Title	Prediction of NB-UVB phototherapy treatment response of psoriasis patients using data mining
Authors(s)	Mohamed, Sharifa, Huang, Bingquan, Kechadi, Tahar
Publication date	2017-12-16
Publication information	Mohamed, Sharifa, Bingquan Huang, and Tahar Kechadi. "Prediction of NB-UVB Phototherapy Treatment Response of Psoriasis Patients Using Data Mining." IEEE, 2017.
Conference details	IEEE International Conference on Bioinformatics and Biomedicine, (BIBM-BHI 2017), Kansas, MO, USA, November 13-16, 2017
Publisher	IEEE
Item record/more information	http://hdl.handle.net/10197/9112
Publisher's statement	© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Prediction of NB-UVB Phototherapy Treatment Response of Psoriasis Patients using Data mining

S. Mohamed, B. Q. Huang, M-T. Kechadi
Insight Centre for Data Analytics, University College Dublin,
Email : sharifa.mohamed@ucdconnect.ie

Abstract—**NB-UVB Phototherapy is one of the most common treatments administrated by dermatologists for psoriasis patients. Although in general, the treatment results in improving the condition, it also can worsen it. If a model can predict the treatment response before hand, the dermatologists can adjust the treatment accordingly. In this paper, we use data mining techniques and conduct four experiments. The best performance of all four experiments was obtained by the stacked classifier made of hyper parameter tuned Random Forest, kSVM and ANN base learners, learned using L1-Regularized Logistic Regression super learner.**

Keywords: Dermatology, Phototherapy, Treatment Response, Datamining, Prediction, Classification

I. INTRODUCTION

Psoriasis is a chronic, immune-mediated inflammatory disease of the skin [1] which have affected approximately 120,000 people in Ireland and 125 million worldwide [2]. Patients with severe psoriasis have high tendency of also being affected by psoriatic arthritis, cardiovascular disease, hypertension, obesity, diabetes, other immune-related ailments such as Crohns disease, and an excess risk of mortality [3]. These patients suffer physically, emotionally and socially because of this disease and have reported reduction in physical functioning and mental functioning comparable to that seen in cancer, arthritis, hypertension, heart disease, diabetes, and depression [4].

Phototherapy involves repeated exposure of the skin to ultraviolet (UV) light to treat various inflammatory skin conditions such as psoriasis, eczema and vitiligo. This therapy is one of the oldest treatment modalities in dermatology, dating back to the ancient Egyptians, who used natural light in combination with herbal extracts to treat skin diseases [5], [8]. Phototherapy continues to be a highly preferred treatment by dermatologists [6], [8].

Mainly three types of phototherapy treatments are used for the treatment of psoriasis: broadband ultraviolet B (BB-UVB), narrowband ultraviolet B (NB-UVB) and psoralen plus ultraviolet A (PUVA) [8]. UVB is most commonly preferred as a first-line treatment for moderate-to-severe psoriasis in healthy adults [6]. NB-UVB is the most commonly used phototherapy modality today as it has a wider application across various dermatologic conditions, it is easier to use and has fewer adverse events when compared with BB-UVB or PUVA [7], [8].

Although phototherapy treatment is effective in treating skin diseases such as psoriasis [9], [10], it also can cause flare ups and worsening of the disease [11]. If the clinicians know the treatment response beforehand, it would enable them to adjust the treatments appropriately if required. This can help to avoid the additional physical and emotional suffering which may cause by worsening of the disease.

Exploiting stored data to extract previously hidden but useful and actionable information, predicting future trends and behaviors are the overall goals of the generic process referred as "data mining" [12]. Recently, data mining techniques have been successfully applied in health-care [13], [14], [15] and particularly in dermatology [16]. However, there is no literature which introduces the prediction of phototherapy treatment response using data mining techniques.

This paper utilizes classification [20] models in order to predict the treatment response of psoriasis patients treated by NB-UVB phototherapy. The proposed prediction model first selects a number of attributes, then prepares data, finally four experiments were conducted to find out the best model to predict the treatment response of NB-UVB phototherapy.

