How to Cite:
Rashi, A., & Madamala, R. (2022). Minimum relevant features to obtain AI explainable system for predicting breast cancer in WDBC. *International Journal of Health Sciences*, *6*(S9), 1312–1326.
https://doi.org/10.53730/ijhs.v6nS9.12538

Minimum relevant features to obtain AI explainable system for predicting breast cancer in WDBC

Agarwal Rashi
Department of Computer Science and Engineering, Harcourt Butler Technical University, Kanpur
Email: dr.rashiagrawal@gmail.com, rashi@hbtu.ac.in

Revanth Madamala
Computer Science Department, University of Southern California, Los Angeles, CA, U.S.A
Email: madamala@usc.edu

***Abstract*---**The potential to explain why a machine learning model produces a certain prediction in incomprehensible terms is becoming increasingly crucial, as it provides accountability and confidence in the algorithm's decision-making process. The interpretation of complex models is difficult. Various approaches to dealing with this issue are being offered.These problems are typically handled in tree ensemble methods by assigning priority levels to input features globally or for a specific prediction. We show that current feature attribution approaches are inconclusive, and develop solutions using SHAP (SHapley Additive Explanation) values, LIME (Local Interpretable Model-Agnostic Explanations), and the Skope Rules package. We employ feature selection methods from SHAP and LIME in this work, which uses the Breast cancer Wisconsin data sets. In the suggested method, features are chosen at the first level of feature selection using Decision tree entropy values. Based on the SHAP and LIME reports, level 2 features are chosen from fewer options. The features are tested on a Decision Tree (DT) model and a DT and Support Vector Machine (SVM) ensemble. Experiments suggest that the ensemble works better as compared to DT. We have also used the Skope Rules package to generate global rules for generalization. A radiologist specializing in mammography has been consulted to verify the SHAP, LIME, and Skope package results. The results of SHAP are most promising.

***Keywords*---**XAI, SHAP, LIME, Skope Rules, feature selection, breast cancer, WDBC, Decision Tree, Ensemble.

International Journal of Health Sciences ISSN 2550-6978 E-ISSN 2550-696X © 2022.
Manuscript submitted: 9 May 2022, Manuscript revised: 18 July 2022, Accepted for publication: 27 August 2022

Introduction

Breast cancer is the second biggest cause of death in women, coming in second only to lung cancer [26]. It is the most common cancer among Indian women, with a prevalence of 25.8 per 100,000 women and a fatality rate of 12.7 per 100,000 women. The incidence and mortality statistics from various recent national cancer registries were compared. Delhi had the highest age-adjusted incidence rate of breast carcinoma at 41 per 100,000 women, followed by Chennai (37.9), Bangalore (34.4), and Thiruvananthapuram District (33.7). In rural registries, the mortality-to-incidence ratio was as high as 66, whereas, in urban registrations, it was as low as 8. In addition, youthful age has been identified as a significant risk factor for breast cancer in Indian women [19]. A pleasant and encouraging clinical picture in the country would result from increased health awareness and the availability of breast cancer screening programs and treatment facilities.

The use of data mining in medical domains is quickly growing as the usefulness of these techniques to classification and prediction systems improves, particularly in assisting medical practitioners in their decision-making. As a result, several research studies have been conducted to create suitable risk ratings and Clinical Prediction Models (CPM). In addition to its importance in discovering ways to improve patient outcomes, lower the cost of medicine, and improve clinical studies, it is also important in finding ways to improve patient outcomes, reduce the cost of medicine, and help with clinical trials.

Explainability has become a hot topic in Artificial Intelligence (AI) research to improve confidence in models like Deep Learning (DL) networks. However, explainable AI (XAI) methods can only provide trustworthy models if the XAI methods themselves can be trusted. The evaluation of XAI methods arose from this need. Several research studies have been conducted to create suitable risk ratings and Clinical Prediction Models (CPM). Many complicated machine learning and prediction modeling techniques, such as neural networks and ensemble methods, have been demonstrated to beat linear modeling approaches in terms of accuracy and precision, but they are sometimes difficult to understand. This trade-off impedes the adoption of such models in high-stakes environments ([12,16,28]).

The Wisconsin Diagnostic Breast Cancer (WDBC) dataset [9] was used to create our CPM. The dataset includes characteristics computed from a digitized image of a fine needle aspirate (FNA) of a breast lump, according to [9]. The characteristics of the cell nuclei in figure 1 are described by the mentioned features [9].

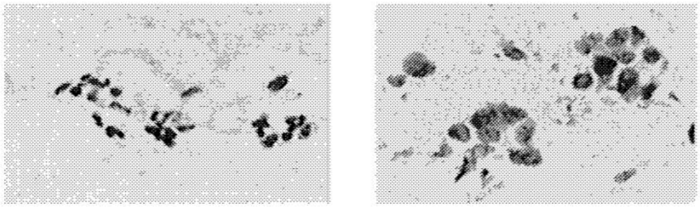


Fig. 1. Image from [9]. (a) Benign, (b) Malignant FNA digitized pictures

We examine the performance of two distinct prediction algorithms for this task: a more easily understandable Decision Tree (DT) model and an ensemble Decision Tree with an SVM model. Internal validation of the models is done using a dataset that has been held out. To give intelligibility to the algorithms' predictions, we select specific classification cases to be explained by Local Interpretable Model-Agnostic Explanations (LIME), SHapley Additive exPlanation (SHAP) values and Skope-rules package.

