Machine Learning Approaches to Multi-Class Human Skin Disease Detection

Asim Saket Samad[1], Mukund Srivastava[2], Souvik Banerjee[3]
[1,2,3]Department of Information and Technology
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY, CHENNAI, INDIA.

***ABSTRACT*- Human Cancer is a standout amongst the most perilous sickness which is fundamentally brought about by hereditary shakiness of various sub-atomic modifications. Among numerous types of human malignant growth, skin disease is the most widely recognized one. To distinguish skin disease at a beginning time we will ponder and break down them through different procedures named as division and highlight extraction. Here, we center harmful melanoma skin disease, (because of the high centralization of Melanoma-Hier we offer our skin, in the dermis layer of the skin) identification. In this, we utilized our ABCD rule dermoscopy innovation for dangerous melanoma skin disease identification. In this framework diverse advance for skin injury portrayal i.e, first the Image Acquisition Technique, pre-handling, division, characterize include for skin Feature Selection decides sore portrayal, characterization strategies. In the Feature extraction by computerized picture handling technique incorporates GLCM and ABCDE highlights and furthermore we utilized DRLBP. Here we proposed the Recurrent Neural Network to group the kind of ailment.**

I. INTRODUCTION

In a country like INDIA, skin diseases are increasing at a very fast rate. There are many factors like economy, differences in climate, illiteracy , personal hygiene issues, lack of awareness, social backwardness, dearth of primary healthcare centres in rural area, pollution and life style in urban area that contribute to the growth of skin related diseases. The prevalence of skin diseases in the general population has varied from 7.86% to 11.16% in various studies. Due to discolouring and other problems with these diseases, the sufferers face many challenges physically and socially [4].Many skin diseases need screening by an expert dermatologist. Normally, manual diagnosis by experts is subjective and is based on the personal expertise in the area. To achieve objective accuracy of diagnosis, computer aided diagnosis is normally used.

With the advancement in computing technologies and digitization of medical field, computer aided diagnosis has become a reality. Machine learning is widely in the field of medical diagnosis.

II. MACHINE LEARNING IN CANCER DETECTION

Machine Learning in Cancer detection has helped in a more effective way than any other available methods. The last two decades a variety of different ML techniques and feature selection algorithms have been widely applied
to disease prognosis and prediction. When dealing with cancer prognosis/prediction one is concerned with three predictive tasks: (i) the prediction of cancer susceptibility (risk assessment), (ii) the prediction of cancer recurrence/local control and (iii) the prediction of cancer survival. In the first two cases one is trying to find (i) the likelihood of developing a type of cancer and (ii) the likelihood of redeveloping a type of cancer after complete or partial remission. In the last case, the prediction of a survival outcome such as disease-specific or overall survival after cancer diagnosis or treatment is the main objective. The prediction of cancer outcome usually refers to the cases of (i) life expectancy, (ii) survivability, (iii) progression and (iv) Treatment sensitivity.
Major types of ML techniques including ANNs (Artificial Neural network) and DTs (Decision Tree) have been used for nearly three decades in cancer detection. In many publications most of them use one or more ML algorithms and integrates data from heterogeneous sources for the detection of tumours as well as for the prediction/prognosis of a cancer type. A growing trend is noted the last decade in the use of other supervised learning techniques, namely SVMs (Support Vector Machine) and BNs (Bayesian Network), towards cancer prediction and prognosis.

III. METHODOLOGY

General methodology of the proposed system is depicted fig. 1. The system can be broadly categorized into data collection, pre-processing, feature extraction and classification modules.

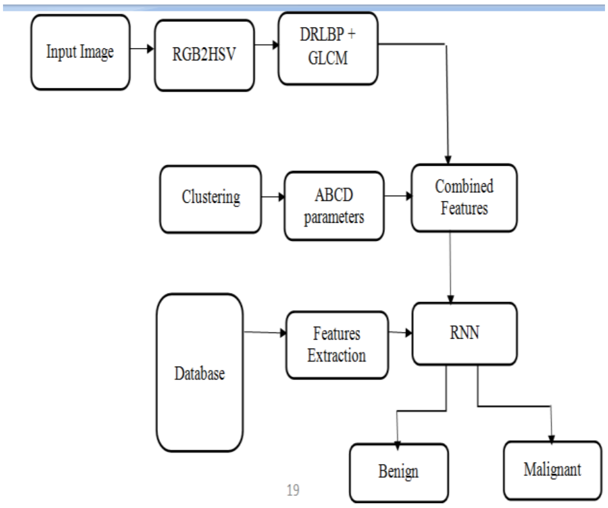

Fig.1

Perception Tasks: Deep neural networks have boosted the performance of computers at perception tasks to previously unimaginable levels. It is very likely that the performance of machine learning algorithms, in the areas where perception tasks are central, will improve further and both enhance the productivity of physicians and improve the level of healthcare. For example in radiology, where the task of the physician is to diagnose a patient using medical imaging, computers have been taught to identify pathologies from such images at either a comparable or even higher level than human doctors.