II. METHODOLOGY

A. Data-set Collection

The data set used in this paper was obtained from University Hospital Limerick (UHL) phototherapy database. The data set was professionally anonymized by OpenApp Computer Support and Services after obtaining ethical approval from UHL research ethics committee.

The NB-UVB phototherapy database consisted of 9083 treatment records of 400 psoriasis patients treated since June 2000. There were 221 females & 179 males among the patients. The information of each patient includes patient's personal details (e.g. Gender, year of birth, skin type) and treatment details such as treatment date, Minimal Erythemal Dose(MED), acitretin given or not, if its ReUVB(retinoid plus UVB) or not, dose measured using mjcm-2, exposure time in ms, stool used or not, visor used or not,erythema grade, erythema skin site, pruritus grade, pruritus skin site, treatment response grade etc .

Table 1 shows the number of treatment records belonged to each response grade of treatment.

B. Methods

The prediction model of treatment response grade consists of the preprocessing and classification processes. The preprocessing process first selects significant attributes from the

Response grade	No. of records
No change	3299
Improved	5488
Cleared	176
Worsened	120

data-set, then filters the noise data and normalises data. The classification process applies modeling algorithms to predict the response grade. Following section further describes them.

1) Preprocessing: The goal of this phase is to provide cleaned data for the classification step. In order to achieve that, 1) first, we derived new attributes from the data-set and applied the information gain technique [17] to select the significant attributes or features 2) then imputed the missing values based on domain knowledge or mode of the attribute, 3) next used Local Outlier Factor technique (LOF) [18] to deal with the local outliers 4) finally normalized the data-set; For normalization, new binary attributes were created for categorical attributes; numerical attribute values were scaled to fall between 0 and 1 using min max normalization technique [19].

The features selected by information gain technique for the classification process are treatment number, dosage, cumulative dosage, cumulative exposure time, previous dosage, ratio between dosage and MED, difference between current and previous treatment and percentage of difference between current and previous treatment.

2) Classification: Prediction of response grade of NB-UVB phototherapy treatment is a multiclass problem. Therefore ANN (Artificial Neural network), C5.0 decision tree, K-NN (K Nearest neighbors classifier), kSVM (Support Vector Machines)& Random Forest (an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees)classifiers that support multiclass classification were selected to conduct experiments. In addition to the above mentioned classifiers, L1-Regularized Logistic Regression classifier was used as the super learner when creating stacked classifiers.

Four experiments were run in order to find out the best classifier model. First experiment was conducted to evaluate the prediction performance of the classifiers when used with default settings (See Figure 1).

Second experiment was conducted to evaluate the prediction performance of the classifiers when hyperparameters were tuned (See Figure 2 and Table 2).

The third experiment evaluates prediction performance of stacked classifiers of size three made of base classifiers with default settings (See Figure 3).

The last experiment evaluates prediction performance of stacked classifiers of size three, made of combinations of hyperparameter tuned base learners (See Figure 4).

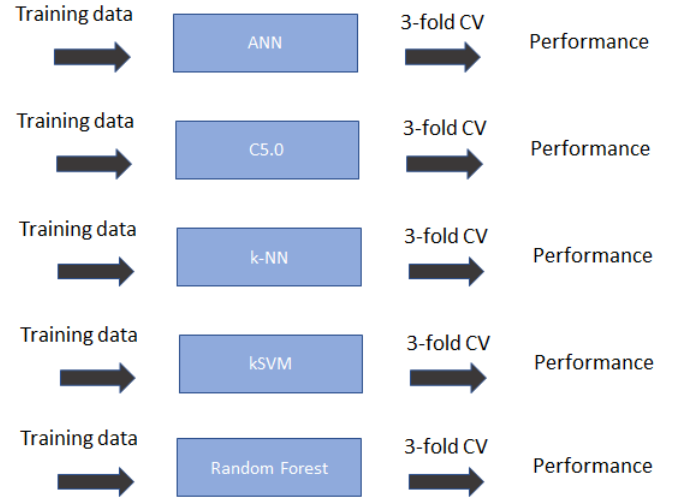In both experiments 3 & 4, L1-Regularized Logistic Regression classifier was used as the super learner.

