International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:6, No:10, 2012

Analysis of Palm Perspiration Effect with SVM for Diabetes in People

Hamdi Melih Saraoğlu, Muhlis Yıldırım, Abdurrahman Özbeyaz, and Feyzullah Temurtaş

Abstract**—**In this research, the diabetes conditions of people (healthy, prediabete and diabete) were tried to be identified with non-invasive palm perspiration measurements. Data clusters gathered from 200 subjects were used (1.Individual Attributes Cluster and 2. Palm Perspiration Attributes Cluster). To decrase the dimensions of these data clusters, Principal Component Analysis Method was used. Data clusters, prepared in that way, were classified with Support Vector Machines. Classifications with highest success were 82% for Glucose parameters and 84% for HbA1c parametres.

Keywords**—**Palm perspiration, Diabetes, Support Vector Machine, Classification.

I. Introduction

DIABETES MELLITUS (DM) is the medical name of the disease which is generally known as diabetes. DM, characterized by hyperglycemia (high blood/plasma glucose level), is a metabolism disorder that emerges due to insulin deficiency or resistance towards insulin. It is sufficient to diagnose DM when glucose level is once ascertained above 200 mg/dl in random or oral glucose test [1].

It is estimated that diabetics around the world will be 366 million in 2030 [2]. In addition, it is determined that 6% of people between 20 – 79 age range have diabetes according to a research which was conducted by International Diabetes Federation (IDF) in IDF member countries in 2007. Every year, 3.8 million people die due to diabetes related reasons [3, 4].

Diabetes is a major health problem in both industrial and developing countries, and its incidence is rising [5,6]. The people with diabetes must control their glucose levels. Because glucose is very important for them [7]. The long-term excess of glucose (hyperglycemia) can cause many problems in diabetes such as blindness, damaged nerves and kidney failure (renal failure), or even increase the heart diseases, strokes and birth defects. On the contrary, the long term low glucose rate (hypoglycemia) can cause confusion, coma and even death. Monitoring the glycemic state of patients is the cornerstone of diabetic care [5-7].

Glucose sticks to the hemoglobin to make hemoglobin A1C (HbA1C). The most widely used tests for monitoring the glycemic status of patients with diabetes are blood glucose and HbA1C. Hemoglobin A1C value reflects mean glucose level during the previous 2-3 months' period and is a useful indicator for risk assessment of diabetic complications [7].

Fasting blood glucose level of a healthy person should be <100 mg/dl (5,6 mmol/l). If this level is between 100-125 mg/dl and (5,6-6,9 mmol/l), then there is a risk of diabetes (pre-diabetes). When the fasting blood glucose level is greater than 126 mg/dl (7.0 mmol/I), pre-diagnosis for diabetes is established [8]. The limit value of HbA1c for the diagnosis of diabetes is determined as 6,5% by the American Diabetes Association and by International Diabetes Federation [9,10].

Invasive techniques in glucose and HbA1c measurement are widely used in diagnosis and treatment process. However, they cause damage and pain to human body in the application process. Therefore, non-invasive measurement techniques which are carried using other body parameters instead of blood have been searched [11]. In the development of non-invasive reception models, sweat, saliva and breath are the most widely used parametres [12,13].

Compared to other parameters, it is easier to obtain perspiration samples [12,13,14,15,16]. In researches which have been conducted recently, the relation between perspiration and illnesses is clearly seen [14,17]. Perspiration rate, perspiration speed and chemicals in perspiration help determining illnesses and body changes. Within this frame, there have been conducted many researches regarding perspiration [11,18,19,20,21,22].

Support Vector Machines (SVM) were suggested by Vapnik in 1998. Built on machine learning methods, SVM adopts a different structure when compared to other instructor supported learning techniques [23]. The structure of SVM aims to find the optimum separating hyper plane between members which belong to a class in an abstract space which was known in advance and members which do not [23,24]. SVM is frequently used in a wide range of fields [25]. In researches, SVM has been frequently used due to its establishing a model by learning statistically gathered information, its preciseness in prospective estimation with present data and its ability to estimate.

