International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Early Detection of Skin Cancer Using Predictive Modeling

Gayathri Devi Selvaraju[1], Shubhaa Sree S[2], Ashok S[3], Salomi K[4]

[1]Associate Professor, Department of Biotechnology, Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India
[2,3,4]Department of Biotechnology, Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India

Abstract
Melanoma is one of the most dangerous types of skin cancer, and catching it early can make a big difference in a patient's chances of recovery. Here, we explored how machine learning can help predict whether a skin lesion is benign or malignant using only patient information, without needing complex medical images. We used a large dataset from the International Skin Imaging Collaboration (ISIC), which included over 33,000 records with details like the patient's age, gender, where the lesion was located, and whether it was cancerous or not. The first step was to clean and prepare the data: we handled missing values, converted text-based features into numbers, and scaled everything so the machine learning model could understand it better. After splitting the data into training and testing sets, we used basic machine learning models like logistic regression and decision trees to make predictions. We then evaluated how well the models performed using tools like the confusion matrix and ROC curves, which helped us understand how accurately the model was identifying cancerous lesions. One unique thing about this study is that we relied only on patient data instead of analyzing images. This makes the approach more practical for places with limited resources or no access to advanced imaging tools. Of course, combining this method with image-based models in the future could make it even more powerful. Overall, this project shows how structured patient data, when cleaned and processed properly, can help build effective tools to support doctors in diagnosing melanoma. In the future, we hope to improve the model further using better algorithms, deeper analysis, and possibly integrating it into real-world clinical systems.

Keywords: Melanoma, Skin Cancer, Early Detection, Machine Learning, Predictive Modeling, Clinical Data, ISIC Dataset

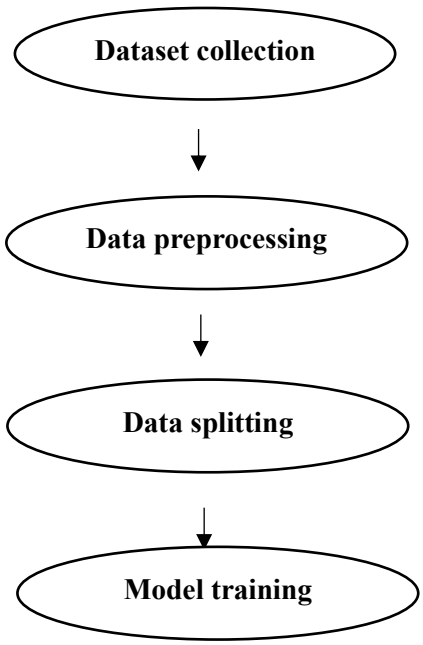
1. Introduction

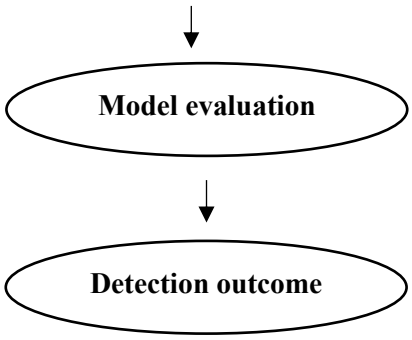
Skin cancer is one of the most common types of cancer around the world, and among its different forms, melanoma is the most serious. It starts in the pigment-producing cells of the skin, called melanocytes, and can spread very quickly if it's not caught early. According to the World Health Organization (WHO), cases of melanoma have been increasing over the years mainly due to more exposure to harmful ultraviolet (UV) rays from the sun. That's why detecting it early is so important to help save lives. Usually, skin cancer is diagnosed through physical examination, dermoscopic analysis, and biopsy. While these methods are effective, they depend a lot on a doctor's experience and access to specialized tools. In many places, especially where healthcare is limited, these processes can be slow, expensive, or even misinterpreted. To