Fig. 1. Experiment 1 - Evaluation of prediction performance of the classifiers under default settings
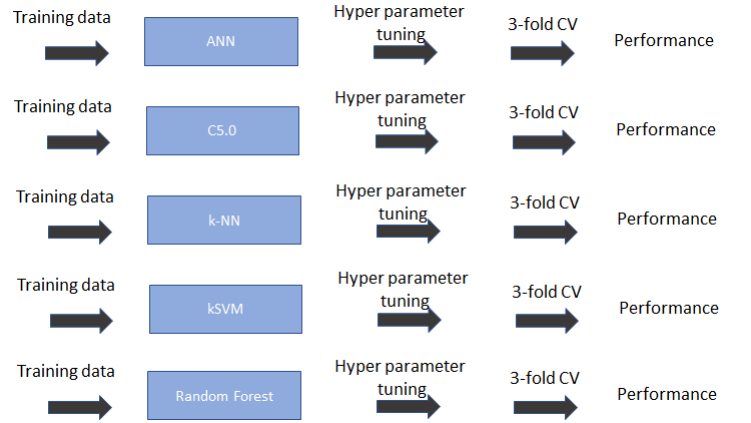

Fig. 2. Experiment 2 - Evaluation of prediction performance of the classifiers after tuning hyperparameters

As the distribution of the response grade attribute was imbalanced, simple oversampling technique was used to balance the data set [23] before conducting experiments and 3 fold cross validation was used to evaluate the performance.

III. RESULTS AND DISCUSSION

This paper uses accuracy, multiclass.au1u and multiclass.brier to evaluate the results. Accuracy is the percentage of correctly classified instances. Multiclass.au1u takes the Average 1 vs. 1 multiclass AUC. It computes AUC of c(c - 1) binary classifiers (all possible pairwise combinations) while considering uniform distribution of the classes [24]. multiclass.brier is defined as: $(1/n) \sum_i \sum_j (y_{ij} - p_{ij})^2$ where $y_{ij} = 1$ if observation i has class j (else 0), and p_{ij} is the predicted probability of observation i for class j [25]. Range of multiclass.brier is from 0 to 2 where a value close to 0 is better.

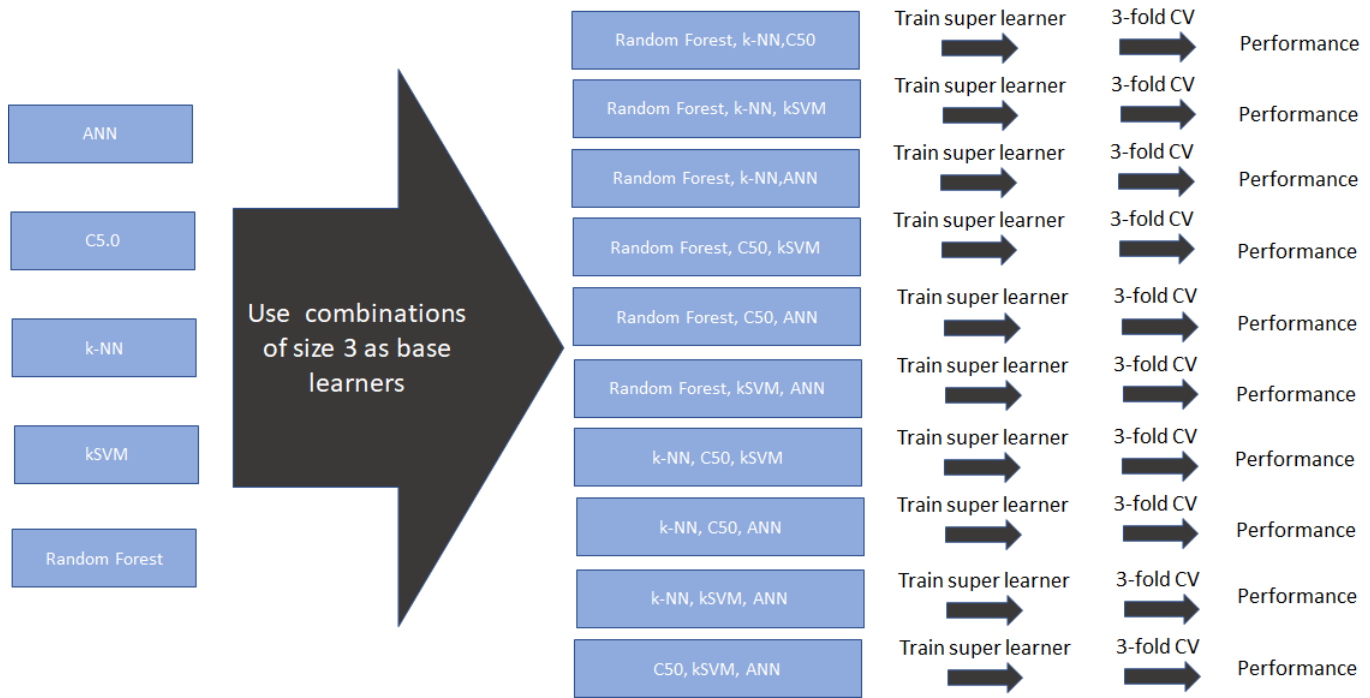The experimental results of this paper are presented in Tables III to VI.