In this research, it is aimed to determine the diabetes condition of people using humidity variance measurements taken from the palms during 80 seconds. Data clusters

Saraoğlu H.M. is with the Electrical and Electronics Engineering Department, Dumlupınar University, Kütahya, Turkey (e-mail: saraoglu@dpu.edu.tr).

Yıldırım M. is with the Bio-medical Device Technology Program, Fatih University, Ankara, Turkey (e-mail: myildirim@fatih.edu.tr).

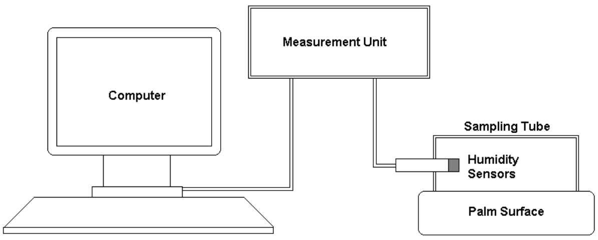
Özbeyaz A. is with the [3]Computer Programming Program, Adıyaman University, Adıyaman, Turkey (e-mail: aozbeyaz@adiyaman.edu.tr).

Temurtaş F. is with the Electrical and Electronics Engineering Department, Bozok University, Yozgat, Turkey (e-mail: feyzullah.temurtas@bozok.edu.tr).

gathered from 200 subjects were used (1.Individual Attributes Cluster and 2. Palm Perspiration Attributes Cluster). Principal Component Analysis Method was used to decrase the dimensions of these data clusters. Data clusters, prepared in that way, were classified with Support Vector Machines.

II. Material

General structure of the system which is designed with the support of TUBITAK for this research and that can measure palm perspiration changes is shown in Figure 1(a) and the measurement system is shown in Figure 1(b).



(a)

(b)

Fig. 1 a) General structure of measurement system b) Measurement system

There are cylindrical data detection reservoirs on measurement system, on which right and left palms can be positioned. There are two SHT and one ChipCap humidity-temperature sensor in these reservoirs. Therefore, it is possible to obtain perspiration signs of right and left hands. Subjects position their palms on measurement reservoirs and perspiration sign data are recorded by measurement system for eighty seconds. During the recording, it is important to be careful to avoid air entrance to measurement reservoir.

Measurement system is both able to show humidity and temperature changes with its graphic touch screen and assures recording to PC via RS-232 serial port. Blood tension and pulse measurements were conducted with OMRON.

A. Gathering and Preparing Palm Data

In this study, data which were obtained from 200 subjects who applied to Kutahya State Hospital for the determination of blood glucose and HbA1c values were used. 135 of these 200 are female and 65 are male. Fasting plasma glucose (FPG) values of these subjects range from 65mg/dl to 590 mg/dl and their HbA1c values are within the range of 4% and 13.77%.

Palm perspiration data are recorded every 2 seconds in 80 seconds and 40 points are obtained and the curve is shown in Figure 4. Each data point is named as $RH_1, RH_2, ..., RH_{40}$. As shown in Equity 1, ΔRH_i values are obtained by taking differentials between both data point. The set which is constituted by ΔRH_i difference values is as shown in equity 2 and signified by the letter D.

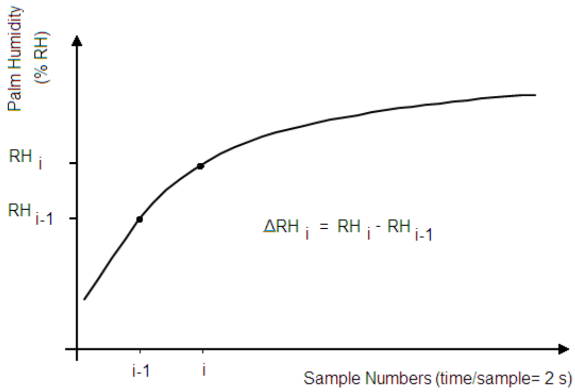


Fig. 2 Short-term palm humidity change diagram

$$\begin{aligned} \Delta RH_1 &= RH_2 - RH_1 \\ \Delta RH_2 &= RH_3 - RH_2 \\ &\vdots \\ \Delta RH_{39} &= RH_{40} - RH_{39} \end{aligned} \tag{1}$$

