The Erythemato-Squamous Dermatology Diseases Severity Determination using Self-Organizing Map

Haryanto¹, Bain Khusnul Khotimah², Miftahul Ulum¹, Diana Rahmawati¹, Koko Joni¹, Ahmad Ubaidillah¹, Riza Alfita¹, and Lilik Anifah¹

Abstract—A new approach based on the implementation of Self Organizing Map is presented for automated detection of erythemato-squamous diseases. The purpose of clustering techniques is in order to determinate the severity of erythemato-squamous dermatology diseases. The studied domain contained records of patients with known diagnosis. Self-Organizing Map algorithm's task was to classify the data points, in this case the patients with attribute data, to one of the six clusters (psoriasis, seboreic dermatitis, lichen planus, pityriasis rosea, cronic dermatitis, dan pityriasis rubra pilaris). The algorithm was used to detect the six erythemato-squamous diseases when 33 features defining five disease indications were used. The purpose is to determine an optimum classification scheme for this problem. The present research demonstrated that the features well represent the erythemato-squamous diseases and SOM algorithm's task achieved high classification accuracies. The best accuration for psoriasis 85,94%, seboreic dermatitis 40,48%, lichen planus 56,25%, and pityriasis rosea 82,61%, with learning rate value were 0,1, 0,2, 0,9, and 0,4.

Keywords— Erythemato-squamous, Self-Organizing Map, classification, dermatology.

I. INTRODUCTION

The differential diagnosis of erythemato-squamous diseases is a real problem in dermatology. They all share the clinical features of erythema and scaling, with very little differences. The diseases in this group are psoriasis, seboreic dermatitis, lichen planus, pityriasis rosea, cronic dermatitis, and pityriasis rubra pilaris. Usually a biopsy is necessary for the diagnosis but unfortunately these diseases share many histopathological features as well [1].

Research about erythemato-squamous diseases has been researched many researcher in the world. Automatic detection of erythemato-squamous diseases using k-means clustering has been investigated by Ubeyli [2]. Classification of erythemato-squamous diseases using Support Vector Machine investigated by Derya [3] and another investigated by Loris [4].

Self-Organizing Map (SOM) invented by Kohonen [5], [6]. SOM network learning some grop pattern and a pattern made into a single cluster that has certain characteristics [7]. SOM procedure is divided into three parts, namely: Normalization, Trainings, and retrieval of information from the training. The goal of normalization is to avoid variable that has value too high against the value of other data. Normalization process variable value seeing a value between 0 and 1, is expected to be detraining variable is normal with other data. After data normalization, data matrix detraining shaped by a particular procedure, recommended training held at least 500 times. The third step, namely the use trainings results in the form of weights (weight) for each cluster. Weight (weight) obtained is used to test the data that will be tested.

In this paper we present the application of a selforganizing map to the problem of differential diagnosis of erythemato-squamous diseases, which represents a major issue in dermatology.

This paper is outlined as follows: In section 2 we briefly describe material and the proposed method. Experimental results and discussions will be presented in section 3, followed by discussions in section 4, conclusions in section 5, and future work in section 6.

II. MATERIAL AND METHOD

A. Material

We observed 231 erythemato-squamous diseases data obtained from Machine Learning Repository (UCI) dataset [1], included 30 training data and 201 testing data. Training data contain 5 data for psoriasis, 5 data for seboreic dermatitis, 5 data for lichen planus, 5 data for pityriasis rosea, 5 data for cronic dermatitis, and 5 data for pityriasis rubra pilaris.

The dataset constructed for this domain, the family history feature has the value 1 if any of these diseases has been observed in the family, and 0 otherwise. The age feature simply represents the age of the patient. Every other feature (clinical and histopathological) was given a degree in the range of 0 to 3.

Hardware specification used in this research is Dell Inspiron N3010 with Processor Intel(R) Core(TM) i5 CPU M430 @2,27 GHz installed memory (RAM) 2,00 GB.