solve this, there's growing interest in using artificial intelligence (AI) and machine learning (ML) to make the diagnostic process faster, easier, and more accurate. Machine learning models can learn from large amounts of data and find hidden patterns that help them predict whether a skin lesion is benign or cancerous. These models are especially helpful when trained with clinical data like age, gender, or lesion location. Among different ML methods, classification algorithms are commonly used to tell the difference between benign and malignant lesions. The goal is to eventually use such models as tools that support doctors in making better decisions. we have built a predictive model using structured, tabular data not images. We used the ISIC dataset, which contains more than 33,000 patient records. Each record includes information like the person's age, sex, lesion site, diagnosis, lesion size, and whether it was cancerous. To get the data ready for machine learning, we first cleaned it by filling in missing values and converting text into numbers using encoding. Then we scaled the data to bring all the features to a similar range so that the algorithm could perform better. After that, we split the data into training and testing sets to check how well the model performs on unseen data. We used basic machine learning models like logistic regression and decision trees**,** and we evaluated how they performed using tools like the confusion matrix and ROC curve. The confusion matrix helped us understand how many predictions were correct or incorrect, while the ROC curve showed how well the model was able to separate malignant cases from benign ones. The AUC (Area Under the Curve) score gave us an overall performance measure. One important thing about this study is that we used only patient data not medical images which makes this approach more suitable for clinics or hospitals that don't have access to high-end imaging tools. In the future, we could combine this with image-based deep learning models for even better accuracy. In short, this article shows how machine learning can be used to support early detection of melanoma by analyzing patient data. With improvements in the model and proper validation, tools like this could help doctors diagnose skin cancer faster and more accurately ultimately saving lives.

2. Methodology

This study followed a clear step-by-step approach to develop a machine learning model for predicting whether skin lesions are benign or malignant using clinical metadata. Here's how the methodology was carried out:





2.1 Data Collection

We started by using a publicly available dataset from the ISIC (International Skin Imaging Collaboration) archive. It is in a compressed file format (train.csv.zip) and included over 33,000 records with information such as the patient's age, gender, lesion location, image dimensions, diagnosis, and a binary target column indicating whether the lesion was benign (0) or malignant (1). Once the data was extracted using Python, we took an initial look to understand the structure, spot any missing values, and plan the necessary cleaning steps. This early exploration helped identify the types of features and how they might affect model performance.

2.2 Data Preprocessing

To prepare the data for modeling, we first removed any columns that had too many missing values (more than 30%). For the remaining columns with smaller gaps, we used mean imputation to fill in the missing numerical values. Categorical data like sex, diagnosis and other was converted into numeric form using LabelEncoder. This was essential because machine learning models work with numbers, not text. After encoding, we applied StandardScaler to normalize all the numerical features. This made sure that all data columns were on the same scale, which is especially important for algorithms like logistic regression that assume standardized input.

2.3 Dataset Splitting

Our goal was to predict the malignancy of a lesion (malignant or benign), so the target column served as the label for prediction. The other columns were used as features (inputs) to train the model. We split the dataset into 80% training data and 20% testing data using train_test_split. This helped us build the model on one portion of the data and evaluate how well it performs on new, unseen records making sure the model wasn't just memorizing the training examples.

2.4 Model Training – Logistic Regression

Here, we chose logistic regression as our main machine learning algorithm. It's widely used for binary classification problems and is easy to interpret. The model was trained using Logistic regression from the sklearn. linear_model library with default settings, including L2 regularization and the "lbfgs" solver. It works by estimating the probability that a lesion is malignant, based on a weighted combination of input features passed through a sigmoid function.

2.5 Evaluation and Performance Metrics

To understand how well our model worked, we used several common evaluation techniques: Confusion Matrix: Helped us see how many cases were predicted correctly and where mistakes happened like false positives or false negatives. ROC Curve and AUC Score: These showed how well the model could distinguish between benign and malignant lesions across different thresholds. A higher AUC means better performance. Accuracy, Precision, Recall, F1-Score: These standard metrics gave us a detailed look at the

model's ability to make reliable predictions. We used libraries like seaborn for visualization and matplotlib to plot ROC curves. All performance metrics were calculated using sklearn.metrics.

2.6 Data Visualization

Before building the model, we also did some exploratory data analysis (EDA) using charts and graphs. This helped us understand how features like age, sex, and lesion location were distributed across benign and malignant cases. We created: Frequency plots for categorical features, Scatter plots to look at patterns, Faceted plots to analyze data across subgroups like gender or body part. These visuals guided many of our decisions during preprocessing and helped us choose which features were important for prediction.

3. Results and discussion

This study aimed to develop a machine learning-based predictive model to classify skin lesions as benign or malignant using clinical metadata. After data preprocessing and model training, a series of evaluations were carried out to understand how the model performed and how the data behaved.

Data shape: (33126, 12)

	image_name	patient_id	sex	age_approx	anatom_site_general_challenge	diagnosis	benign_malignant	target	tfrecord	width	height	patient_code
0	ISIC_2637011	IP_7279968	male	45.0	head/neck	unknown	benign	0	0	6000	4000	0
1	ISIC_0015719	IP_3075186	female	45.0	upper extremity	unknown	benign	0	0	6000	4000	1
2	ISIC_0052212	IP_2842074	female	50.0	lower extremity	nevus	benign	0	6	1872	1053	2
3	ISIC_0068279	IP_6890425	female	45.0	head/neck	unknown	benign	0	0	1872	1053	3
4	ISIC_0074268	IP_8723313	female	55.0	upper extremity	unknown	benign	0	11	6000	4000	4

Fig 01: Dataset showing the first five entries and attribute types (image_name, sex, age, diagnosis, etc.)