$$D = \{\Delta RH_1, \Delta RH_2, ..., \Delta RH_{39}\} \tag{2}$$

There are two basic clusters in the classification done by the Support Vector Machine. The first cluster is the individual attributes and the second one is palm perspiration attributes

Age:	Year
Systole:	mmHg
Diastole :	mmHg
Body Mass Index (BMI):	Weight/Height^2(kg/m^2)

The elements of individual attributes cluster is as follows:

The elements of palm perspiration attributes cluster is as follows:

Maximum difference:	Highest ΔRH_i value
Minimum difference :	Lowest ΔRH_i value
Average :	Average ΔRH_i values
Standard deviation:	ΔRH_i values standard deviation
Humidity change (ΔRH):	$\Delta RH = RH_{40} - RH_1$

B. Pretreatment of Data Clusters with PCA

SVM is a statistical learning method. Data which are used in the study keep past statistics, form a model and search belonging state of available data. Conducting a statistical process on data which will be classified will guarantee a successful classification. For that purpose, many statistical methods such as Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Linear Discrimination Analysis (LDA) were developed [27].

Principal Component Analysis (PCA) is a multivariate statistical method which describes the variance covariance structure of a variation set via linear combination of these variances and provides data reduction and interpretation. PCA, which is known as feature extraction, will decrease the available data, enable to avoid unnecessary data and will positively effect classification success. It is difficult to make various evaluations as data used in the study consist multivariate and connected information [27]. PCA was used before classification in the conducted study and many interconnected variables were converted into independent, more significant and less variables. Therefore, it is considered that classification success would be increased.

C.Diabetes Diagnosis Values

The reference value of fasting blood glucose (FBG) for classification is given in Table 1 [8,9]. The input parametres of 200 subjects were measured. According to the lab test results of fasting blood glucose, 50 of the subjects were healthy, 64 of them were under the risk of diabetes and 86 of them were diabetic. Another parametre used in diagnosis of diabetes and in monitoring the conditions of diabetics is the HbA1c which is measured in blood sample. The level of HbA1c in healthy people should be below 5,7%. If a diabetic patient's HbA1c value is below 6,5% (the measurement should be carried out in every three months), the glucose level is thought to be under control for that patient and the patient is diognosed as healthy [9,10]. That is to say, the patient has managed to control his glucose level. The reference range values used in the classification of HbA1c are presented in Table 2. When the 200 subjects are classified in accordance with the lab HbA1c results, it is seen that 76 of them are 'good', 54 of them are 'tolerable', and 70 of them are 'bad'.

TABLE I
CLASSIFICATION RANGE OF BLOOD GLUCOSE LEVEL FOR DIAGNOSIS

Diagnosis	Healthy	Diabetes Risk	Diabetes
Fasting Plasma Glucose	<100 mg/dl (50 subjects)	100-125 mg/dl (64 subjects)	>126 mg/dl (86 subjects)

TABLE II
HBA1C VALUE CLASSIFICATION RANGE FOR DIAGNOSIS

Diagnosis	Good	Tolerable	Bad
HbA1c	<%6,5 (76 subjects)	%6,6-%7,4 (54 subjects)	>%7,5 (70 subjects)

D. Support Vector Machine

Support Vector Machine (SVM) which is among the classification techniques based on optimization is a method used in data mining classification problems. In this method classification is done by linear and non-linear functions. The goal of SVM is to find the best hyperplane in feature space. As it is seen in Figure 3, there are lots of hyperplane possibilities in data groups which can be seperated from each other. To achieve a minimum level of fault tolerance, a hyperplane remote to each of the groups should be found. These lines are named as Support Vectors. When bilateral and multilateral classification methods are considered, it is seen that the classes in data clusters have appropriate results as:

$D = (x_1, y_1),...,(x_i, y_i), x \in R^n, y \in \{1,-1\}$ [23,24].

