of the Logistic Regression reveal the significance of different phases of promoting model efficiency. A baseline model that reliably performed with an accuracy of 80 % formed a good starting point in evaluating the effects of each optimization measure. Scaling the features, both with StandardScaler and MinMaxScaler was an effective solution to enhancing model performance. At least the minimal improvement in accuracy due to the use of StandardScaler (80.4%) shows that it will standardize the features so that no single variable takes control over learning process. Nonetheless, the more critical enhancement minMaxScaler achieved, which was 83.9, indicates that scaling attributes to zero-to-one range can culminate in improved convergent and effective algorithms, particularly when these algorithms become sensitive to attribute scale.

SelectKBest and Recursive Feature Elimination (RFE) feature selection methods showed poor performance. The SelectKBest, which employed the f_classif score function to select the best 10 features, gave an accuracy of 79 %. This decrease implies that the procedure could have lost crucial attributes prompting minor decline in performance. On the one hand, RFE, as an iterative searching algorithm that sequentially removed less significant features with the references to Logistic Regression coefficients, demonstrated a respectable level of accuracy equal to 84%. It can be assumed that the capacity of RFE to keep the most significant features of the model was a factor that led to model performance improvement. The findings highlight the attention that should be paid to the selection of this or that feature selection method since it may improve or worsen the Accuracy of the model.

The dimensionality reduction through PCA also proved very crucial in optimal modeling. The 80.9 % accuracy level was obtained using the PCA method that downscaled the features to five principal components. Although this was not a large difference, it shows that dimensionality reduction can be useful, especially in large dimensional sets of data, in reducing the complexity of a model without making it lose needed variance. The most significant information was saved with the help of CA, as it is important to deal with the complexity of data but not to decrease the model accuracy. One should however remember that the main idea behind PCA is not to directly optimize prediction but rather to maximize data variance.

The grid search via GridSearchCV also assisted in tuning the hyperparameter to boost the model performance. A further refinement of the regularization parameter (C value) of the Logistic Regression algorithm increased the accuracy to 81.9%. This shows the relevance of having optimal hyperparameters and their importance in the performance of the model. Lastly, the voting algorithm, in particular Voting Classifier, was found to be the most successful optimization algorithm with an accuracy score of 87.8 %. The ensemble model of combining Logistic Regression with Random Forest and SVM deploying a soft voting technique proved to have better generalization and robustness than other individual models. It shows that the use of an ensemble of several different classifiers can result in more stable predictions and contribute to a greater accuracy of the model in general.

6. Conclusion

The research comparisons show models attain their best performance depending on which optimization methods get selected.The decision between preprocessing methods and optimization techniques impact the overall influence on results. While the base Logistic Regression model achieved 80% accuracy during assessment but subsequent methods like All three optimization elements of feature scaling together with hyperparameter tuning and ensemble learning technique respectively produced noteworthy performance improvements.Maximization of performance Metrics occurred through MinMax Scaling which provided significant benefits over other preprocessing methods. techniques, increasing accuracy to 83.9%. Although SelectKBest slightly reduced performance, The model received enhancement through RFE because it used performance-based feature selection. Dimensionality Application of PCA for dimension reduction provides minor improvements that showcase its worth when dealing with high-dimensional problems. dimensional scenarios. GridSearchCV provided hyperparameter tuning which served to enhance accuracy to 81.9% through its approach.The optimal parameter search process led to a 81.9 % accuracy rate. Ultimately,The VotingClassifier ensemble method reached an accuracy level of 87.8% which confirmed its powerful capabilities.Multiple classifiers united together boost both generalization capabilities and resistance capabilities. These findings Multiple approaches work better together rather than each on its own in terms of successful decision making.The customization of methods according to the dataset leads to the discovery of optimal performance results.

References

[1] G. Chugh, S. Kumar, and N. Singh, “Survey on machine learning and deep learning applications in breast cancer diagnosis,” Cognitive Computation, vol. 13, no. 6, pp. 1451–1470, 2021.

[2] Kushwaha, S. (2023). A futuristic perspective on artificial intelligence. In Proceedings of the IEEE OPJU International Technology Conference on Emerging Technologies For Sustainable Development (pp. 1-6). O.P. Jindal University, Raigarh, Chhattisgarh, India.

[3] Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An automated diagnostic system for heart disease prediction based on χ^2 statistical model and optimally configured deep neural network. IEEE Access, 7, 34938–34945.

[4] Gkontzis, A. F., Kotsiantis, S., Panagiotakopoulos, C. T., & Verykios, V. S. (2022). A predictive analytics framework as a countermeasure for attrition of students. Interactive Learning Environments, 30(6), 1028–1043.

[5] Duncker, D.(2021). Smart wearables for cardiac monitoring—real-world use beyond atrial fibrillation. Sensors, 21(7), 2539.

[6] Isakadze, N., & Martin, S. S. (2020). How useful is the smartwatch ECG? Trends in Cardiovascular Medicine, 30(7), 442–448.