B. Method

Method of Self Organizing Map (SOM) has been progressing rapidly both in the substance of the SOM and its application to clustering in various fields. Modular network SOM developed by Kazuhiro Tokunaga and Tetsuo Furukawa [9]. Networks Artificial Counterpropagation developed by Davide Ballabio, Viviana Consonni, and Roberto Todeschini [10]. SOM of SOMs developed by Tetsuo Furukawa [11], Dieter Müller developed a self-organized mapping of clusters of

¹Haryanto, Miftahul Ulum, Diana Rahmawati, Koko Joni, Ahmad Ubaidillah, Riza Alfita, and Lilik Anifah are with Departement of Electrical Engineering, Trunojoyo University Madura, Bangkalan, Indonesia. E-mail: haryanto_utm@yahoo.com;;;;;;.

²Bain Khusnul Khotimah are with Departement of Informatics Engineering, Trunojoyo University Madura, Bangkalan, Indonesia.E-mail: bain_kk@gmail.com.

data into groups of neurons [12]. Comparision between SOM-based optimization and particle swarm optimization for minimized construction time from a secant pile wall is researched by Jieh-Haur Chen [13]. Still a lot of research about other SOM. Figure of classic SOM configuration is figured in fig 1.

Erythemato-squamous diseases classification in this research using classic SOM. This research is divided into 9 experiments described in Table 1.

The clasters of this research are divided into 6 clasters described in Table 2. Clasters of this research are Psoriasis, Seboreic dermatitis, Lichen planus, Pityriasis rosea, Cronic dermatitis and Pityriasis rubra pilaris.

Algorithm of this research is figured in fig 2. The step of learning using SOM are initialization, determination the number of classes, setting a learning rate parameter, determination of the iteration value, and calculating the distance between random data into each input weight.

The classic SOM network learning algorithm can be formulated as follows [6]:

- a. Set learning iteration number t = 0. Then initialize all weights wij with small random values (or initialize the weights using input data). Set the initial topological neighborhood (d0). Set the initial learning rate (0) and set the total number of iterations (T)
- b. While iteration number (t) is less than T, repeat Step 3–Step 6.
- c. Choose an input vector x randomly in the training set.
- d. Determine the neuron j so that its weight vector wj is the closest to the input vector and call it as the winner neuron. The winner neuron j has the closest distance, Dmin(t) to the input pattern x(t), where Dmin(t) is given:

e.
$$D_{\min}(t) = \min D_i(t) = \min \sum_j (x_{ij}(t) - w_{ij}(t))^2$$

f. Update the weight vectors of both the winner neuron j and its neighbors as:

g.
$$W_i(t+1) = W_i(t) + \alpha(t) \bullet (x - W_i(t))$$
$$\forall i \in N_i$$

h. where $\alpha(t)$ is a learning rate function that exponentially decreases with time. Further, we define a neighborhood order function d(t) which exponentially decreases with time. In this paper, the following equations are used for $\alpha(t)$ and d(t):

i.
$$\begin{aligned} \alpha(t) &= \alpha_0 e^{-t/3T} \\ d(t) &= \left[d_0 e^{-t/3T} \right] \end{aligned}$$

- j. where [y] denotes the largest integer less or equal to y.
- k. Set t = t+1 and if t < T go to Step 3, and otherwise STOP.

III. RESULT AND DISCUSSION

Difficulty for the differential diagnosis is that a disease may show the features of another disease at the beginning stage and may have the characteristic features at the following stages.

Patients were first evaluated clinically with 12 features. Afterwards, skin samples were taken for the evaluation of 22 histopathological features. The values of the histopathological features are determined by an analysis of the samples under a microscope. The Clinical Attributes areerythema, scaling, definite borders, itching, koebner phenomenon, polygonal papules, follicular papules, oral mucosal involvement, knee and elbow involvement, scalp involvement, family history, and Age. Histopathological Attributes melanin incontinence, eosinophils in the infiltrate, PNL infiltrate, fibrosis of the papillary dermis, exocytosis, acanthosis, hyperkeratosis, parakeratosis, clubbing of the rete ridges, elongation of the rete ridges, thinning of the suprapapillary epidermis, spongiform pustule, munro microabcess, focal hypergranulosis, disappearance of the granular layer, vacuolisation and damage of basal layer, spongiosis, saw-tooth appearance of retes, follicular horn plug, perifollicular parakeratosis, inflammatory monoluclear inflitrate, and band-like infiltrate.