Diagnostic Assistance: Machine learning has been used to augment the physician's capacity by looking at the total available data on the patient and making recommendations based on this information - This approach is only limited by the total number of patients treated in the world. Diagnostic assistance is used for assisting the physicians in real time by either data retrieval or diagnosis recommendations. Multiple companies also offer a prediction service for identifying most at-risk patients.

Treatment process: As an important part of healthcare is also the process a patient goes through and how they are treated. Improvements in this process can produce gains in both the quality and cost of care. For this end, a branch of data science called process mining has been utilized with quite a bit of success.

IV. K-MEANS CLUSTERING

K-implies bunching is a kind of unsupervised realizing, which is utilized when you have unlabeled information (i.e., information without characterized classifications or gatherings). The objective of this calculation is to discover bunches in the information, with the quantity of gatherings spoken to by the variable K. The calculation works iteratively to appoint every datum point to one of K bunches dependent on the highlights that are given. Information focuses are bunched dependent on highlight comparability. The after-effects of the K-implies grouping calculation are:

1. The centroids of the K bunches, which can be utilized to mark new information
2. Labels for the preparation information (every datum point is allocated to a solitary bunch)

Instead of characterizing bunches before taking a gander at the information, bunching enables you to discover and examine the gatherings that have framed naturally. The "Picking K" segment underneath depicts how the quantity of gatherings can be resolved.

Every centroid of a bunch is an accumulation of highlight esteems which characterize the subsequent gatherings. Inspecting the centroid include loads can be utilized to subjectively decipher what sort of gathering each bunch speaks to.

The K-implies grouping calculation utilizes iterative refinement to create a last outcome. The calculation inputs are the quantity of bunches K and the informational index. The informational index is an accumulation of highlights for every datum point. The calculations begin with introductory evaluations for the K centroids, which can either be haphazardly created or arbitrarily chosen from the informational collection.

V. RESULTS

We trained individual classifiers mentioned above on training symptoms dataset. A neural network is trained on a dataset to predict possible disease as an output in R with sigmoid activation function. The architecture of neural network is 13-0-10 with 13 input neurons, 10 output neurons and no neurons hidden layer as shown in fig 2 which becomes a simple logistic regression.

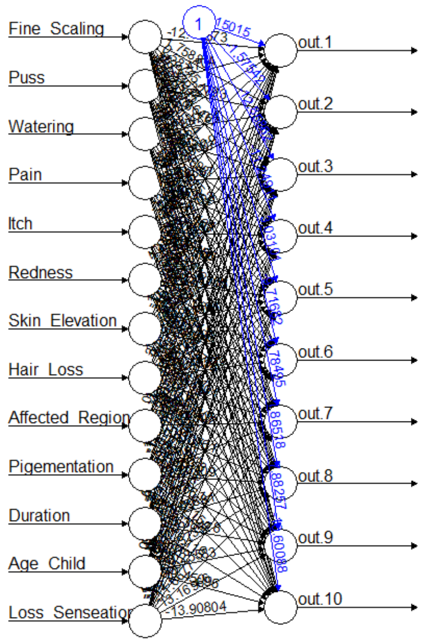

Fig. 2. Neural Network Architecture

Fig. 3 shows decision tree on the training dataset. All leaf nodes represent the disease class. As can be seen highest importance is given to redness of skin lesion being a root node of the tree.

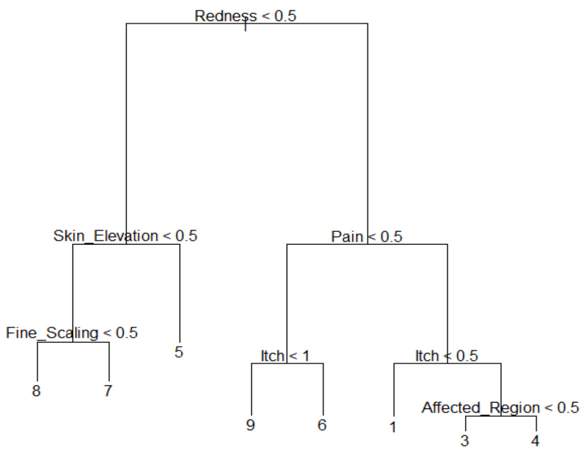

Fig 3 Decision tree on training dataset

Performance of a classifier model is measured in terms of precision, recall and accuracy. Accuracies of these classifiers on validation dataset which are calculated by equation 1 are summarized in table 3

We also developed a simple GUI which can be used for detection of disease for new patient as shown in fig. 5. The system is verified successfully by concerned doctor in the hospital.