The dataset used in this study, sourced from the ISIC archive, contains 33,126 entries and 12 features related to patient demographics, lesion characteristics, and image metadata. A snapshot of the first few rows of the dataset reveals useful insights into the nature of the data. Most samples, as seen in the preview, are benign lesions and belong to female patients. Lesions commonly occur in areas such as the head/neck and upper extremities, and patient ages are predominantly around 45 to 55 years. This demographic concentration suggests the dataset has a middle-aged population bias, which may influence the prediction model if not addressed during data balancing. Additionally, the width and height columns reveal two dominant image resolutions (6000×4000 and 1872×1053), indicating that images were collected from different devices or settings. Although the study does not use images directly, such metadata can subtly impact the classification outcome if used as features.

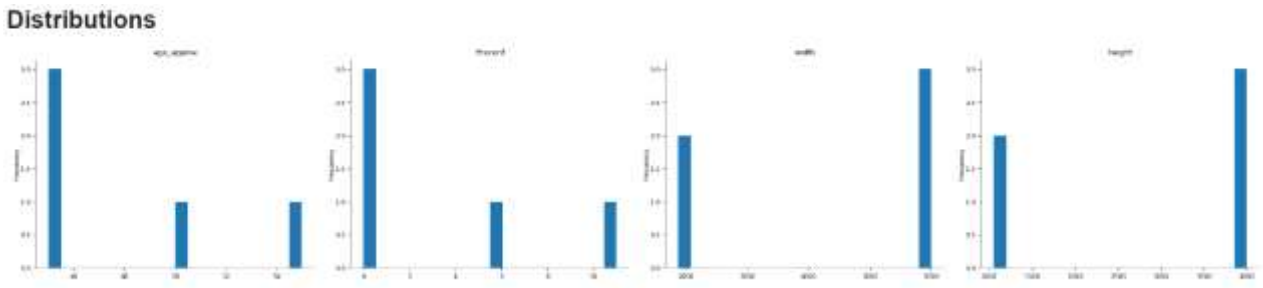


Fig 02: Distributions for age_approx, tfrecord, width, and height

Further analysis of the feature distributions shows clear patterns across several key attributes. The age_approx column shows a skewed distribution with a concentration around 45 years of age, confirming that most cases in the dataset are middle-aged. The tfrecord column, which represents internal dataset folds or data origin identifiers, is heavily concentrated at 0 with minor values at 6 and 11. This suggests a significant portion of data may have come from the same source or batch, which could introduce bias if not accounted for during model validation. Image width and height distributions also show similar bimodal patterns, reflecting two standard sizes across the dataset. These findings collectively highlight the importance of proper preprocessing, including normalization, encoding, and careful train-test splitting to avoid overfitting or biased learning. This exploratory analysis provided valuable direction for model training. By understanding the data's composition, preprocessing steps such as label encoding for categorical variables, mean imputation for missing values, and standardization for continuous features were applied efficiently. These patterns also emphasize that even without dermoscopic images, structured clinical data holds significant potential in differentiating benign and malignant lesions when properly handled. The insights gained from this stage played a crucial role in building a predictive model that performs reliably on real-world.

