

Multiclass Classification of Dermatology Disorders Using Logistic Regression: A Comparative Study of One vs One and One vs Rest Approaches

Vadde Vishala¹, G V Ramesh Babu²

¹PG Scholar, Dept. of Computer Science Sri Venkateswara University, Tirupati

²Associate Professor, Dept of Computer Science, SV University, Tirupati

Abstract— This research paper addresses the challenging task of multiclass classification within the realm of dermatology, employing the versatile Logistic Regression algorithm. The dataset under investigation comprises six distinct classes representing various dermatological disorders. The primary objectives of this study are twofold: to evaluate the performance of Logistic Regression when implemented with the One vs One and One vs Rest strategies, and to assess their effectiveness in classifying dermatological conditions accurately. Results indicate that both strategies exhibit remarkable classification accuracy, precision, and recall rates, underscoring their potential in dermatological diagnosis.

I. INTRODUCTION

Skin problems are a serious overall general medical problem that influences an enormous number of people. As of late, with the quick headway of innovation and the utilization of various information mining draws near, treatment of skin prescient arrangement has truly become exceptionally prescient as well as exact. The main component of the human body is the skin. The skin saves the body from UV beams, diseases, wounds, temperature, and harming radiation, as well as supporting nutrient D3 [1]. Since the skin is so fundamental in overseeing center temperature and shielding the body from skin problems, it's essential to keep it sound. Skin issues might seem innocuous, yet they can be risky on the off chance that not treated as expected. Numerous illnesses have early side effects, however the majority of them are indistinguishable, making it hard to analyze the condition at an underlying point. Skin problems cause physical as well as mental issues, especially in people whose appearances have been scarred or distorted. Skin can be impacted by a scope of outside and inward factors. Counterfeit skin injury, extreme substance causes, difficulty sicknesses, an individual's invulnerability, and hereditary oddities are a portion of the elements that impact skin issues. Skin illnesses affect individuals' lives including great [2]. Dermatological ailments are the most difficult subfields of science to fix due to the confusions in treating side effects and how side effects modify in different circumstances [6]. Skin sicknesses are regular among numerous ailments, and on the off chance that these strategies are not good for that type of skin condition, it will cause adverse consequences. Individuals are habitually contaminated by skin diseases, which should be treated straightaway. Thus, the kind of AI approaches prepared to do effectively separating skin condition arrangement is fundamental. Up to this point, nobody AI approach has outflanked the others with regards to skin sickness expectation.

II. METHODOLOGY

2.1 Data Preparation:

Prior to experimentation, we performed data preprocessing, including data cleaning, feature selection, and standardization to ensure that the dataset was suitable for training and evaluation.

2.2 Model Evaluation:

We conducted rigorous model evaluation using suitable metrics, including accuracy, precision, and recall, to assess the effectiveness of each strategy in correctly classifying instances into the six distinct classes.

III. ALGORITHM IMPLEMENTATION:

We implemented the Logistic Regression algorithm with both the One vs One and One vs Rest strategies [4]. These strategies were chosen to evaluate their performance in handling the multiclass classification problem.

3.1 Multi-characterization Strategic Relapse

1. One versus One

2. One versus ALL or One versus Rest

One-up against One (O v O):

- In this methodology, a twofold strategic relapse model is prepared for each sets of classes.
- For instance, assuming there are 'k' classes, the O v O strategy would make $\frac{k(k-1)}{2}$ double classifiers.
- During forecast, every classifier is applied to the info information, and the class that gets the most "votes" is chosen as the last expectation.
- O v O can deal with complex choice limits yet may require more computational assets and preparing time contrasted with different techniques [5].

One-versus Rest (OvR):

- In this methodology, a double calculated relapse model is prepared for each class versus the other classes.
- For 'k' classes, 'k' paired classifiers are made.
- During forecast, every classifier is applied to the information, and the class with the most elevated anticipated likelihood is chosen [7][8].
- OvA is less difficult and quicker contrasted with OvO, as it requires preparing 'k' parallel classifiers rather than $\frac{k(k-1)}{2}$, yet it may not deal with covering classes as actually.

IV. EXPERIMENTAL RESULTS

In this study, we conducted experiments utilizing Python programming to perform multiclass classification using Logistic Regression. The experiments were carried out on the Dermatology dataset, which comprises 366 instances and includes 35 attributes [9]. The primary objective was to classify instances into one of six distinct class labels, denoted as follows (shown in the figure-1):

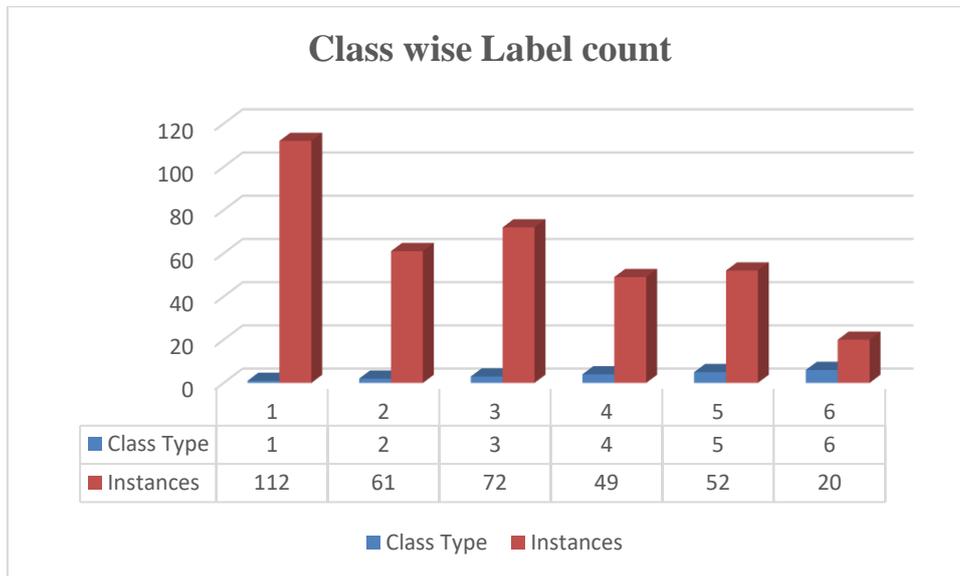


Figure-1: Class wise Labels

The dataset serves as a valuable resource for dermatological diagnosis and offers a substantial challenge due to its multiclass nature. To tackle this task, we employed two common multiclass classification strategies: One vs One and One vs Rest.