Fig. 3. Experiment 3 - Evaluation of prediction performance of the stacked classifier with default parameters

TABLE II
HYPERPRAMETER TUNING SETTINGS FOR EXPERIMENT 2

Classifier	Param name	Param type	Param descrption	Tuned values
Random forest	ntree	integer	Number of trees to grow.	5-1000
	mtry	integer	Number of variables randomly sampled as candidates at each split.	1-9
k-NN	K	integer	Number of nearest neighbours to be used	1-250
C50	winnow	logical	Should predictor winnowing (i.e feature selection) be used?	True, False
	noGlobalPruning	logical	Global pruning should be applied or not	True, False
kSVM	kernel	discrete	The kernel function used in training and predicting.	"rbfdot" , "polydot" , "tanhdot" , "laplacedot, "besseldot","anovadot"
	scale	numeric	Used with "tanhdot" and "polydot" kernels	1-10
	offset	numeric	Used with "tanhdot" and "polydot" kernels	1-10
	sigma	numeric	Used with "besseldot", "anovadot", "rbfdot" and "laplacedot" kernels	1-10
	degree	integer	Used with "besseldot", "anovadot" and "Polydot" kernels	1-6
	order	integer	Used with "besseldot" kernel	1-10
ANN	maxit	integer	maximum number of iterations	10-1000
	size	integer	number of units in the hidden layer.	2-50

TABLE III
PERFORMANCE OF EXPERIMENT 1

Classifier	Accuracy	multiclass.au1u	multiclass.brier
C5.0	0.885	0.966	0.183
k-NN	0.916	0.969	0.160
ANN	0.586	0.837	0.529
Random Forest	**0.934**	**0.987**	**0.113**
kSVM	0.695	0.890	0.466

TABLE IV
PERFORMANCE OF EXPERIMENT 2

Classifier	Accuracy	Best hyperparameter values
C5.0	0.883	winnow=TRUE; noGlobalPruning=TRUE
k-NN	0.913	K = 1
ANN	0.835	maxit=780; size=50
Random Forest	**0.934**	ntree=447; mtry=4
kSVM	0.929	kernel=laplacedot; sigma=10

TABLE V
PERFORMANCE OF EXPERIMENT 3

Classifier	Accuracy	multiclass.au1u	multi.brier
Random Forest, k-NN,C50	0.942	0.99	0.481
Random Forest, k-NN,kSVM	0.942	0.99	0.481
Random Forest, k-NN,ANN	**0.944**	**0.991**	**0.479**
Random Forest, C50, kSVM	0.937	0.988	0.483
Random Forest, C50, ANN	0.94	0.989	0.482
Random Forest, kSVM, ANN	0.94	0.989	0.483
k-NN, C50, kSVM	0.931	0.985	0.491
k-NN, C50, ANN	0.927	0.984	0.492
k-NN, kSVM, ANN	0.916	0.981	0.497
C50, kSVM, ANN	0.886	0.972	0.515