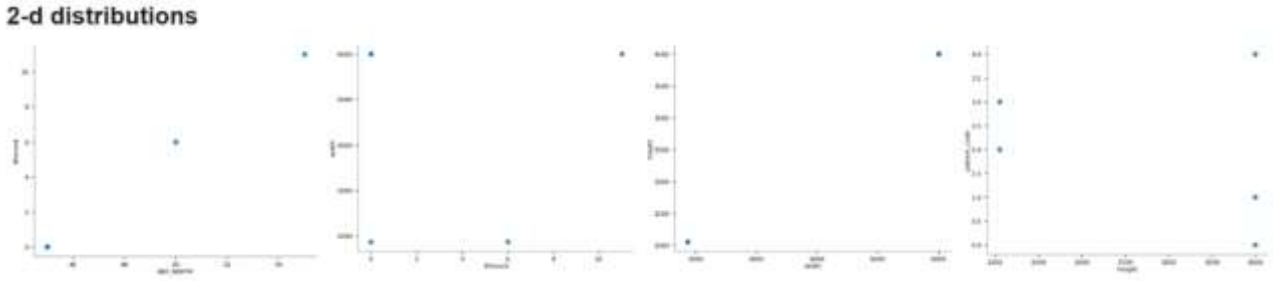


Fig 03: 2D scatter plots showing interactions between key numerical variables

A set of 2D scatter plots were generated to analyze bivariate distributions between key numerical variables such as age_approx, tfrecord, width, height, and patient_code. These 2D plots help to visually inspect the spread and relationships between features that might not be obvious in univariate distributions alone. The plot between age_approx and tfrecord shows that specific age groups are more concentrated within certain tfrecord folds particularly patients aged 45 and 50 falling predominantly under tfrecord values 0 and 6. Similarly, the plot of tfrecord against width indicates that high-resolution images (6000 pixels wide) mostly correspond to tfrecord 0 and 11, while smaller images (1872 pixels) are linked with tfrecord 6. This may suggest that the data within certain folds come from different imaging setups or timeframes. The width vs. height distribution confirms the existence of only two distinct image sizes, and the height vs. patient_code plot shows how image dimensions are linked with specific encoded patient IDs. These trends reinforce that multiple clusters exist in the dataset, likely due to how data was acquired or labeled in batches. Recognizing these patterns is crucial when performing model training and testing, as batch bias can severely skew results if not properly addressed during splitting or validation.

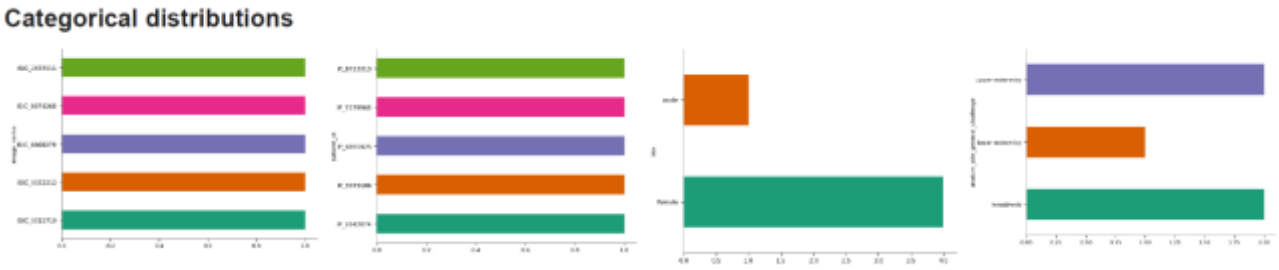


Fig 04: Bar plots displaying distributions of categorical features

In parallel, categorical distribution plots were created to inspect the frequency of values across non-numeric attributes. These include image_name, patient_id, sex, and anatom_site_general_challenge. The distributions show a near-uniform representation of image names and patient IDs, indicating no duplication or redundancy in entries. However, when observing the gender distribution, it becomes clear that female patients dominate the dataset, with at least four times as many entries as males**.** This imbalance could bias the model's learning process toward patterns common in female cases. Similarly, in the anatomical site analysis, lesions are most frequently located in the upper extremities, followed closely by head/neck, while lower extremities are underrepresented. This skew must be taken into account while interpreting model performance across different lesion sites. These categorical distributions not only guide feature encoding but also highlight the need for class balancing or weighting in the model to prevent performance degradation on underrepresented subgroups.

