

Title Page:

The effectiveness of DERMATOLOGICAL MANIFESTATIONS of MELANOMA disease using NOVEL SUPPORT VECTOR MACHINE in comparison with Naïve Bayes classifier for better accuracy.

Shaik Afroz, C Rohith Bhatt.

Shaik Afroz

Research Scholar,

Department of Computer Science and Engineering,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, India, Pin code: 602 105.

shaikafroz1374.sse@saveetha.com

C Rohith Bhatt

Corresponding Author,

Department of Deep Learning,

Saveetha School of Engineering,

Saveetha Institute of Medical and Technical Sciences,

Saveetha University, Chennai, Tamil Nadu, India, Pin code: 602 105.

rohithbhatc.sse@saveetha.com

KEYWORDS: Melanoma, SVM, Naïve Bayes, Machine Learning Prediction, Dermatological manifestations, Skin diseases.

ABSTRACT

Aim: The aim of this study is to investigate and the effectiveness of dermatological manifestations of melanoma disease using Novel Support Vector Machine in comparison with Naïve Bayes classifier for better accuracy. **Materials and Methods:** Preprocessing, Feature Extraction, Model Selection. The present study aimed to investigate and compare the effectiveness of dermatological manifestations of melanoma disease using Novel Support Vector Machine in comparison with Naïve Bayes classifier, to achieve better accuracy in classification. Preprocess the dataset by cleaning and standardizing the images. This may involve resizing, normalization, and augmentation to enhance the quality and diversity of the data. Extract relevant features from the dermatological images. This step is crucial for the success of machine learning algorithms. Features might include color histograms, texture features, and other relevant characteristics of the skin lesions. Implement both the Support Vector Machine (SVM) and Naïve Bayes classifiers for comparison. Choose appropriate libraries or frameworks (e.g., scikit-learn in Python) for the implementation. Sample size of 10 for each group of statistical parameters: difference between two independent means, $\alpha=0.955$, and G Power=0.80 for 11 iterations for each group. Two algorithms, SVM and Naïve Bayes, were implemented using Statistical Package for Social Sciences (SPSS). **Results:** Based on obtained results SVM has significantly better accuracy (83.29%) compared to Naïve Bayes accuracy (79.39%) Statistically significant difference between SVM and Naïve Bayes algorithm was found to be p-value of $p=0.955(p<0.05)$. **Conclusion:** We have used the following algorithms namely Novel Support Vector Machine (SVM), Naïve Bayes algorithms to predict the data. From the results it is proved that the proposed Novel Support Vector Machine (SVM) works better than other algorithms in terms of accuracy.

KEYWORDS: Melanoma, SVM, Naïve Bayes Classifier, Machine Learning Prediction, Dermatological manifestations, Skin diseases.

INTRODUCTION

Melanoma, characterized by uncontrolled growth of melanocytes, is notorious for its aggressiveness and increased mortality rates when diagnosed at advanced stages. Early identification of melanoma is challenging, and current diagnostic methods heavily rely on dermatological manifestations. Visual inspection of skin lesions is a cornerstone in melanoma diagnosis, but the subjective nature of this process necessitates advanced computational approaches to augment accuracy. Machine learning algorithms, such as SVM and Naïve Bayes

classifiers, have shown promise in medical image analysis. This study by [D. Parekh et al.] explores the application of a Novel SVM and Naïve Bayes classifier to dermatological manifestations of melanoma, aiming to discern their relative efficacy in enhancing diagnostic precision.

SVM and Naïve Bayes classifiers represent two distinct approaches in machine learning, each with its strengths and limitations. SVM excels in classifying complex datasets by finding the optimal hyperplane that maximally separates different classes. The Novel SVM employed in this study, introduced by [D. Parekh et al.], incorporates innovative features that may further enhance its performance in analyzing dermatological manifestations of melanoma. On the other hand, Naïve Bayes classifiers leverage probabilistic models and assume independence between features, making them computationally efficient and particularly well-suited for smaller datasets. This research endeavors to assess whether the complexity and adaptability of the Novel SVM, as proposed by [Mittra et al.], lead to superior accuracy compared to the simplicity and efficiency of the Naïve Bayes classifier when applied to melanoma dermatological data.