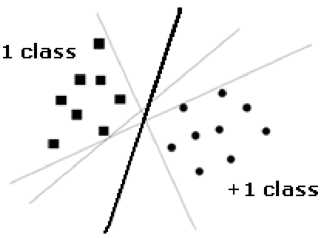


Fig. 3 Optimum Separator Planes

In situations when the data can not be seperated with linear methods, non-linear classifications rather than the linear ones can be used. Non-linear classifications actualize classifications by carrying the data in a multidimensional space. [24]. In Figure 4, a sample data cluster which can not be classified with linear method is presented. In Figure 4 (b), one dimensional data that can be seperated by Support Vectors by being carried into two dimensional space is presented.

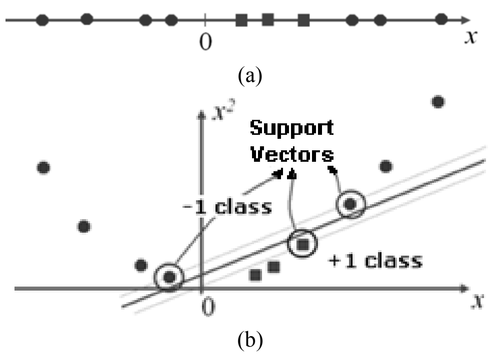


Fig. 4 a) Data cluster that can not separated with linear method in one dimensional space b) Linear separation by being carried into two-dimensional space [28].

The solution of non-linear problems is actualized by carryingn the samples with Kernel functions (that is Ø) to a space which is multidimensional and where they can be seperatedwith linear methods [24].

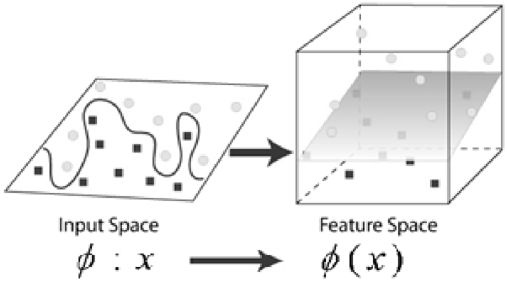


Fig. 5 The classification of a data cluster in multi-dimensional space [29].

In Figure 5, x inputs with Ø functions are carried to a multidimensional space as Ø(x). Thus, optimum hyperplane is obtained with the solution of optimization problems in sample space. This classification (obtaining optimum hyperplane) is done by *f*(x) optimization function. The optimization Formula of the Support Vector Machines is generally stated as;

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x_j) + b$$

As there is not a hyperplane in this situation b is omitted. The b term in question takes place in kernel functions. A more detailed information about SVM exists in literature and it is summarised briefly in this section [23,24].

Various kernel functions are used in Support Vector classifications. Most widely used ones are as follows;

Linear Kernel:

$$K(x_i, x_j) = x_i^T x_j$$

Polynomial Kernel:

$$K(x_i, x_j, c, d) = (c + x_i^T x_j)^d$$

Sigmoid Kernel:

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + \Phi)$$

Radial Based Kernel:

$$K(x_i, x_j, \sigma) = e^{\frac{\|x_i - x_j\|^2}{2\sigma^2}}$$

In machine learning applications, data is mainly grouped into two as education and testing. First, machine learning operation is actualized with education data and then the success of learning (classification) is tested with test data. In this study, it is aimed to design a machine that can make statistical analysis by using the data separated for education and testing. As the data can not be classified with linear method, it is carried into high level sample space wih Radial based kernel function and tried to be classified [23,26,27].