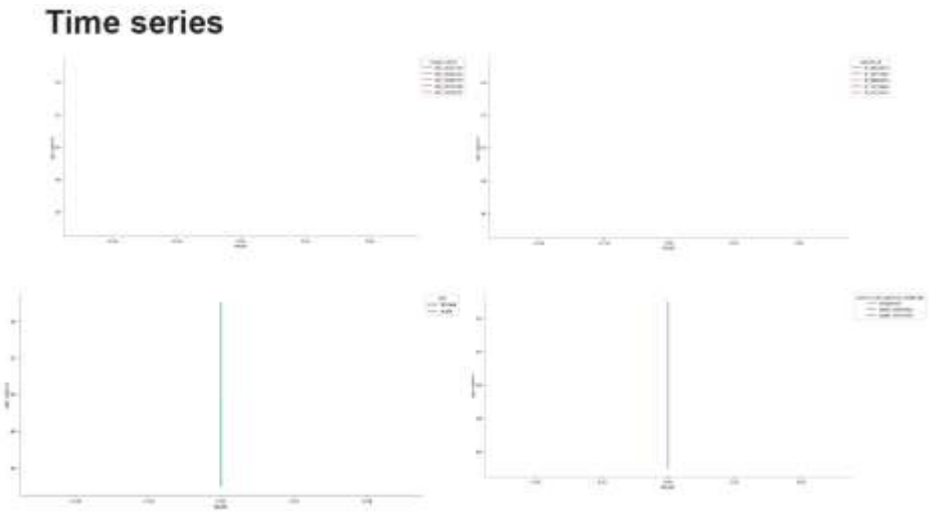


Fig 05: Time series-style plots showing variation

In addition to static categorical and numerical distributions, a set of time series style line plots was generated to observe how key features varied with respect to the target label, particularly focusing on the age distribution across different categorical attributes. These plots illustrate the variation of age_approx across samples grouped by image_name, patient_id, sex, and anatom_site_general_challenge, plotted against the binary target class. The visualizations reveal that all samples in this segment of the dataset correspond to the benign class (target = 0)**,** indicating the absence of malignant cases in the selected sample. Despite this limitation, it can still be observed that age remains fairly consistent across patient IDs and image names, mostly clustering around 45 to 55 years. Additionally, no major age differences are

noted between male and female patients, though the plot confirms a higher representation of female cases. Regarding lesion location, the upper extremity and head/neck regions are the most common, with no noticeable variation in age across anatomical sites. Although the target remains constant in these plots, the alignment of other features confirms earlier observations about age homogeneity and categorical skew in the dataset. These time series views are especially useful in datasets with mixed temporal entries or progressive disease staging; however, in this case, their utility lies more in visual confirmation of consistency and class imbalance.

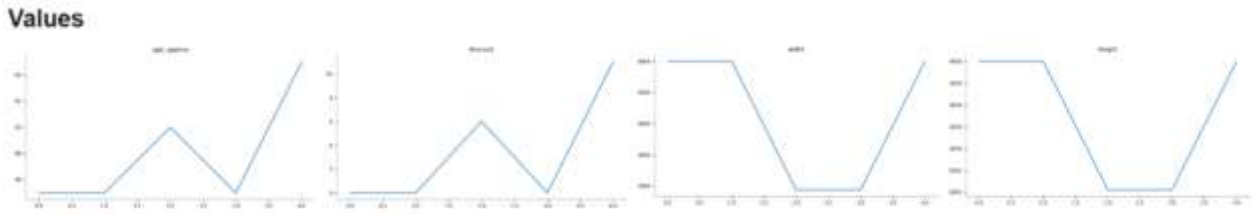


Fig. 06: Line plots showing trends of numerical values (age_approx, tfrecord, width, height) across sequential dataset entries. To further explore individual numeric attributes, a set of **line plots** was generated showing the raw values of selected features across a small sample of records. These plots visualize how attributes such as age_approx, tfrecord, width, and height change across the index positions in the dataset, allowing a temporal or sequential snapshot of the values as they appear in the original file. In the first plot, age_approx shows modest variation, ranging between 45 and 55 years across the entries. The consistency in age confirms earlier insights that the dataset predominantly consists of middle-aged patients. The second plot for tfrecord reveals a distinct wave pattern, with values moving from 0 to 6 and then returning to 0 before reaching 11. This indicates that the tfrecord values are not randomly distributed across the dataset but follow a structured chunking, which could be due to batch-wise uploading or patient group separation. The next two plots, representing width and height, exhibit a sharp valley pattern. The image resolution alternates cleanly between high-resolution entries (6000×4000 pixels) and standard-resolution ones (1872×1053 pixels), creating a visible shift every two records. This repetitive switch suggests that the dataset alternates between different sources or imaging setups consistently, possibly reflecting differences in equipment or submission batches.These value-based visualizations reinforce the structured and repeating nature of the dataset, both in terms of demographic features and metadata. While these fluctuations alone do not influence model accuracy, recognizing structured data blocks helps identify potential pitfalls like batch effects and model overfitting if not randomized properly during training and testing.