Materials and Methods

We focus on locally interpretable models, which encourage understanding of the learned relationship between the input variable and target variable over small regions, which is in line with the purpose of this research. Local interpretability is typically used to generate explanations for specific predictions produced by a classifier for a given instance.

Lack of "stability" is one of the problems with present locally interpretable models. This is defined as "explanation level uncertainty" in [13]. The authors show that due to the simplicity of the black-box model, explanations provided by different locally interpretable models have some uncertainty associated with them. This paper takes a more detailed look at the problem. Can explanations provided by a locally interpretable model provide consistent findings for the same instance? That is the primary question we wish to answer. LIME [23] is one of the earliest locally interpretable models, which uses random perturbation to produce simulated data points surrounding an instance and then fits a sparse linear model over projected responses from the perturbed points to provide explanations. LIME was extended using decision rules in [23].

Similarly, leave-one-covariate-out (LOCO) [17] is a popular method for building local explanatory models with local variable importance measures. Instead of perturbing the data points around an instance being explained, the authors advocated utilizing K-means to segment the dataset. The authors of [11] developed a supervised tree-based strategy for partitioning the dataset. In [15], the authors employed LIME in precision medicine and addressed the necessity of interpretability in understanding the role of key aspects in decision making.

Following the inclusion of irrelevant features to the explained instance in a sentiment analysis task, the authors of [28] investigate the robustness of Local Interpretable Model-agnostic Explanations (LIME), SHapley Additive exPlanations (SHAP), and Learn To eXplain (L2X) [16]. According to the findings, L2X explanations outperform other explanation strategies on all of the study's recommended robustness measures. Changes in the feature values of instances can create out-of-distribution data, which is a shortcoming of this family of assessment procedures. Hence, as the authors of [6,22] mentioned, black-box model predictions might have much uncertainty.

In [18], the authors developed a method for decomposing a classifier's predictions into individual contributions of each feature. The difference between original forecasts and predictions produced by removing a set of features is computed using this methodology. The authors of [17] established the equivalence of several local interpretable models [10,11,22], as well as a game theory-based strategy to explain the model dubbed SHAP (Shapley Additive exPlanations). [4] suggests a method for obtaining local explanations based on local gradients, which illustrate the movement of data points to change their predicted label. [28,30,31] adopted a similar method to explain and understand the behavior of image categorization models.The Figure depicts the workflow of our study. A literature review of various studies was carried out for the same dataset. The focus was on XAI models used for CPM. The other steps, as depicted, are explained hereafter.

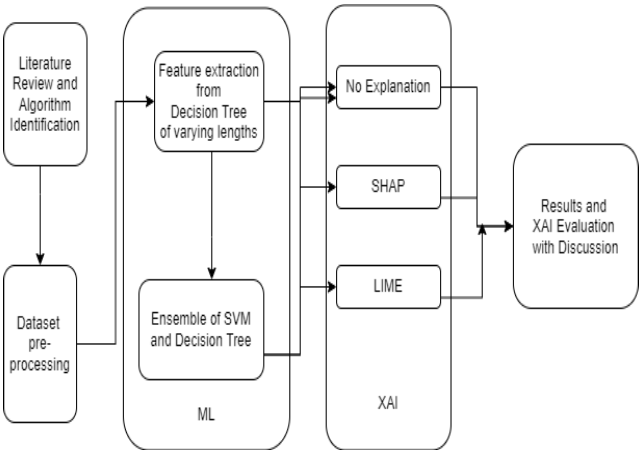


Fig. 2. Flowchart of the Methodology

Algorithm Identification

Feature selection (FS) is a crucial topic in pattern recognition and data mining. Its goal is to pick the most differentiating qualities from a list of options while removing the ones that aren't necessary. The breast cancer Wisconsin Diagnostic dataset is used in this investigation. The dataset was examined before the ML models were chosen. The dataset contains 569 data points, 212 malignant and 357 benign data points. As a result, the dataset's characteristics are as follows: (1) radius, (2) texture, (3) perimeter, (4) area, (5) smoothness (local variation in radius lengths), (6) compactness (perimeter^2 / area - 1.0), (7) concavity (severity of concave portions of the contour), (8) concave points (number of concave portions of the contour), (9) symmetry, and (10) fractal dimension ("coastline approximation" - 1) are all factors to consider. Three pieces of information are computed for each feature: (1) mean, (2) standard error, and (3) "worst" or largest (mean of the three largest values). As a result, there are a total of 30 dataset features.