Table III shows the results of first experiment, where C5.0, k-NN, ANN, Random Forest and kSVM classifiers were evaluated using default settings. We can note that the Random Forest classifier was the best performer. It scored 0.934 for accuracy, 0.987 for multiclass.au1u and 0.113 for
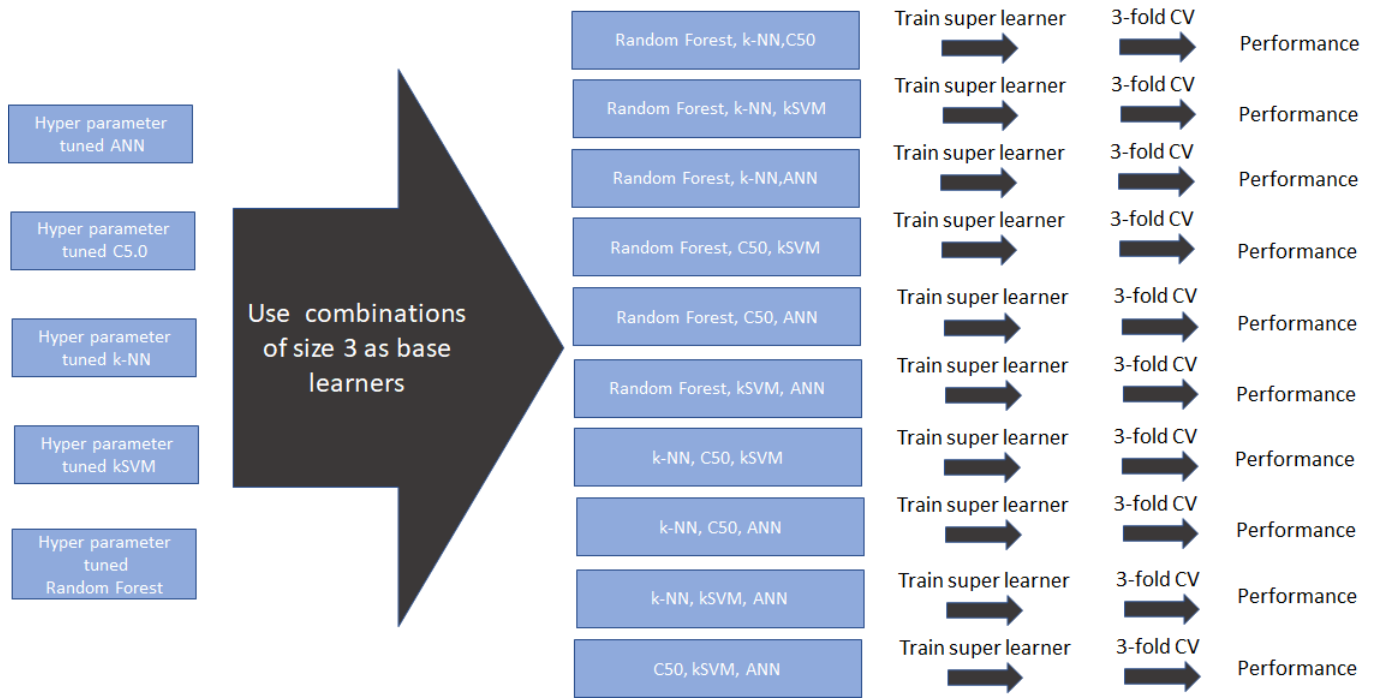
Random Forest, k-NN,C50	Train super learner	3-fold CV	Performance
Random Forest, k-NN, kSVM	Train super learner	3-fold CV	Performance
Random Forest, k-NN,ANN	Train super learner	3-fold CV	Performance
Random Forest, C50, kSVM	Train super learner	3-fold CV	Performance
Random Forest, C50, ANN	Train super learner	3-fold CV	Performance
Random Forest, kSVM, ANN	Train super learner	3-fold CV	Performance
k-NN, C50, kSVM	Train super learner	3-fold CV	Performance
k-NN, C50, ANN	Train super learner	3-fold CV	Performance
k-NN, kSVM, ANN	Train super learner	3-fold CV	Performance
C50, kSVM, ANN	Train super learner	3-fold CV	Performance

Fig. 4. Experiment 4 - Evaluation of prediction performance of the hyperparameter tuned stacked classifiers