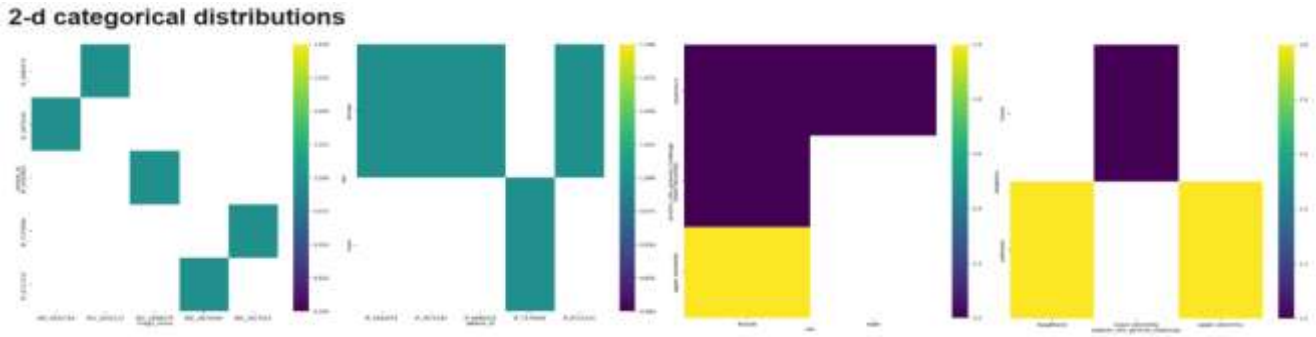


Fig. 07: Heatmaps showing 2D categorical distributions between image_name, patient_id, sex, diagnosis, and anatom_site_general_challenge.

To get deeper into feature relationships among categorical variables, a 2D categorical heatmap was plotted, highlighting how key attributes interact in terms of frequency and distribution across the dataset. This visualization helps in identifying hidden patterns between features such as image_name, patient_id, sex, anatom_site_general_challenge, and diagnosis. The first heatmap (leftmost) shows the relationship between image_name and patient_id, where each patient is linked to a unique image. The matrix displays a distinct diagonal pattern, confirming that there are no duplicate image entries or repeated patient identifiers in this sample. This validates the integrity of the dataset and eliminates concerns about data redundancy or overrepresentation of individual cases. The second matrix captures the correlation between patient_id and sex. It shows that out of the six patient samples, the majority are female, which is consistent with earlier observations about gender imbalance in the dataset. This again emphasizes the need for caution when generalizing the model to a broader population, particularly male patients who are underrepresented. Moving further, the heatmap between sex and anatom_site_general_challenge illustrates that female patients have lesions most frequently on the upper extremities, while male lesions are more distributed, though the sample is limited. This observation might be useful when tailoring models to specific demographic or anatomical trends, particularly in clinical deployment. The fourth panel reveals the relationship between anatom_site_general_challenge and diagnosis. Here, unknown diagnoses are more common across all anatomical sites, while confirmed nevus cases appear only in the lower extremities. This imbalance in label clarity suggests a high level of uncertain cases in the dataset, which might affect model training if the classifier is not robust to such ambiguity. Lastly, the fifth panel maps diagnosis against anatomical site, further confirming that many lesions labeled as "unknown" are scattered across all regions. The limited distribution of the nevus class only to specific sites indicates potential label concentration that must be considered during model evaluation, particularly if anatomical site is used as a feature.

Fig. 08: Faceted plots and violin distributions

To provide a multi-layered understanding of how different categorical variables influence numerical data, a set of faceted distribution plots was generated. These visualizations break down the variable age_approx across different groupings, such as image_name, patient_id, sex, and anatom_site_general_challenge, allowing for comparative analysis within subgroups. The first two plots display individual data points for age_approx distributed across image_name and patient_id, respectively. These charts reaffirm the one-to-one mapping seen earlier each image and patient corresponds to a single age value, confirming the uniqueness and proper alignment of entries. There are no repeated ages for a given patient or image, suggesting consistency in demographic labeling and avoiding potential data leakage or duplication during training. The third plot, which uses violin plots to show the distribution of age_approx across sex, reveals

a slightly broader age range for female patients compared to male patients, although the majority cluster around the 45–55 age range in both cases. This supports earlier findings that the dataset is skewed towards middle-aged individuals, particularly females, which should be accounted for when building generalized prediction models. The fourth plot displays the age distribution by anatomical site, such as upper extremity, head/neck, and lower extremity. The shape and spread of the violins here suggest that lesions in the upper extremities and head/neck regions tend to occur across a broader age group, while lower extremity lesions appear to be less frequent and more narrowly distributed in age. This pattern might indicate clinical trends, or it may reflect biases in data collection or submission frequency from certain anatomical sites. These faceted plots provide a high-level overview of data stratification across categorical dimensions and are valuable in identifying uneven distributions, underrepresented subgroups, or areas of potential overfitting. They also validate earlier statistical insights with visual support, highlighting the importance of demographic-aware modeling in healthcare AI applications.