Because the dataset has a binary output, supervised learning; specifically classification approaches, are the most appropriate types of ML models to explore [1-3]. Regression approaches can generate a likelihood for a data point to belong to a class, which can then be translated into a class label. Unsupervised approaches may produce some results, but they were not investigated further because the binary output is already available.

For our study, two levels of feature reductions are proposed. Level 1 of feature reduction selects a minimal feature subset based on a Decision Tree without pruning, whereas level 2 selects features from the minimal feature subset acquired from level 1 based on pruning. These features are also used in an ensemble model of the Decision Tree with SVM.

Dataset Preprocessing

After discretizing the data using the discretized filter, missing values were removed from the dataset. Second, instances were resampled using the resample filter to preserve the subsample's class distribution and nudge it toward a uniform distribution.

Machine Learning Models

Only two machine learning models were chosen to keep the project scope to a minimum. The two models that were chosen were chosen based on their respective complexity. We examine the performance of two distinct prediction algorithms for this task: a more easily understandable Decision Tree (DT) model and an ensemble Decision Tree with an SVM model. Internal validation of the models is done using a dataset that has been held out. To give intelligibility to the algorithms' predictions, we select specific classification cases to be explained by Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanation (SHAP) values.

We used the Gini impurity measure to determine optimal splits for the Decision Tree. Using a parameter grid-search over a parameter grid of 3 to maximum depth for tree depth, 1 to 10 for minimum samples per split, and 1 to 16 for minimum samples per leaf, and 5-fold cross-validation, we determined an optimal maximum tree depth of 6, a minimum of 1 sample for each leaf node, and a minimum of 2 samples for each split.

In PAC (probably approximately correct) learning theory, most available empirical ensemble analyses use weak learners (e.g., decision trees, neural networks, or naive Bayes). Because learning classification methods aim to enhance classification accuracy while also increasing generalization ability, ensembles based on non-weak classifiers, such as support vector machines, are vital to investigate (SVM). Support vector machines (SVMs) are a new generation of learning systems based on statistical learning theory developments [13]. SVMs calculate a separating hyperplane that maximizes the margin between data classes to achieve strong generalization abilities. SVMs have proven to be an effective learning machine in various applications [7,29]. SVMs, while their superior performance, have significant drawbacks. Multiclass classification, for example, cannot equal the performance of binary classification because SVMs utilize approximation algorithms to reduce computing complexity, but they degrade classification performance [13]. As a result, researchers have tried to improve SVMs using ensemble techniques.

Explainable AI

Interpretability, a qualitative measure of how comprehensible an explanation is to humans, can be used to evaluate XAI technologies [10]. While this is necessary to ensure proper interaction between humans and the model, interpretability frequently includes end-users, resulting in substantial biases [20]. A qualitative evaluation alone cannot ensure coherency with reality (i.e., model behavior), because incorrect explanations can be easier to understand than accurate ones. We also require quantitative and objective evaluation criteria that validate the relationship between the explanations given by the XAI approach and the trained model being assessed to enable trust in XAI methods.

Understanding why a model generated a prediction is critical for various activities, including trust, actionability, accountability, and debugging. Important values are often assigned to each input feature to comprehend predictions using tree ensemble methods like gradient boosting trees or random forests. These significance values might be generated for a single prediction or a whole dataset to describe a model's overall behavior. Current feature attribution approaches for tree ensembles are inconsistent, which means they can give higher value to characteristics that have a smaller impact on the model's output. Because tree ensemble methods are widely used in research and industry, this inconsistency affects many users.

By combining tree ensemble features with the recently defined class of additive feature attribution methods [17], we can justify usingSHapley Additive exPlanation (SHAP) values as the sole consistent feature attribution method with acceptable qualities.

SHAP was implemented using the `shap' python package [24]. Local explanations were generated using automatic mode for the number of perturbations (samples). Force plots (figure 5) were generated for the Decision Tree and the ensemble models for local explanations. The inference here is displayed in the explanations for the ensemble model.

LIME was implemented using the `lime' python package [22]. Explanations were generated to show the 10 highest impact labels, using the probability of malignant tumor. LIME was implemented for both the Decision Tree and the ensemble model. Similar to SHAP, the inference is shown only for the ensemble model.

Skope-rules is another explainability package that can be used to construct rules from the training model and then make predictions on any new dataset. The decision tree model may be trained, and the following rules can be drawn using the *clf.fit()* technique.