Classifier	Accuracy	multiclass.au1u	multi.brier
Random Forest, k-NN,C50	0.939	0.989	0.0978
Random Forest, k-NN,kSVM	0.942	0.989	0.0923
Random Forest, k-NN,ANN	0.939	0.989	0.0995
Random Forest, C50, kSVM	0.941	0.99	0.0934
Random Forest, C50, ANN	0.938	0.989	0.1
Random Forest, kSVM, ANN	0.945	0.99	0.0912
k-NN, C50, kSVM	0.934	0.987	0.105
k-NN, C50, ANN	0.925	0.986	0.115
k-NN, kSVM, ANN	0.933	0.988	0.105
C50, kSVM, ANN	0.921	0.986	0.117

multiclass.brier. K-NN was the second best performer with 0.916 accuracy, 0.969 multiclass.au1u and 0.160 multiclass brier. The worst performance was by ANN with an accuracy of 0.586, multiclass.au1u of 0.837 and multiclass.brier of 0.529.

The results of experiment 2 is presented in Table IV. Hyperparameters of the 5 classifiers namely C5.0, k-NN, ANN, Random Forest and kSVM were tuned to obtain the highest accuracy. Accuracy of ANN classifier increased from 0.586 to 0.835 for the hyperparameters maxit=780 and size=50. Accuracy of kSVM also increased by hyperparameter tuning. It increased from 0.695 to 0.929 when kernel=laplacedot and sigma=10 were used. No improvements were obtained by tuning hyperparameters of Random Forest algorithm but it remained the best performer in experiment 2 as well. Accuracy of C5.0 and k-NN dropped slightly and obtained lower accuracy than for the default settings.

Table V illustrates the performance of the third experiment where classifier combinations of size 3 was used to create a stacked learner. These combinations were made using C5.0, k-NN, ANN, Random Forest and kSVM classifiers with default settings. L1-Regularized Logistic Regression classifier was used as the super learner of stacked classifiers. The best performance of experiment 3 was shown by Random Forest,k-NN and ANN classifier combination with 0.944 accuracy, 0.991 multiclass.au1u and 0.479 multiclass.brier. It is interesting to note that, it is a combination made up of the best two and the worst performers of experiment number one. Not only this combination, but also all the other combinations of base learners which contained Random Forest algorithm had scored a higher accuracy and multiclass.au1u than the best accuracy of experiment 1 which was obtained by Random Forest algorithm. The lowest performance of the combinations was by C5.0, kSVM and ANN combination, which were made up by the worst three performers of experiment one. Although, the said combination scored 0.886 accuracy and 0.972 multiclass.au1u which were higher than their individual performance in experiment 1. Although experiment 3 showed an overall improvement in accuracy and multiclass.au1u, performance of multiclass.brier turned out to become worse being less closer to zero than before.

As shown in Table VI, the best performance of all four experiments was obtained by the hyper parameter tuned Random Forest, kSVM and ANN combination of stacked classifiers in experiment 4. This combination scored 0.945 accuracy, 0.99 multiclass.au1u and 0.0912 multiclass.brier. Although the improvement of accuracy and multiclass.au1u is very small compared to the best performance of experiment 3, performance of the multiclass.brier measure has improved

from 0.479 to 0.0912 which is remarkable. The worst performance of experiment 4 was obtained by C5.0, kSVM and ANN combination scoring 0.921 for accuracy, 0.986 for multiclass.au1u and 0.117 for multiclass.brier. This is also a notable improvement from experiment 3, where 0.886, 0.972 and 0.515 for accuracy, multiclass.au1u and multiclass.brier were recorded as the worst performance.

IV. CONCLUSION

In this paper we introduced the prediction of NB-UVB phototherapy treatment response of psoriasis patients using data mining algorithms. We identified the most significant feature set and used four experiments to evaluate the performance. We noted that the performance of the classifiers improved by hyperparameter tuning and by using stacking technique. The best performance of all four experiments was obtained by the hyper parameter tuned Random Forest, kSVM and ANN combination of stacked classifiers in experiment 4. This combination scored 0.945 accuracy, 0.99 multiclass.au1u and 0.0912 multiclass.brier. Worst performance of all four experiments was by ANN when evaluated using default settings in experiment 1.