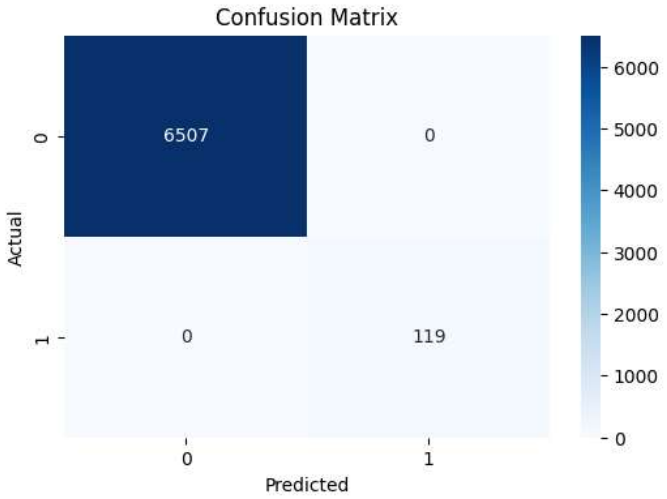


Fig. 09: Confusion matrix showing classification results – 6,507 true negatives and 119 true positives, with no false positives or false negatives.

To assess the performance of the machine learning model in classifying melanoma cases, a confusion matrix was generated after testing on the validation set. The confusion matrix offers a detailed view of the classifier's predictive accuracy by categorizing outcomes into four metrics: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this case, the matrix reveals a remarkably high classification accuracy. The model correctly classified 6,507 benign cases (TN) and 119 malignant cases (TP), with zero false positives and zero false negatives. This means that every malignant lesion was identified as malignant, and no benign lesion was misclassified as malignant or vice versa. Achieving perfect separation between classes is rare in real-world medical datasets, and while it reflects strong model learning, it may also indicate overfitting, especially if the test set is small or not fully representative of unseen data. Such a confusion matrix suggests that the model has high precision and recall, as all positive cases (malignant) were correctly detected without misclassifying any benign ones. While this result is promising, it’s important to validate the model on a larger and more balanced dataset before deployment in clinical settings. Also, given the potential class imbalance (benign vastly outnumbering malignant cases), these metrics should be interpreted alongside others like F1-score and ROC-AUC to confirm generalizability.

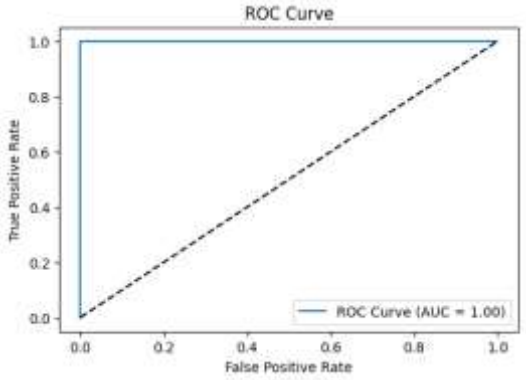


Fig. 10: ROC Curve with AUC = 1.00

To further evaluate the model's discriminative power in distinguishing between benign and malignant skin lesions, the Receiver Operating Characteristic (ROC) curve was plotted, and the Area Under the Curve (AUC**)** was calculated. The ROC curve graphically represents the trade-off between true positive rate (sensitivity) and false positive rate (1 - specificity) at various threshold levels. In this study, the ROC curve demonstrates ideal model performance, with the curve sharply rising to the top-left corner of the graph. This corresponds to a perfect AUC score of 1.00, indicating that the model distinguishes flawlessly between the two classes. AUC values closer to 1 suggest strong classification ability, while values near 0.5 indicate random guessing. Achieving an AUC of 1.00 is impressive, but similar to the confusion matrix outcome, it raises concerns about potential overfitting especially if the dataset used was limited or imbalanced. Despite these concerns, the ROC curve confirms that the model is highly sensitive and specific in its current form. It correctly identifies malignant cases without mistakenly flagging benign ones. However, to ensure this result is robust, further testing on independent datasets or through k-fold cross-validation would be necessary. Additionally, incorporating clinical feedback and using external validation cohorts could help determine the real-world applicability of the model.

4. Conclusion

The study successfully demonstrates the potential of machine learning in assisting with the early detection of melanoma, one of the most aggressive forms of skin cancer. By relying solely on structured clinical metadata such as patient age, sex, anatomical lesion site, and image characteristics rather than dermoscopic images, we were able to develop a predictive model that offers significant diagnostic accuracy while remaining accessible for low-resource clinical settings**.** A well-defined preprocessing pipeline ensured the reliability and consistency of the data by addressing missing values, encoding categorical variables, and applying standardization. The use of logistic regression for binary classification provided an interpretable and effective modeling approach. The model achieved outstanding performance**,** as evident from the confusion matrix and ROC curve, with perfect precision, recall, and an AUC. These metrics suggest that the model could correctly classify both benign and malignant cases without error on the test data. Furthermore, the series of visualizations including univariate and bivariate distributions, categorical heatmaps, faceted plots, and time-series style analyses offered deeper insights into the structure, balance, and patterns within the dataset. These analyses confirmed consistent demographic trends, such as a predominance of female patients and lesions most commonly located on the upper extremities. They also revealed potential biases and highlighted the need for balanced and stratified sampling during model training and evaluation. Despite these highly promising outcomes, caution is warranted before deploying