Results

This section may be divided into subheadings. It should provide a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.The Decision Tree model was iterated for various depths, and the following result was observed:

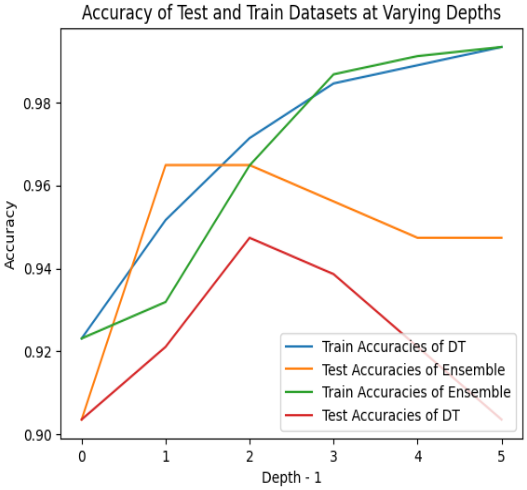


Fig. 3. Accuracy for the two models at varying depths

The test accuracy for the Ensemble method of Decision Tree with SVM gave us better results, and the increased depth resulted in overfitting, which gave better training results but worse test results, as expected. The exact values of precision, accuracy and F1 score for depth =4 are shown.

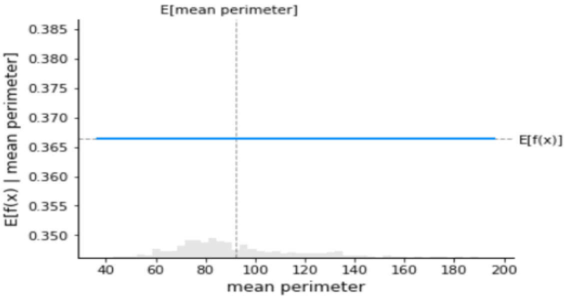
Model	Precision	Accuracy	F1-Score
Decision Tree	0.882	0.9033	0.8718
Ensemble of Decision Tree and SVM	0.9858	0.9561	0.9436

The top 10 features extracted from the Decision Tree are:mean radius, mean texture, mean perimeter, mean area, mean smoothness, mean compactness, mean concavity, mean concave points, mean symmetry and mean fractal dimension

The top 10 features extracted from the Ensemble model are: mean texture, mean smoothness, mean symmetry, radius error, concave points error, symmetry error, worst texture, worst perimeter and worst concave points.

The ensemble seems to point to features more logically, as it is looking at the worst values in some features, whereas the Decision Tree extracts the mean of all features as the most important ones.

A comprehensive assessment of all 569 patients was conducted by creating partial dependence plots (PDP) for each feature, as shown in Figures 4-9. The blue line represents the average probability of a malignant tumor is present. For example, in the case of patient number 25, the PDP for mean perimeter, mean compactness and mean area does not affect the probability of malignant tumors, as illustrated by the straight blue line in Figure. The most critical parameters that contribute to the malignant tumor prediction process are mean texture, smoothness, worst concave points, worst texture, and worst perimeter. Because the other features are unconnected, their partial dependency plots are parallel lines. Thus a focused study for each patient is carried out, and we have details available of the factors pertinent to the patient. These factors were analyzed and discussed with domain experts, and there is a reasonable agreement between the features suggested by the SHAP package and the ones a radiologist looks for when referring to further oncologists.