The experimental results are satisfactory and inspires further research in this area.

REFERENCES

[1] Tsoi, Lam C et al. "Identification of Fifteen New Psoriasis Susceptibility Loci Highlights the Role of Innate Immunity." Nature genetics 44.12 (2012): 1341-1348. PMC. Web. 10 May 2017.

[2] http://www.irishtimes.com/life-and-style/damage-limitation-is-the-only-cure-for-psoriasis-1.1583239

[3] Gelfand JM, Troxel AB, Lewis JD, et al. The risk of mortality in patients with psoriasis: results from a population-based study. Arch Dermatol. 2007 Dec;143(12):1493-1499.

[4] Rapp, Stephen R., et al. "Psoriasis causes as much disability as other major medical diseases." Journal of the American Academy of Dermatology 41.3 (1999): 401-407.

[5] Racz E, Prens EP. "Phototherapy and photochemotherapy for psoriasis". *Dermatol Clin.2015;33(1):79-89.*

[6] Wan J, Abuabara K, Troxel AB, et al. "Dermatologist preferences for first-line therapy of moderate to severe psoriasis in healthy adult patients".*J Am Acad Dermatol.2012; 66(3): 376-86.*

[7] Honigsmann H. "Phototherapy for psoriasis". *Clin Exp Dermatol.2001;26(4):343-50.*

[8] Nakamura M, Farahnik B, Bhutani T. "Recent advances in phototherapy for psoriasis". *F1000Research. 2016;5:F1000 Faculty Rev-1684. doi:10.12688/f1000research.8846.1.*

[9] PaÅąiÄŐ, Aida, et al. "Phototherapy in pediatric patients." Pediatric dermatology 20.1 (2003): 71-77.

[10] Dawe, R. S., et al. "A randomized controlled trial of narrowband ultraviolet B vs. bath-psoralen plus ultraviolet A photochemotherapy for psoriasis." British Journal of Dermatology 148.6 (2003): 1194-1204.

[11] http://www.bad.org.uk/shared/get-file.ashx?id=175&itemtype=document

[12] Jain, Nikita, and Vishal Srivastava. "Data mining techniques: a survey paper." IJRET: International Journal of Research in Engineering and Technology 2.11 (2013): 2319-1163.

[13] Patel, Sheenal, and Hardik Patel. "Survey of data mining techniques used in healthcare domain". International Journal of Information 6.1/2 (2016).

[14] Jothi, Neesha, and Wahidah Husain. "Data mining in healthcare-a review." Procedia Computer Science 72 (2015): 306-313.

[15] Ahmad, Parvez, Saqib Qamar, and Syed Qasim Afser Rizvi. "Techniques of data mining in healthcare: a review." International Journal of Computer Applications 120.15 (2015).

[16] Rambhajani, Madhura, Wyomesh Deepanker, and Neelam Pathak. "A survey on implementation of machine learning techniques for dermatology diseases classification." International Journal of Advances in Engineering & Technology 8.2 (2015): 194.

[17] https://en.wikipedia.org/wiki/Information_gain_ratio

[18] MM Breunig, HP Kriegel, RT Ng, J Sander, "LOF: identifying density-based local outliers", ACM sigmod record, 2000

[19] L Al Shalabi, Z Shaaban, B Kasasbeh ,"Data mining: A preprocessing engine" *Journal of Computer Science*, 2006

[20] https://en.wikipedia.org/wiki/Statistical_classification

[21] https://en.wikipedia.org/wiki/Ensemble_learning#Stacking

[22] https://en.wikipedia.org/wiki/Cross-validation_(statistics)

[23] Kubat, Miroslav, and Stan Matwin. "Addressing the curse of imbalanced training sets: one-sided selection." ICML. Vol. 97. 1997.

[24] Ferri, CÃľsar, JosÃľ HernÃ ąndez-Orallo, and R. Modroiu. "An experimental comparison of performance measures for classification." Pattern Recognition Letters 30.1 (2009): 27-38.

[25] Brier, Glenn W. "Verification of forecasts expressed in terms of probability." Monthly weather review 78.1 (1950): 1-3.