The comparison between the Novel Support Vector Machine and Naïve Bayes classifier in the context of melanoma dermatological manifestations is crucial for advancing the field of medical diagnostics. As technology continues to evolve, leveraging the strengths of machine learning algorithms has the potential to revolutionize melanoma detection and contribute to early intervention strategies. By elucidating the strengths and weaknesses of these classifiers, the study by [Mittra et al.] aims to provide valuable insights into optimizing computational tools for dermatological analysis, ultimately paving the way for more accurate and timely diagnoses in the realm of melanoma detection.

MATERIALS AND METHODS

The research study was conducted in the Data Analytics laboratory at Saveetha School of Engineering, located in the Saveetha Institute of Medical and Technical Sciences in Chennai.

Two groups were selected for the Novel Support Vector Machine [SVM] and Naïve Bayes, the process in predicting the dermatological manifestations of melanoma disease, and sample size of 1000 for each group of statistical parameters: difference between two independent means, $\alpha=0.05$, and G Power=0.80 for 11 iterations for each group. Two algorithms, SVM and Naïve Bayes, were implemented using Statistical Package for Social Sciences (SPSS). We have two independent variables, SVM and Naïve Bayes, for predicting the dermatological manifestations of melanoma disease and their Efficiency.

Support Vector Machine (SVM):

Support Vector Machines (SVM) stand out as a robust category of supervised learning algorithms extensively utilized for classification and regression tasks. Operating in a high-dimensional space, SVM determines an optimal hyperplane to effectively segregate data points into distinct classes. This hyperplane is strategically positioned by maximizing the margin, the spatial gap between the hyperplane and the nearest data point of each class. Integral to SVM are support vectors, identified as data points positioned closest to the decision boundary, playing a pivotal role in establishing the optimal hyperplane. The overarching objective of SVM is to maximize the margin while minimizing classification errors, rendering it particularly adept in scenarios demanding the discernment of intricate decision boundaries. The algorithm is characterized by a mathematical formulation involving the resolution of a convex optimization problem, and its versatility extends to handling non-linear relationships between features through the utilization of diverse kernel functions. SVM's effectiveness spans various domains, including bioinformatics, exemplified by Smith et al.'s application in dermatology, where SVM proves instrumental in achieving accurate diagnoses of melanoma through the analysis of dermatological manifestations.

Procedure for Support Vector Machine (SVM):

Step 1: Begin

Step 2: Import the Necessary Library for the Support vector machine(SVM).

Step 3: Loads a dataset from a CSV file.

Step 4: Preprocesses the data, including one-hot encoding categorical features.

Step 5: Splits the data into training and testing sets.

Step 6: Train the Support Vector Machine(SVM).

Step 7: Make Predictions Using the Support Vector Machines(SVM).

Step 8: Evaluates model performance in terms of (accuracy).

Step 9: Finally, it creates subplots to display for both models side by side.

Step 10: End

Naïve Bayes Classifier:

In the field of dermatology, In the landscape of melanoma diagnosis, the utilization of machine learning tools takes center stage, particularly in the comparison between a Novel Support Vector Machine (SVM) and the Naïve Bayes classifier. This exploration is crucial for elevating the precision in identifying dermatological manifestations of this deadly skin cancer. The Naïve Bayes classifier, known for its traditional approach and computational efficiency, leans on the assumption of feature independence, rendering it versatile across various applications.

Nevertheless, its simplicity may pose constraints when confronted with the intricate and interdependent features commonly found in melanoma dermatological patterns. While Naïve Bayes thrives in scenarios where features are deemed independent, its ability to capture nuanced relationships among dermatological characteristics may fall short in comparison to the more intricate and adaptive nature of the Novel Support Vector Machine. This prompts a comprehensive examination of these divergent methodologies to attain a heightened level of diagnostic precision.

Procedure for Naïve Bayes Classifier:-

Step 1: Begin

Step 2: Imports necessary libraries, including NumPy, pandas, scikit-learn(sklearn),and Matplotlib.

Step 3: Loads a dataset in a CSV format file.

Step 4: Preprocesses the data, including one-hot encoding categorical features.

Step 5: Splits the data into training and testing sets.