Parameters of learning are described in Table 1 for minimal value, Table 2 for maximal value and Table 3 for average value. Table 3, 4, and 5 shows that the some parameters had the overlap value among clasters and it is not linear. That is why we try to overcome to classify automatically using SOM.

Accuracy of each group is shown in fig 4, fig 5, fig 6, fig 7, fig 8, and fig 9. Fig 4 is figured the accuray of the fisrt group (Psoriasis), fig 5 is figured the accuray of the second group (Seboreic dermatitis), fig 6 is figured the accuray of the second group (Lichen planus dermatitis), fig 7 is figured the accuray of the second group (Pityriasis rosea), fig 8 is figured the accuray of the second group (Cronic dermatitis) and fig 9 is figured the accuray of the second group (Pityriasis rubra pilaris).

Fig 4 shows that the accuracy of cluster 1 (Psoriasis group), first experiment and second experiment are 85,9375%. Thirth, fourth, and fifth experiment have accuracy 84,375% but sixth, seventh, eighth and ninth experiment give accuracy 0%.

Fig 5 shows that the accuracy of second cluster (Seboreic dermatitis group), first experiment and second experiment are 40,476%, thirth has accuracy 0%, fourth and sixth 21,428%, fifth 40,476, seventh and ninth have accuracy 35,714%, and eighth experiment has accuracy 38,095%.

Fig 6 shows that the accuracy of thirth cluster (Lichen planus group), first and second experiment are 31,25%, thirth has accuracy 18,75%, fourth, fifth and sixth 12,5%, seventh 0%, eighth has accuracy 18,75% and ninth experiment has accuracy 56,25%.

Fig 7 shows that the accuracy of fourth cluster (Pityriasis rosea group), first experiment and second experiment are 78,26%. Thirth is 69,565%, fourth has accuracy 82,61%,fifth experiment has accuracy 4,34 %, sixth has accuracy 60,869%, eighth has 17,39% but seventh and ninth 0%.

Fig 8 shows that the accuracy of fifth cluster (Cronic dermatitis group), first experiment and second experiment are 86,49%. Thirth is 43,24%, fourth has accuracy 35,13%, fifth experiment has accuracy 13,51%, sixth has accuracy 37,84%, seventh and eighth have 48,65% but ninth has accuracy 10,81%.

Fig 9 shows that the accuracy of sixth cluster (Pityriasis rubra pilaris group), first, second, thirth, fourth, and seventh experiment 93,33%, fifth has accuracy 7,14%, sixth experiment has accuracy 6,67%, eighth has accuracy 100%, and ninth has accuracy 0%.

Time needs to process of each experiment is described in Table 7. First, seventh, eighth, and ninth experiment need 0,2028 second to process. Second and fourth experiment need 0,187 second, thirth experiment needs 0,218 second, fifth and sixth experiments consume 0,265 and 0,234 second. Graph of time-consuming for each experiment is figured in fig 10.

IV. CONCLUSION

In this research, we proposed an effective Erythematosquamous diseases classification system without the need for user interaction. Experimental results have been provided to show the successful of the system that could help medical doctors as decision support system to determine the claster of erythemato-squamous diseases.

A. Future Work

The system could classify the six claster (Psoriasis, Seboreic dermatitis, Lichen planus, Pityriasis rosea, Cronic dermatitis and Pityriasis rubra pilaris), the future work we develop the more effective and robust classification machine.

ACKNOWLEDGEMENT

This research was supported by UC Irvine Machine Learning Repository (UCI).

REFERENCES

- [1]. http://archive.ics.uci.edu/ml/
- [2]. Ubeyli, Elif Derya and Erdoğan Doğdu 2008, Automatic detection of erythemato-squamous diseases using k-means clustering has been investigated Journal of Mecidal System