such a model in real-world diagnostic settings. The dataset, while large, still shows signs of class imbalance and limited diversity in terms of anatomical coverage and demographic representation. Moreover, the perfect performance metrics suggest that further testing is essential through cross-validation, external validation datasets**,** and inclusion of more varied patient populations. Looking forward, future work could expand the model's scope by integrating image-based data using deep learning, employing ensemble classifiers like (Random Forests**)** and incorporating explainable AI techniques to improve clinical trust. This study lays a strong foundation for building AI-assisted diagnostic tools that are effective, especially in under-resourced healthcare environments.

Reference

1. Marchetti, M. A., Liopyris, K., Dusza, S. W., Codella, N. C. F., Gutman, D. A., Helba, B., … Halpern, A. C. (2020). Diagnostic accuracy of computer algorithms vs dermatologists for cutaneous melanoma: Results from the ISIC 2017⁄2019 challenge. Journal of the American Academy of Dermatology, 82(3), 622–627.
2. Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., … von Kalle, C. (2020). Deep neural networks are superior to dermatologists in melanoma image classification. European Journal of Cancer, 119, 11–17.
3. Bellomo, D., et al. (2020). Machine-learning logistic regression combining clinicopathologic and gene-expression variables predicts sentinel-lymph-node metastasis in melanoma. JCO Precision Oncology.
4. Arora, A., et al. (2020). Multiple machine learning algorithms outperform clinicopathologic models in predicting overall survival from TCGA melanoma RNA expression data. Bioinformatics, 37(12), 1751–1761.
5. Eguchi, R. T., et al. (2023). Prognostic modeling of cutaneous melanoma stage I patients using registry data: Logistic regression identifies risk subsets. Cancer, 129(3).
6. Jia, G., Song, Z., Xu, Z., et al. (2021). Gene marker screening for metastatic melanoma prognosis using logistic regression. BMC Medical Genomics, 14, 96.
7. Wong, Y., et al. (2022). Dermotonet/CNN fusion improves melanoma diagnosis accuracy over clinicians. JMIR Dermatology, 4, e39113.
8. Lu, X., Liu, S., & Weinberger, K. (2021). XGBoost with 5-methylcytosine signatures for cutaneous melanoma prognosis. Cancers, 13(7), 2179.
9. Hekler, A., et al. (2022). Automated melanoma detection via ResNet50 using HAM10000 and ISIC datasets. Life, 14(12), 1602.
10. Rezk, A., et al. (2024). Comprehensive review of AI in skin cancer diagnosis. Healthcare Analytics, 100259.
11. Bechelli, S., & Delhommelle, J. (2022). Machine and deep learning algorithms for skin cancer classification from dermoscopic images. Bioengineering, 9(3), 97.
12. Acosta, A., Zhang, X., Wang, S., Liu, J., & Tao, C. (2024). Advancements in skin cancer classification: ML techniques in clinical image analysis. Multimedia Tools and Applications.
13. Krohling, B., Castro, P. B. C., Pacheco, A. G. C., & Krohling, R. A. (2021). A smartphone-based application for skin cancer classification using deep learning with clinical images and lesion information. arXiv preprint.

14. Loss, F. P., et al. (2024). Skin cancer diagnosis using NIR spectroscopy and ML algorithms. Scientific Reports, 14, 45567.
15. Pacheco, A. G. C., & Krohling, R. A. (2019). The impact of patient clinical information on automated skin cancer detection. arXiv preprint.
16. Rotemberg, V., et al. (2020). A patient-centric dataset of images and metadata for identifying melanomas. arXiv preprint.
17. Acosta, A., Tao, C., Zhang, X., et al. (2023). MiSC: A hybrid multimodal model integrating metadata and images for skin cancer detection with MobileNetV2 and logistic regression. Multimedia Tools and Applications.
18. Giotis, A., et al. (2023). Skin cancer classification using machine learning on digital images: The MED-NODE system. Multimedia Tools and Applications.
19. Behara, K., Bhero, E., & Agee, J. T. (2024). AI in dermatology: A comprehensive review into skin cancer detection. PeerJ Computer Science, 10, e2530.