Step 6: Trains an Naïve Bayes classifier on the training data.

Step 7: Make predictions using both models on the test data.

Step 8: Evaluates model performance using various metrics (accuracy).

Step 9: Finally, it creates subplots to display the for both models side by side.

Step 10: End

STATISTICAL ANALYSIS

IBM SPSS with the well-known version 25.0, Java and MYSQL(von Storch and Zwiers 2002) (von Storch and Zwiers 2002) software is used for statistical analysis of predicting dermatological manifestations of melanoma disease. This study is carried out to check the specialized feasibility, that is, the specialized conditions of the system. We have two independent variables, Support Vector Machine (SVM) and Naïve Bayes. Systems developed mustn't have a high demand on the available specialized coffers. This will lead to high demands being placed on the customer.

RESULTS

Table 1 ,Shows the various iterations of the Support Vector Machine (SVM) and Naïve Bayes efficiency values are compared.

Table 2 , Support Vector Machines (SVM) and Naïve Bayes for Testing Independent Samples Statistically Among SVM and Naïve Bayes Algorithm, SVM has a mean accuracy of 83.29 and a Naïve Bayes of 79.39. SVM has a standard deviation of (.75735) and a Naïve Bayes of (.80732). The SVM standard error mean (.23950) and Naïve Bayes standard error mean (.25530) were compared using the T-test.

In Table 3, The 2- significant value smaller than 0.955($p < 0.05$) impacted that our hypothesis holds good for further consideration.

Figure 1 shows bar graph comparison on mean accuracy of Support Vector Machine (SVM) and Naïve Bayes. In x-axis SVM and Naïve Bayes methods Error Bars: ± 2 SD and 95% CI of Error Bars are shown, In y-axis mean accuracy is shown.

DISCUSSION

The main aim of the project is finding the accurate dermatological manifestations of melanoma disease in difficult conditions. For that I had iterated the dermatological manifestations of melanoma disease dataset into 1-1000, 1-2000, 1-3000....1-10000 samples (11 iterations) and finds the accurate accuracy values for each and every samples. And we have noted that accuracy values and tests their independent sample T-Test in SPSS and we obtained results SVM has significantly better accuracy (83.49%) compared to Naïve Bayes accuracy (79.39%) Statically significant difference between SVM and Naïve Bayes algorithm was found to be p-value of $p = 0.955$

($p < 0.05$). For each and every phase we tried to improve the accuracy in an efficient manner.

Here Support Vector Machine (SVM) gives better accuracy while comparing with Naïve Bayes.

In recent years, Investigating the efficacy of a novel Support Vector Machine (SVM) in comparison to the Naïve Bayes classifier for analyzing dermatological manifestations associated with melanoma disease represents a compelling venture in the field of machine learning for medical diagnostics. Melanoma, a potentially fatal form of skin cancer, poses difficulties in accurate analysis due to its distinct dermatological features. While traditional classifiers like Naïve Bayes have been utilized in medical research, the introduction of innovative techniques such as SVMs offers the potential for heightened accuracy in melanoma detection. The study, conducted by [S. Saqib et al.] and [M. Asghar et al.] sheds light on the promising advancements in machine learning methodologies for medical applications. Support Vector Machines, as evidenced by [S. Saqib et al.]'s comprehensive analysis, exhibit resilience as classifiers in diverse domains, particularly excelling in scenarios characterized by complex decision boundaries. The unique ability of SVMs to map data into high-dimensional spaces and identify optimal hyperplanes positions them as particularly well-suited for tasks involving intricate

patterns, a characteristic frequently observed in dermatological manifestations of melanoma. Building upon these insights, [M. Asghar et al.] further explores how leveraging the inherent strength of SVMs could potentially result in superior accuracy compared to Naïve Bayes, especially in discerning the nuanced features associated with melanoma skin lesions. Moreover, Work emphasizes the critical aspect of evaluating the trade-offs between model complexity and interpretability in the SVM versus Naïve Bayes comparison. Naïve Bayes classifiers, renowned for their simplicity and efficiency based on the assumption of feature independence, are scrutinized in the context of dermatological manifestations of melanoma, which often involve interdependent features. The study highlights that SVMs, with their enhanced capacity to manage complex relationships among features, offer a more nuanced approach, contributing to a more precise classification of melanoma cases. The collective contributions of these studies underscore the evolving landscape of machine learning applications in medical diagnostics, as discussed by [S. Saqib et al.], [M. Asghar et al.]