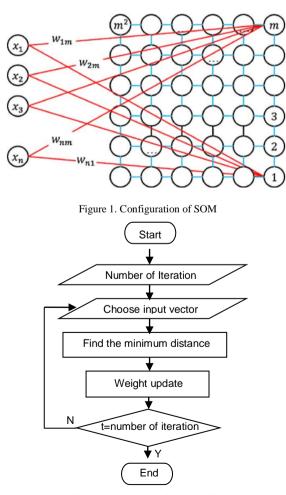


Figure 2. Algorithm of classic SOM

Volume 34, Number 2, 179-184, DOI: 10.1007/s10916-008-9229-6

- [3]. Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biological Cybernetics 43, 59e69.
- [4]. Kohonen, T., 1982. Analysis of a simple self-organizing process. Biological Cybernetics 44, 135e140.
- [5]. Kohonen, T., 2001. Self-Organizing Maps. Springer-Verlag, Berlin.
- [6]. A.M. Kalteh, 2008, Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application, Environmental Modelling & Software 23 (2008) 835e845, Science Direct Elsevier.
- [7]. Haykin, S., 1999, Neural Networks: A Comprehensive Foundation, Prentice Hall Upper Saddle River, New Jersey.
- [8]. Vesanto, J., Alhoniemi, E., 2000, Clustering of the selforganizing map. IEEE, Transactions on Neural Networks 11 (3), 586e600. Woolf AD, Pfleger B. "Burden of major musculoskeletal conditions", Bull World Health Organ 2003;81:646-656.
- [9]. Kazuhiro Tokunaga, Tetsuo Furukawa, Modular network SOM, Neural Networks 22 (2009) 82-90, Science Direct Elsevier, 2009.
- [10]. Davide Ballabio, Viviana Consonni, Roberto Todeschini, The Kohonen and CP-ANN toolbox: A collection of MATLAB modules for Self Organizing Maps and Counterpropagation Artificial Neural Networks, Chemometrics and Intelligent Laboratory Systems xxx (2009) xxx–xxx, Science Direct Elsevier, 2009.
- [11]. Tetsuo Furukawa, SOM of SOMs, Neural Networks 22 (2009) 463§478, Science Direct Elsevier, 2009.
- [12]. Dieter Müller, Self organized mapping of data clusters to neuron groups, Neural Networks 22 (2009) 415-424, Science Direct Elsevier, 2009.
- [13]. Jieh-Haur Chen, Comparison of SOM-based optimization and particle swarm optimization for minimizing the construction time of a secant pile wall, Automation in Construction 18 (2009) 844–848, Science Direct Elsevier, 2009.

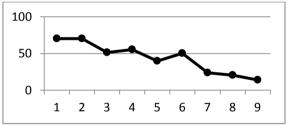


Figure 3. System accuracy using learning rate value 0,1 to 0,9

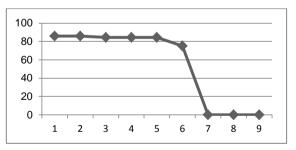
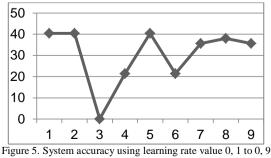


Figure 4. System accuracy using learning rate value 0,1 to 0,9 of cluster 1 (Psoriasis group)



of cluster 2 (Seboreic dermatitis group)

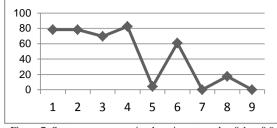
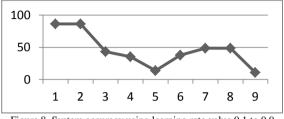
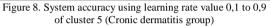
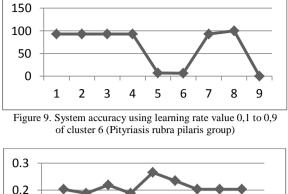


Figure 7. System accuracy using learning rate value 0,1 to 0,9 of cluster 4 (Pityriasis rosea group)







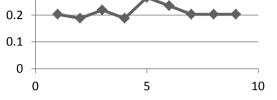


Figure 10. Time-consuming for each experiment

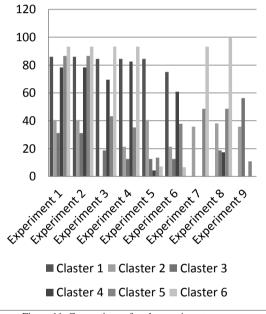


Figure 11. Comparison of each experiment accuracy

		BLE 1. OF THE RESEAR	СН		TABLE 2. Claster of the research				
No	Experiment	Learning rate	Iteration	No	Claster	Group			
1	Experiment 1	0,1	500	1	Claster 1	Psoriasis			
2	Experiment 2	0,2	500	2	Claster 2	Seboreic Dermatitis			
3	Experiment 3	0,3	500	3	Claster 3	Lichen Planus			
4	Experiment 4	0,4	500	4	Claster 4	Pitriasis Rosea			
5	Experiment 5	0,5	500	4	Claster 4				
6	Experiment 6	0,6	500	5	Claster 5	Cronic Dermatitits			
7	Experiment 7	0,7	500	6	Claster 6	Pityriasis Rubra Pilaris			
8	Experiment 8	0,8	500						
9	Experiment 9	0,9	500						