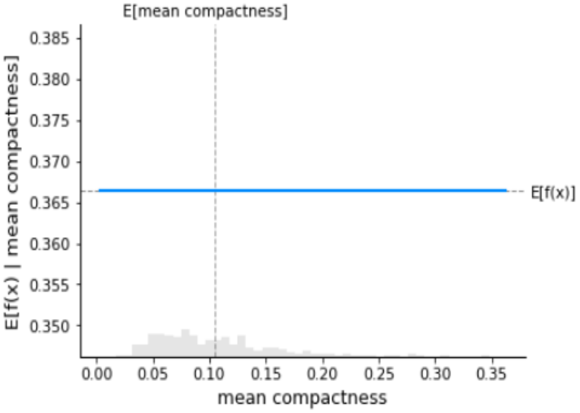


Fig 4 and 5. PDP for mean perimeter and mean compactness

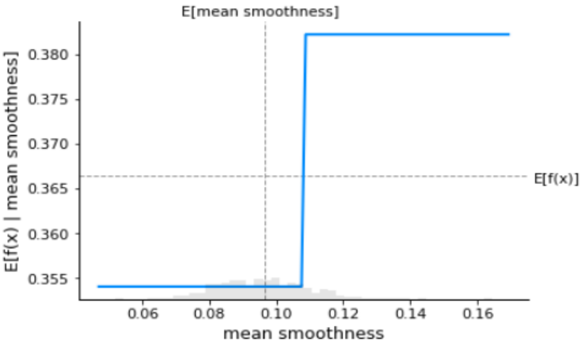


Fig 6 and 7: PDP for mean smoothness and mean area

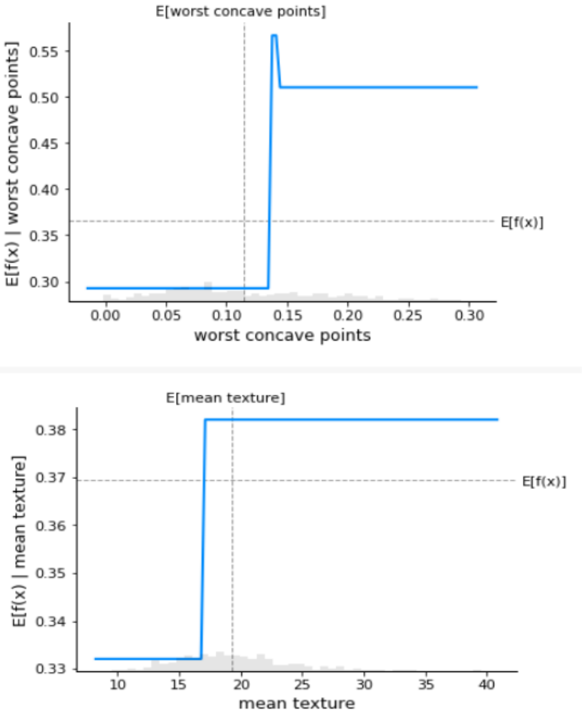
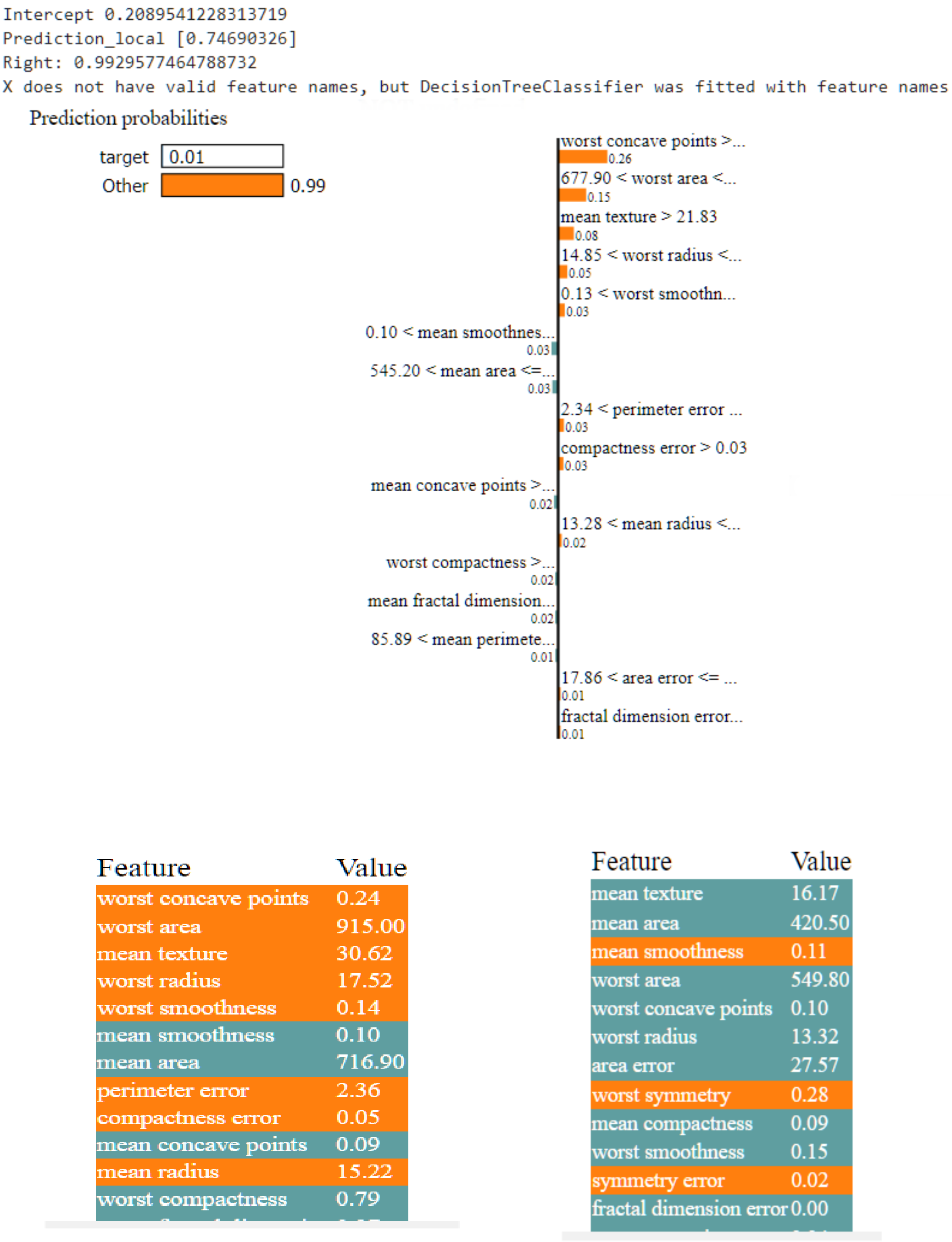


Fig 8 and 9. PDP for mean texture and worst concave points

SHAP Values give the following top 10 features for detection of malignant tumors based on taking mode for varied patients on a local level:
worst concave points, worst area, mean smoothness,worst concavity, area error, mean radius, concavity error, worst symmetry, worst compactness, and worst smoothness. These are pertinent, as confirmed by the radiologist consulted for the same.