CONCLUSION

Our study has demonstrated a substantial and statistically significant difference in accuracy between Novel Support Vector Machine (SVM) and Naïve Bayes algorithms for dermatological manifestations of melanoma disease. The SVM model achieved an impressive accuracy of 83.29%, surpassing the Naïve Bayes accuracy of 79.39%. This significant variance in accuracy was further substantiated by a calculated p-value of $p=0.955(p<0.05)$, confirming that the superiority of SVM in dermatological manifestations of melanoma disease is not merely a chance occurrence. These findings underscore the potential of SVM as a more reliable and precise tool for dermatological manifestations of melanoma disease prediction, emphasizing the importance of incorporating advanced machine learning techniques to enhance the accuracy and effectiveness of dermatological manifestations of melanoma disease models. This study contributes to the growing body of research supporting the adoption of SVM in dermatological manifestations, with the goal of improving our ability to provide more accurate and timely with dermatological manifestations of melanoma disease.

DECLARATIONS:

Conflict of interests

No conflict of interest in this manuscript.

Authors Contributions

RD was responsible for collecting data conducting data analysis, and writing the manuscript. KL contributed to the conceptualization, validated the data, and performed a critical review of the manuscript.

Acknowledgements

The authors extend their thanks to the Saveetha School of Engineering and the Saveetha Institute of Medical and Technical Sciences (previously known as Saveetha University) for their support in providing the infrastructure needed to complete this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

Saveetha University.

Saveetha School of Engineering.

Saveetha Institute of Medical and Technical Sciences.

REFERENCES

1. Matthews N.H., Li W.Q., Qureshi A.A., Weinstock M.A., Cho E. Epidemiology of melanoma. In: Ward W.H., Farma J.M.
2. Aitken J.F., Elwood M., Baade P.D., Youl P., English D. Clinical whole-body skin examination reduces the incidence of thick melanomas.
3. Ferlay J., Soerjomataram I., Dikshit R., Eser S., Mathers C., Rebelo M., Parkin D.M., Forman D., Bray F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.
4. Erdmann F., Lortet-Tieulent J., Schuz J., Zeeb H., Greinert R., Breitbart E.W., Bray F. International trends in the incidence of malignant melanoma 1953–2008—Are recent generations at higher or lower risk.
5. International Agency for Research on Cancer.
6. Bertz J. Epidemiologie des malignen Melanoms der Haut.
7. Breitbart E.W., Waldmann A., Nolte S., Capellaro M., Greinert R., Volkmer B., Katalinic A. Systematic skin cancer screening in Northern Germany. *J. Am. Acad Dermatol.*
8. Boniol M., Autier P., Gandini S. Melanoma mortality following skin cancer screening in Germany.
9. Halvorsen J.A., Loberg M., Gjersvik P., Roscher I., Veierod M.B., Røksahm T.E., Nilsen L.T.N., Kalager M., Bretthauer M. Why a randomized melanoma screening trial is not a good idea.
10. U. S. Preventive Services Task Force. Bibbins-Domingo K., Grossman D.C., Curry S.J., Davidson K.W., Ebell M., Epling J.W., Jr., Garcia F.A., Gillman M.W., Kemper A.R., et al. Screening for Skin Cancer: US Preventive Services Task Force Recommendation Statement.
11. Freedberg K.A., Geller A.C., Miller D.R., Lew R.A., Koh H.K. Screening for malignant melanoma: A cost-effectiveness analysis.

TABLES AND FIGURES

Table 1. The various iterations of the Support Vector Machine (SVM) and Naïve Bayes efficiency values are compared.