IPTEK, Journal of Proceeding Series, Vol. 1, 2014 (eISSN: 2354-6026)

MINIMAL VALUE OF LEARNING PARAMETERS Parameter 1 2 3 4 5 6						MAXIMAL VALUE OF LEARNING PARAMETERS Parameter 1 2 3 4 5 6							
1: Erythema	0	1	0	1	0	1	1: Erythema	3	3	3	3	3	
2: Scaling	1	1	0	1	0	1	2: Scaling	3	3	3	2	3	
2: Scaling 3: Definite borders		0	0				2: Scaling 3: Definite borders		2	3			
	1			0	0	0		3			2	3	
4: Itching	0	0	0	0	0	0	4: Itching	3	3	3	3	3	
5: Koebner phenomenon	0	0	0	0	0	0	5: Koebner phenomenon	3	2	3	3	0	
6: Polygonal papules	0	0	0	0	0	0	6: Polygonal papules	0	0	3	0	0	
7: Follicular papules	0	0	0	0	0	1	7: Follicular papules	2	1	0	0	2	
8: Oral mucosal involvement	0	0	0	0	0	0	8: Oral mucosal involvement	0	2	3	0	0	
9: Knee and elbow involvement	0	0	0	0	0	0	9: Knee and elbow involvement	3	1	2	0	1	
10: Scalp involvement	0	0	0	0	0	0	10: Scalp involvement	3	0	1	0	0	
11: Family history, (0 or 1)	0	0	0	0	0	0	11: Family history, (0 or 1)	1	1	0	0	0	
12: Melanin incontinence	0	0	0	0	0	0	12: Melanin incontinence	0	0	3	0	0	
13: Eosinophils in the infiltrate	0	0	0	0	0	0	13: Eosinophils in the infiltrate	2	2	2	1	1	
14: PNL infiltrate	0	0	0	0	0	0	14: PNL infiltrate	3	3	0	1	0	
15: Fibrosis of the papillary dermis	0	0	0	0	1	0	15: Fibrosis of the papillary dermis	0	0	2	0	3	
16: Exocytosis	0	0	0	0	0	0	16: Exocytosis	2	3	3	3	2	
17: Acanthosis	0	0	0	0	1	1	17: Acanthosis	3	3	3	2	3	
18: Hyperkeratosis	0	0	0	0	0	0	18: Hyperkeratosis	3	3	2	2	2	
19: Parakeratosis	0	0	0	0	0	0	19: Parakeratosis	3	3	3	2	2	
20: Clubbing of the ridges	0	0	0	0	0	0	20: Clubbing of the ridges	3	0	0	0	0	
21: Elongation of the rete ridges	1	0	0	0	0	0	21: Elongation of the rete ridges	3	2	0	0	3	
22: Thinning of the suprapaillary							22: Thinning of the suprapaillary						
Epidermis	0	0	0	0	0	0	Epidermis	3	1	0	0	1	
23: Spongiformpustule	0	0	0	0	0	0	23: Spongiformpustule	3	2	0	0	0	
24: Munro microabcess	0	0	0	0	0	0	24: Munro microabcess	3	0	3	1	0	
25: Focal hypergranulosis		0	0	0	0	0	25: Focal hypergranulosis	0	0	3	0	0	
26: Disappearance of the granular							26: Disappearance of the granular						
Layer	0	0	0	0	0	0	Layer	3	0	2	2	0	
27: Vacuolisation and damage of	Ū	Ū	Ū	Ū	Ū	0	27: Vacuolisation and damage of						
Basal layer	0	0	0	0	0	0	Basal layer	1	0	3	0	0	
28: Spongiosis	0	0	0	0	0	0	28: Spongiosis	0	3	3	3	2	
	0	0	0	0	0	0	29: Saw-tooth appearance of retes	0	0	3	1	0	
29: Saw-tooth appearance of retes							30: Follicular horn plug	0	1	1	0	1	
30: Follicular horn plug	0	0	0	0	0	0	31: Perifollicular parakeratos is	0	1	0	0	0	
31: Perifollicular parakeratos is	0	0	0	1	0	1	32: Inflammatory monoluclear	Ŭ	•	2	2	2	
32: Inflammatory monoluclear	~	~	~	~	c	c	Infiltrate	3	3	3	3	3	
Infiltrate	0	0	0	0	0	0	33: Band-like infiltrate	1	2	3	0	1	
33: Band-like infiltrate	0	0	0	0	0	0	55. Dand-fike minuate	1	2	5	0	1	