The LIME explanation displays the results tabularly, including the predicted value, positive and negative feature values, and overall feature values. For example, two rows are shown as a sample in figure 10 and figure 11. Figure 10 is for a benign tumor output predicted with 0.99 probability. The second chart displays the same occurrence's positive and negative feature values. In the second chart, the horizontal bars reflect the significance of the attribute for the record. The third table indicates the LIME value associated with each attribute. The methods and explanations on the ground are pretty intuitive. The sampling locality for the 25th record uniformly and randomly selects a single data point, resulting in altered data points and a matching predicted value from the model. By default, the feature selection is auto. LIME focuses on fitting the interpretable model to the perturbed dataset using the sample weights and providing local explanations by using the newly trained model.

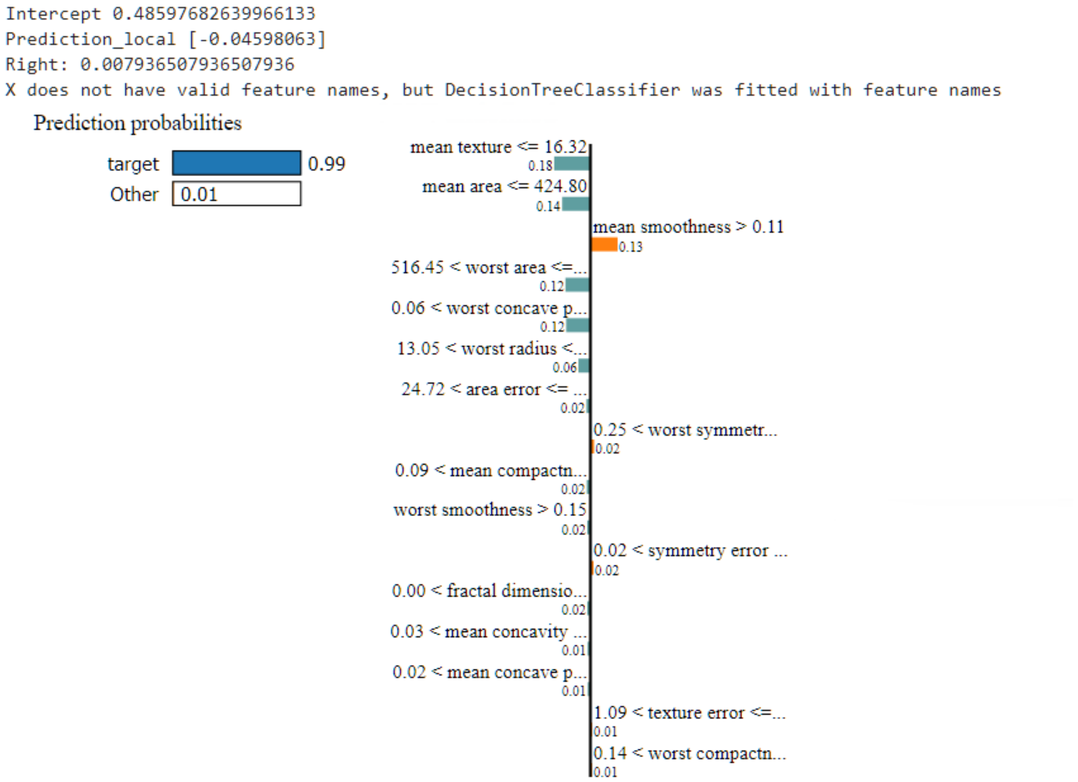


```
Intercept 0.2089541228313719
Prediction_local [0.74690326]
Right: 0.9929577464788732
X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names
```

Feature	Value
worst concave points	0.24
worst area	915.00
mean texture	30.62
worst radius	17.52
worst smoothness	0.14
mean smoothness	0.10
mean area	716.90
perimeter error	2.36
compactness error	0.05
mean concave points	0.09
mean radius	15.22
worst compactness	0.79

Feature	Value
mean texture	16.17
mean area	420.50
mean smoothness	0.11
worst area	549.80
worst concave points	0.10
worst radius	13.32
area error	27.57
worst symmetry	0.28
mean compactness	0.09
worst smoothness	0.15
symmetry error	0.02
fractal dimension error	0.00

Fig 10: Benign Tumor feature values



```
Intercept 0.48597682639966133
Prediction_local [-0.04598063]
Right: 0.007936507936507936
X does not have valid feature names, but DecisionTreeClassifier was fitted with feature names
```

Fig 11. Malignant Tumor feature values

Using the Skope Rules package, we tried to retrain the model, and by adjusting the precision and recall parameters, the rules are curated. If the threshold value is lowered, many rules are generated, which is useless to the user because false-positive rules are introduced into the application. A higher threshold will reduce the number of rules and the usage of only the most pertinent rules in applications. We have generated the top 5 rules to determine a malignant tumor for the decision tree as follows:

area error > 13.929999828338623 and worst concave points > 0.1423499956727028 and mean smoothness > 0.07928499951958656
worst radius > 16.789999961853027 and worst concavity > 0.21620000153779984
mean texture > 16.795000076293945 and worst perimeter > 106.0999984741211 and mean smoothness > 0.08302000164985657
radius error > 0.17840000241994858 and worst area > 729.5499877929688 and worst concave points > 0.1423499956727028
worst texture > 25.670000076293945 and worst area <= 871.7999877929688 and worst concave points > 0.13580000400543213