[7] Kushwaha, S. (2023). An effective adaptive fuzzy filter for speckle noise reduction. Multimedia Tools and Applications, 2023, 1-16. Springer.
[8] Martins, M. V., Tolledo, D., Machado, J., Baptista, L. M., & Realinho, V. (2021). Early prediction of student's performance in higher education: A case study. In A. Rocha, H. Adeli, G. Dzemyda, F. Moreira, & A. M. R. Correia (Eds.), Trends and applications in information systems and technologies (Vol. 9, pp. 166–175).
[9] Bijari, S., Sayfollahi, S., Mardokh-Rouhani, S., Bijari, S., Moradian, S., Zahiri, Z. (2024). Radiomics and deep features: Robust classification of brain hemorrhages and reproducibility analysis using a 3D autoencoder neural network. Bioengineering, 11, 643.
[10] Thammasiri, D., Delen, D., Meesad, P., & Kasap, N. (2014). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. Expert Systems with Applications, 41(2), 321–330.
[11] Tembhurne, J. V., Hebbar, N., Patil, H. Y., & Diwan, T. (2023). Skin cancer detection using ensemble of machine learning and deep learning techniques. Multimedia Tools and Applications, 16, 1–24.
[12] Wang, C., Zhao, H., & Zhang, H. (2020). Chinese college students have higher anxiety in the new semester of online learning during COVID-19: A machine learning approach. Frontiers in Psychology, 11, 587413.
[13] Li, Y., Li, C., Li, X., Wang, K., Rahaman, M. M., Sun, C., Chen, H., Wu, X., Zhang, H., & Wang, Q. (2022). A comprehensive review of Markov random field and conditional random field approaches in pathology image analysis. Archives of Computational Methods in Engineering, 29(1), 609–639.
[14] Islam, A., Belhaouari, S. B., Rehman, A. U., & Bensmail, H. (2022). KNNOR: An oversampling technique for imbalanced datasets. Applied Soft Computing, 115, 108288.
[15] Bulbul, H. I., Usta, N., & Yildiz, M. (2017). Classification of ECG arrhythmia with machine learning techniques. In 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 546–549). IEEE.
[16] Hamim, T., Benabbou, F., & Sael, N. (2022). Student profile modeling using boosting algorithms. International Journal of Web-Based Learning and Teaching Technologies, 17(5), 1–13.
[17] Sharma, P., & Gupta, D. V. (2018). Disease classification from ECG signal using R-peak analysis with artificial intelligence. International Journal of Signal Processing, Image Processing and Pattern Recognition, 11(3), 29–40.
[18] Ullah, A., Anwar, S. M., Bilal, M., & Mehmood, R. M. (2020). Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation. Remote Sensing, 12(10), 1685.
[19] enpipat, W., & Akkarajitsakul, K. (2020). Student dropout prediction: A KMUTT case study. In 2020 1st International Conference on Big Data Analytics and Practices (IBDAP) (pp. 1–5). IEEE.
[20] Khan, M. U., Aziz, S., Naqvi, S. Z. H., & Rehman, A. (2020). Classification of coronary artery diseases using electrocardiogram signals. In 2020 International Conference on Emerging Trends in Smart Technologies (ICETST). IEEE.
[21] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. Computers & Education, 145, 103724.
[22] Abidalkareem, A., Ibrahim, A. K., Abd, M., Rehman, O., & Zhuang, H. (2024). Identification of gene expression in different stages of breast cancer with machine learning. Cancers (Basel), 16, 1864.
[23] Alshawwa, I. A., El-Mashharawi, H. Q., Salman, F. M., Al-Qumboz, M. N. A., Abunasser, B. S., & Abu-Naser, S. S. (2024). Advancements in early detection of breast cancer: Innovations and future directions. (Publisher and journal info missing – complete if available.)
[24] Bostanci, E., Kocak, E., Unal, M., Guzel, M. S., Acici, K., & Asuroglu, T. (2023). Machine learning analysis of RNA-seq data for diagnostic and prognostic prediction of colon cancer. Sensors, 23, 3080.
[25] Kushwaha, S. (2023). Review on artificial intelligence and human computer interaction. In Proceedings of the IEEE OPJU International Technology Conference on Emerging Technologies For Sustainable Development (pp. 1-6). O.P. Jindal University, Raigarh, Chhattisgarh, India.
[26] Chugh, G., Kumar, S., & Singh, N. (2021). Survey on machine learning and deep learning applications in breast cancer diagnosis. Cognitive Computation, 13, 1451–1470.
[27] Mahboubisarighieh, A., Shahverdi, H., Jafarpoor Nesheli, S., Alipoor Kermani, M., Niknam, M., Torkashvand, M. (2024). Assessing the efficacy of 3D Dual-CycleGAN model for multi-contrast MRI synthesis. Egyptian Journal of Radiology and Nuclear Medicine, 55, 1–12.
[28] Rezaeijo, S. M., Rezaei, M., Poursheikhani, A., Mohammadkhani, S., Goharifar, N., Shayankia, G. (2023). Integrative bioinformatics analysis of miRNA and mRNA expression profiles identified some potential biomarkers for breast cancer. Egyptian Journal of Medical Human Genetics, 24, 62.
[29] Sahu, A., Qazi, S., Raza, K., Singh, A., & Verma, S. (2022). Machine learning-based approach for early diagnosis of breast cancer using biomarkers and gene expression profiles. In Computational Intelligence in Oncology Applications: Diagnosis, Prognosis, and Therapy of Cancers (pp. 285–306). Springer.
[30] Asadi, O., Yekkalam, A., & Manthouri, M. (2023). MDDC: Melanoma detection using discrete wavelet transform and convolutional neural network. Journal of Ambient Intelligence and Humanized Computing, 14(9), 12959–12966.