Parameter	1	2	3	4	5	6
1: Erythema	3	3	3	3	3	3
2: Scaling	3	3	3	2	3	2
3: Definite borders	3	2	3	3	3	2
4: Itching	3	3	3	3	3	2
5: Koebner phenomenon	3	2	3	3	1	0
6: Polygonal papules	0	0	3	3	0	0
7: Follicular papules	2	1	0	0	2	3
8: Oral mucosal involvement	0	0	3	2	0	0
9: Knee and elbow involvement	3	3	2	0	1	3
10: Scalp involvement	3	3	1	0	0	2
11: Family history, (0 or 1)	1	1	0	0	0	1
12: Melanin incontinence	0	0	3	2	0	0
13: Eosinophils in the infiltrate	2	2	2	1	1	0
14: PNL infiltrate	3	3	1	1	0	1
15: Fibrosis of the papillary dermis	0	0	2	0	3	3
16: Exocytosis	2	3	3	3	2	3
17: Acanthosis	3	3	3	2	3	3
18: Hyperkeratosis	3	3	2	2	2	2
19: Parakeratosis	3	3	3	2	2	2
20: Clubbing of the ridges	3	3	0	0	0	1
21: Elongation of the rete ridges	3	2	0	0	3	2
22: Thinning of the suprapaillary						
Epidermis	3	3	0	0	1	0
23: Spongiformpustule	3	2	0	0	0	1
24: Munro microabcess	3	2	3	1	0	C
25: Focal hypergranulosis	0	0	3	3	0	1
26: Disappearance of the granular						
Layer	3	2	2	2	0	C
27: Vacuolisation and damage of						
Basal layer	1	0	3	2	0	C
28: Spongiosis	0	3	3	3	2	3
29: Saw-tooth appearance of retes	0	0	3	1	0	C
30: Follicular horn plug	0	1	1	0	1	3
31: Perifollicular parakeratos is	0	1	0	0	0	3
32: Inflammatory monoluclear						
Infiltrate	3	3	3	3	3	3
33: Band-like infiltrate	2	2	3	3	1	1
34: Age (linear)	7 5	7 0	6 5	7 0	7 0	5 6

No	Experiment	Accuracy (%)
1	Experiment 1 (with learning rate 0,1)	70,149
2	Experiment 2 (with learning rate 0,1)	70,149
3	Experiment 3 (with learning rate 0,1)	51,243
4	Experiment 4 (with learning rate 0,1)	55,223
5	Experiment 5 (with learning rate 0,1)	39,800
6	Experiment 6 (with learning rate 0,1)	50,248
7	Experiment 7 (with learning rate 0,1)	23,880
8	Experiment 8 (with learning rate 0,1)	20,398
9	Experiment 9 (with learning rate 0,1)	13,930

	TABLE 7. TIME-CONSUMING FOR EACH EXPERIMENT					
Learning Rate	Time (Second)					
Experiment 1 (learning Rate 0,1)	0,2028013					
Experiment 2 (learning Rate 0,2)	0,1872012					
Experiment 3 (learning Rate 0,3)	0,2184014					
Experiment 4 (learning Rate 0,4)	0,1872012					
Experiment 5 (learning Rate 0,5)	0,2652017					
Experiment 6 (learning Rate 0,6)	0,2348015					
Experiment 7 (learning Rate 0,7)	0,2028013					
Experiment 8 (learning Rate 0,8)	0,2028013					
Experiment 9 (learning Rate 0,9)	0,2028013					