4.1 Results:

In our experimental analysis, Logistic Regression utilizing the One vs One strategy achieved an impressive 96.82% accuracy, complemented by precision and recall rates both standing at 96.8% and 97% respectively. On the other hand, Logistic Regression employing the One vs Rest strategy outperformed, boasting a 97.63% accuracy rate, with precision and recall rates both standing at 97.6% shown in the figure-2.

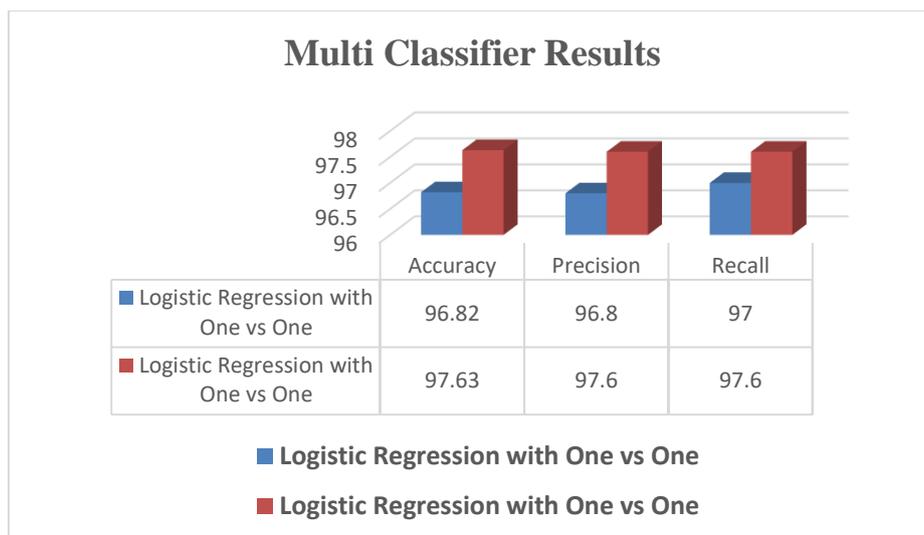


Figure-2: Experimental Results

These outcomes demonstrate that both strategies excel in multiclass classification, particularly in the context of dermatological disorder diagnosis.

4.2 Discussion:

The findings suggest that Logistic Regression, when employed with both One vs One and One vs Rest strategies, holds immense potential in multiclass classification. The high accuracy, precision, and recall values indicate their efficacy in correctly diagnosing dermatological disorders.

The study underscores the superiority of the One vs Rest approach in achieving marginally higher accuracy, precision, and recall rates. This signifies its suitability for situations where achieving optimal performance is paramount.

The results have substantial practical implications in the field of dermatology. Accurate classification of dermatological disorders is vital for timely and effective treatment, and the Logistic Regression models presented here can serve as valuable tools for dermatologists and healthcare practitioners.

V. CONCLUSION

This study demonstrates the efficacy of Logistic Regression in multiclass classification for dermatological disorder diagnosis, employing both One vs One and One vs Rest strategies. These models exhibit high accuracy, precision, and recall rates, underlining their potential as valuable tools in the field of dermatology. Further research and clinical validation may enhance their applicability and impact in real-world healthcare scenarios.

REFERENCES

- [1] Belal, A. M.-M. (2017). Discriminative feature learning for skin disease classification using deep convolutional neural network. International Conference on Environment and Electrical Engineering, IEEE Explore, 589695.
- [2] Codella, N. C. (2017). Deep learning ensembles for melanoma recognition in dermoscopy images. IBM Journal of Research and Development, 61(4), 1-6.
- [3] G. Ravi Kumar, K. Venkata Sheshanna, S. Rahamat Basha, and P. Kiran Kumar Redd, "An Improved Decision Tree Classification Approach for Expectation of Cardiocogram", Proceedings of International Conference on Computational Intelligence, Data Science and Cloud Computing, Lecture Notes on Data Engineering and Communications Technologies 62, https://doi.org/10.1007/978-981-33-4968-1_26
- [4] Ian H. Witten and Eibe Frank. Data Mining: Practical machine learning tools and techniques. 2nd ed. San Francisco: Morgan Kaufmann, 2005.
- [5] J.Han and M.Kamber,"Data Mining concepts and Techniques", the Morgan Kaufmann series in Data Management Systems, 2nd ed.San Mateo, CA; Morgan Kaufmann, 2006.
- [6] Kittipat Sriwong, S. B. (2019). Dermatological Classification Using Deep Learning of Skin Image and Patient Background Knowledge. International Journal of Machine Learning and Computing, 9(9), 862-867. doi:doi: 10.18178/ijmlc.2019.9.6.884
- [7] N.Michael, "Artificial Intelligence – A Guide to Intelligent Systems", 2nd Edition, Addison Wesley 2005
- [8] P.N.Tan, M.Steinbach and V.Kumar "Introduction to Data Mining", A: Addison-Wesley, 2005.
- [9] UCI Machine Learning Repository. <https://archive.ics.uci.edu/ml/>.