Further, we generated the top 5 rules to determine a malignant tumor for the ensemble model as follows:

worst area > 690.3500061035156 and mean concave points > 0.04988499917089939 and worst texture > 20.355000495910645
worst texture > 20.045000076293945 and worst perimeter > 101.64999771118164 and worst concave points > 0.10954999923706055
worst area > 495.1999969482422 and worst concave points > 0.13580000400543213 and worst texture > 20.355000495910645
worst texture > 19.90999984741211 and worst perimeter > 107.5999984741211 and worst concave points <= 0.14239999651908875
worst texture > 21.274999618530273 and worst area > 929.8000183105469 and worst concave points <= 0.1423499956727028

Discussion

Authors should discuss the results and how they can be interpreted from the perspective of previous studies and the working hypotheses. The findings and their implications should be discussed in the broadest context possible. Future research directions may also be highlighted.As seen in the previous section, the ensemble method gives better results in determining the most important features. The Decision Tree gives 100% accuracy in the training set; the accuracy drops to 90% in testing. This is due to the overfitting model when no pruning is done. We have finally used a pruned version with depth =4 for our model.

What is interesting to observe is the performance of the SHAP, LIME, and Skope rules package. LIME is predicted to perform similarly to SHAP. However, metrics may imply that it is less interpretable or trustworthy due to the absence of a global explanation. LIME lends itself well to decision tree models, as it highlights the model features' inflection points when providing a prediction. In certain cases, these inflection points may seem illogical, such as indicating that a patient with an area error bigger than 13.93 increases their likelihood of having a malignant tumor.

In this case, the explanation is analyzing other features before area error, and when it evaluates the 'area error' feature, it may not have a high priority in comparison to other characteristics, or the explanation frames it in a way that is not immediately understandable. Nonetheless, these are examples of explanations that do not fit the evidence. Even if the forecast is accurate, it may be accurate for the wrong reasons, resulting in low user confidence and interpretability. Additionally, interpretability may be weaker for LIME than for SHAP due to the possibility of information overload, as LIME presents three separate visual representations when evaluating a local prediction, but SHAP provides just one. One of LIME's shortcomings is its imprecise definition of neighborhood and proximity. It employs a Gaussian distribution to sample the records in the locality, which ignores the link between various features. If there is a non-linear relationship between the dependent and independent variables, the explanations will be inaccurate.

The PDP gives insight at a local level, and when studied at the individual level, the results are very promising, as per the feedback from the radiologist we consulted. The Skope Rules generated for the ensemble method are pertinent and explainable as per the feedback from the radiologist. The 'area error' being considered in the first set of rules generated for the Decision Tree does not seem convincing to the doctor. Hence, the ensemble method is preferred over here as well.

Conclusions

This section is not mandatory but can be added to the manuscript if the discussion is unusually long or complex.The user study findings indicate that giving a post-hoc explanation for a prediction improves the results' interpretability. LIME,SHAP, and Skope Rules are examples of post-hoc explanations that improve interpretability. The increase in interpretability that has been measured is significant. Additional feedback from the radiologist research indicates that some flaws are inherent in the data and that certain aspects may be completely avoided at the outset to avoid interfering with the model construction process. Additionally, a study with a similar approach might be conducted on a different dataset or in a more general domain using a different sub-domain. Repeating this study with a larger sample of users at a later period could also generate interesting results.

Numerous significant inferences can be drawn from the findings. While both the decision tree and ensemble methods perform admirably, a more advanced form of machine learning algorithm such as Neural Networks should also be investigated further. Because example-based XAI techniques were not examined in this study, they should be studied and appraised compared to SHAP and LIME. Finally, performing a user study in the form of an interview in partnership with domain experts will almost certainly generate complete and meaningful data. While SHAP, LIME, and Skope Rules improved interpretability significantly in this case, whether these results might be applied to a larger user study or if the data is generally applicable to health care systems.

Data Availability Statement

Data used is from UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: the University of California, School of Information and Computer Science

Conflicts of Interest: "The author declares no conflict of interest."