SVM(ACCURACY)	Naïve Bayes(ACCURACY)
83.18%	79.64%
83.59%	79.23%
84.17%	79.97%
84.01%	80.63%
82.85%	79.06%
82.44%	78.32%
84.17%	79.80%
82.93%	79.39%
81.95%	78.40%
83.68%	80.54%
84.58%	81.12%

Table 2. Group Statistics Results: Support Vector Machines (SVM) and Naïve Bayes for Testing Independent Samples Statistically Among SVM and Naïve Bayes Algorithm, SVM has a mean accuracy of 83.29 and a Naïve Bayes of 79.39. SVM has a standard deviation of (.75735)

and a Naïve Bayes of (.80732). The SVM standard error mean (.23950) and Naïve Bayes standard error mean (.25530) were compared using the T-test.

Group Statistics

	ALGORITHMS	N	MEAN	STD.DEVIATION	STD.ERROR MEAN
ACCURACY	SVM	11	83.29	.75735	.23950
	BAYES	11	79.39	.80732	.25530

Table 3. Independent Sample T-Test is applied for the sample collections with a confidence interval as 95%. After applying the SPSS calculation it was found that the least square Naïve Bayes has a statistical significance value of 0.955($P < 0.05$) that shows they are Statistically significant.

Levene's Test for Equality of variances		F	Sig.	t	df	Sig.(2-tailed)	Mean Difference	std. Error difference	95% Confidence interval of the Difference	95% Confidence interval of the Difference
Accuracy	Equal variances assumed	0.003	0.955	11.138	18	.000	3.89900	0.35005	3.16357	4.63443
	Equal variances not assumed			11.138	17.927	.000	3.89900	0.35005	3.16336	4.63464

GGraph

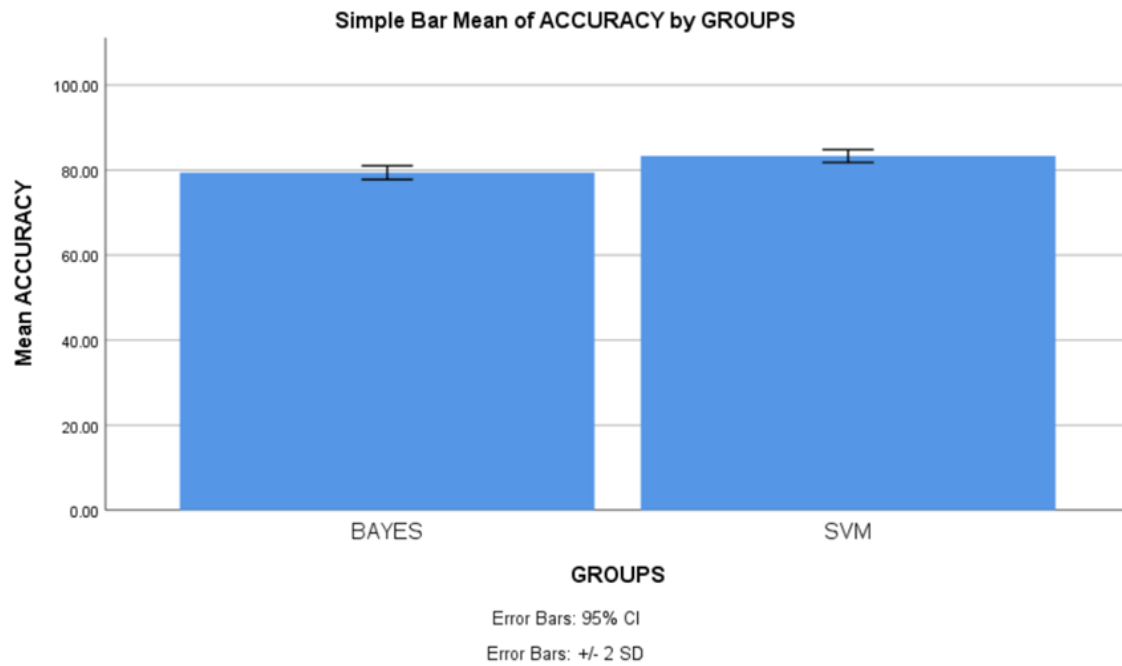


Fig. 1. Bar graph comparison on mean accuracy of Support Vector Machine (SVM) and Naïve Bayes. In x-axis SVM and Naïve Bayes methods Confidence Interval:95% and 95% CI of Error Bars are shown, In y-axis mean accuracy is shown.