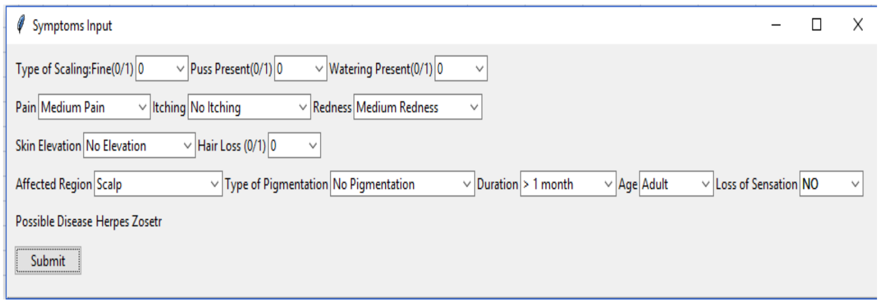

Fig 5.

VI. FUTURE ENHANCEMENT OF MACHINE LEARNING

The future of machine learning might not only help forecast potential issues and diseases, but it can also suggest what steps a patient should take and how medical staff should react. ML in healthcare could also offer "sophisticated clinical decision support," by automatically adding patient data and finding the most effective treatment plan. Machine learning and AI will also become commonplace in healthcare administration for tasks such as adding information to EHRs and billing patients. Meanwhile, natural language processing tools could help physicians to easily create patient medical files by verbally discussing an examination with a patient as it is happening. After the exam is complete, an automated system can send a bill to the patient's Smartphone immediately. As these technologies advance, the use of Machine Learning in healthcare might eventually spur the democratization of healthcare -- making care more

Table 4. Accuracy of Classifiers Sr. No.	**Classifier Model**	**Accuracy**
1.	ANN	100%
2.	KNN	98%
3.	SVM	99%
4.	Decision Tree	97.7%
5.	Random Forest	100%

accessible to more people -- and it could help make care more affordable.

VII. CONCLUSION

In this paper, we discussed different classifiers for skin disease detection. Due to digitalization and improvement in technology, more and more detailed data is becoming available. Medical data mining can be used for diagnosis, and decision making, etc. We collected original data related to 10 common skin diseases and applied five different classifiers for classification of an input record to one of these 10 diseases. Results indicate these classifiers do well in identification of disease. This system can help new medical practitioners to correctly identify disease in case of doubt.

REFERENCES

[1] Jain S, Barambhe M S, Jain J, Jajoo U N, Pandey N. Prevalence of skin diseases in rural Central India: A community-based, cross-sectional, observational study. J Mahatma Gandhi Inst Med Sci 2016; 21:111-5

[2] Grover S, Ranyal RK, Bedi MK. A cross section of skin diseases in rural Allahabad. Indian J Dermatol 2008; 53:179-81

[3] Rao GS, Kumar SS, Sandhya. Pattern of skin diseases in an Indian village. Indian J Med Sci 2003; 57:108-10.

[4] G. K. Jana, A. Gupta, A. Das, R. Tripathy, and P. Sahoo. Herbal treatment to skin diseases: A global approach. Drug Invention Today, 2(8):381–384, August 2010.

[5] E. Barati at el, A Survey on Utilization of Data Mining Approaches for Dermatological (Skin) Diseases Prediction, Journal of Selected Areas in Health Informatics (JSHI): March Edition, 2011

[6] Kwetishe Danjuma, Adenike Osofisan, Evaluation of Predictive Data Mining Algorithms in Erythemato-Squamous Disease Diagnosis, IJCSI International Journal of Computer Science Issues, 11(6), 85-94 (2014)arXiv:1501.00607v1 [cs.LG]

[7] Sunday Olusanya Olatunji and Hossain Arif, Identification Of Erythemato-Squamous Skin Diseases Using Extreme Learning Machine And Artificial Neural Network, ICTACT Journal On Soft Computing, October 2013, Volume: 04, Issue: 01

[8] Haryanto et al, The Erythemato-Squamous Dermatology Diseases Severity Determination using Self-Organization Map, IPTEK, Journal of Proceeding Series, Vol. 1, 2014 pp 279-284

[9] Lejla et al, GA-ANFIS Expert System Prototype forPrediction of Dermatological Diseases, Europian Federation of Medical Informatics(EFMI) 2015, doi:10.3233/978-1-61499-512-8-622

[10] Archana Singh, Ajay Rana and Shipra Srivastava, Comparative Analysis and Automatic detection of various skin diseases and Implementation of leucoderma using support vector machine, International Journal of Computer Applications (0975 – 8887) Volume 48– No.24, June 2012

[11] Dr. Trupti, P. Shah, Pooja J. Shah, Connectionist Expert System for Medical Diagnosis using ANN– A case study of skin disease Scabies, International Journal of Advanced Research in Computer Science and Software Engineering 3(8), August - 2013, pp. 227-230

[12] Nidhal K. Abbadi, Nizar Saadi Dahir, Muhsin A. Al-Dhalimi and Hind Restom, Psoriasis Detection Using Skin Color and Texture Features, Journal of Computer Science 6 (6): 48-652, 2010