References

1. Agarwal R, H, P, SS (2021a) Deep Learning augmented with Contour Detection Designed for Diagnosis of COVID-19 in X-rays. Design Engineering(Toronto) 8:10445–10461
2. Agarwal R, Hariharan S, Nagabhushana Rao M, Agarwal A (2021) Weed Identification using K-Means Clustering with Color Spaces Features in Multi-Spectral Images Taken by UAV. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. IEEE, pp 7047–7050
3. Agarwal RAA (2021b) Decision Support System designed to detect yellow mosaic in Pigeon pea using Computer Vision. Design Engineering(Toronto) 8:832–844

4. Baehrens D, Schroeter T, Harmeling S, et al (2009) How to Explain Individual Classification Decisions
5. Camburu O-M, Giunchiglia E, Foerster J, et al (2019) Can I Trust the Explainer? Verifying Post-hoc Explanatory Methods
6. Chen L-C, Zhu Y, Papandreou G, et al (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation
7. Chih-Wei Hsu, Chih-Jen Lin (2002) A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks 13:415–425. https://doi.org/10.1109/72.991427
8. Dabkowski P, Gal Y (2017) Real Time Image Saliency for Black Box Classifiers Fong R, Vedaldi A (2017) Interpretable Explanations of Black Boxes by Meaningful Perturbation. https://doi.org/10.1109/ICCV.2017.371
9. Frank, A. & Asuncion, A. (2010). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.
10. Gilpin LH, Bau D, Yuan BZ, et al (2018) Explaining Explanations: An Overview of Interpretability of Machine Learning
11. Goncharuk, N., Lysenko, O., Nataliya, P., Kovyda, N., & Tsekhmister, Y. (2022). Management of psychological changes at pregnant women during the COVID-19 pandemic. International Journal of Health Sciences, 6(2), 870–879. https://doi.org/10.53730/ijhs.v6n2.8528
12. Gosiewska A, Biecek P (2019) Do Not Trust Additive Explanations
13. Katuwal GJ, Chen R (2016a) Machine Learning Model Interpretability for Precision Medicine
14. Kim K (2003) Financial time series forecasting using support vector machines. Neurocomputing 55:307–319. https://doi.org/10.1016/S0925-2312(03)00372-2
15. Lakkaraju H, Bastani O (2019) "How do I fool you?": Manipulating User Trust via Misleading Black Box Explanations
16. Lei J, G'Sell M, Rinaldo A, et al (2016) Distribution-Free Predictive Inference For Regression
17. Letham B, Rudin C, McCormick TH, Madigan D (2015) Interpretable classifiers using rules and Bayesian analysis: Building a better stroke prediction model. https://doi.org/10.1214/15-AOAS848
18. Lundberg S, Lee S-I (2017a) A Unified Approach to Interpreting Model Predictions
19. Lundberg SM, Erion G, Chen H, et al (2020) From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence 2:56–67. https://doi.org/10.1038/s42256-019-0138-9
20. Malvia S, Bagadi SA, Dubey US, Saxena S (2017) Epidemiology of breast cancer in Indian women. Asia-Pacific Journal of Clinical Oncology 13:289–295. https://doi.org/10.1111/ajco.12661
21. Mohseni S, Zarei N, Ragan ED (2018) A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems
22. Nurwahidah, H., Maswan, M., & Fathoni, A. (2020). The effect of deep breathing technique and lo'i sto combination on decreasing the symptomps of asthma patients at the area of public health center of Penana'e. International Journal of Social Sciences and Humanities, 4(1), 140–150. https://doi.org/10.29332/ijssh.v4n1.431

23. Rahnama AHA, Boström H (2019) A study of data and label shift in the LIME framework
24. Ribeiro MT, Singh S, Guestrin C (2016a) "Why Should I Trust You?": Explaining the Predictions of Any Classifier
25. Robnik-Sikonja M, Kononenko I (2008) Explaining Classifications For Individual Instances. IEEE Transactions on Knowledge and Data Engineering 20:589–600. https://doi.org/10.1109/TKDE.2007.190734
26. Ross AS, Hughes MC, Doshi-Velez F (2017) Right for the Right Reasons: Training Differentiable Models by Constraining their Explanations. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, California, pp 2662–2670
27. Simonyan K, Vedaldi A, Zisserman A (2013) Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps
28. Singh SD, Henley SJ, Ryerson AB (2017) Surveillance for Cancer Incidence and Mortality - United States, 2013. Morbidity and mortality weekly report Surveillance summaries (Washington, DC : 2002) 66:1–36. https://doi.org/10.15585/mmwr.ss6604a1
29. Sundararajan M, Taly A, Yan Q (2017) Axiomatic Attribution for Deep Networks
30. Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). Get vaccinated when it is your turn and follow the local guidelines. International Journal of Health Sciences, 5(3), x-xv. https://doi.org/10.53730/ijhs.v5n3.2938
31. Valdes G, Luna JM, Eaton E, et al (2016) MediBoost: a Patient Stratification Tool for Interpretable Decision Making in the Era of Precision Medicine. Scientific Reports 6:37854. https://doi.org/10.1038/srep37854
32. Widodo A, Yang B-S, Han T (2007) Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. Expert Systems with Applications 32:299–312. https://doi.org/10.1016/j.eswa.2005.11.031
33. Zeiler MD, Fergus R (2013) Visualizing and Understanding Convolutional Networks
34. Zhou B, Khosla A, Lapedriza A, et al (2015) Learning Deep Features for Discriminative Localization