III. Method

In the study, first the blood glucose levels of the patients were listed using the individual attributes and palm perspiration attributes. The organized data was disseminated with Principal Component Analysis and meaningful signs that have relations with each other were tried to be gathered. Data, used in the study, has three classes. The first one comprises the patient group with a low fasting blood glucose level, the second one comprises the patient group with a normal fasting blood glucose level and the last one comprises the patient group with high fasting blood glucose level. Same classes were re-designed in accordance with HbA1c reference values. During classification, both the fasting blood glucose levels and the HbA1c data clusters were grouped into two as education and testing. Machine was trained with the data used for training and the machine was tested with data used for testing to see how accurately it can diognose. Classifications were done separately in accordance with fasting blood glucose and HbA1c levels. During classification, data was chosen randomly for training and testing.

Studies are tested with a computer equipped with core 2 processor and 4 GB memory. All the algorithms are actualized with Support Vector Machine and Principal Component Analysis in Matlab application environment used widely in the fields of engineering and research.

IV. Results and Discussion

In Figure 6, the success rates of SVM classifications by using the data clusters in this study (a.Classification successes with using the palm perspiration attributes, b.Classification successes with using the individual attributes and c. Classification successes with uniting all the data clusters) is presented.

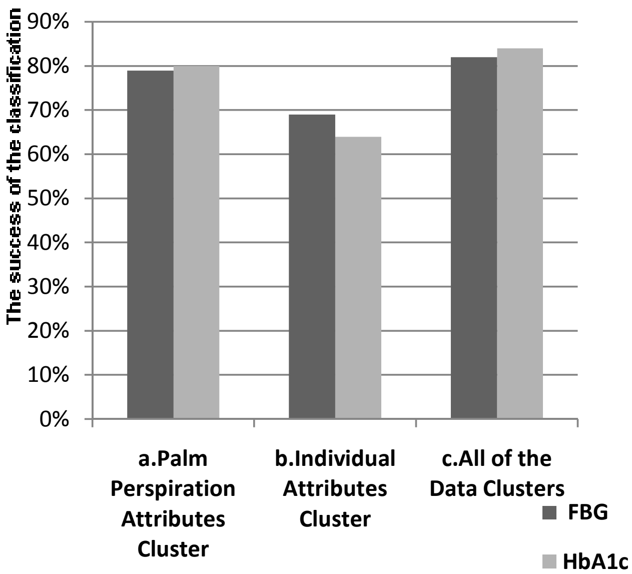


Fig. 6 Results of diabetes diagnosis with of SVM

In accordance with Column a in Figure 6, in the classification done by using the palm perspiration attributes of the subjects;

- Fasting blood glucose reference classification achieved 78% success.

- 80% success is achieved in HbA1c reference classification

In accordance with Column b in Figure 6, in the classification done by using the individual attributes (age, body mass index, systole, diastole) of the subjects;

- Fasting blood glucose reference classification achieved 69% success.
- 64% success is achieved in HbA1c reference classification

In accordance with Column c in Figure 6, in the classification done by using all the data clusters (palm perspiration attributes + individual attributes);

- Fasting blood glucose reference classification achieved 82% success.
- 84% success is achieved in HbA1c reference classification.

As regards the gathered results, perspiration attributes achieved 80% accuracy in the diagnosis of diabetes conditions of people. Therefore, the relation between perspiration in human and diabetes is a subject to be researched.

ACKNOWLEDGMENT

This work was fully supported under TUBITAK (The Scientific and Technological Research Council of Turkey-project no. 109E234). Data used in this study were obtained from Dumlupınar University Kütahya Evliya Çelebi Education and Research Hospital. The authors would like to thank for their kind supports.

REFERENCES

[1] Arısoy, E.S., *T.C. Ministry of Health, Diagnosis and Treatment Guides for Primary Healthcare 2003,*T.C. Miistry of Health, 2003

[2] Wild S., Roglic G, Green A, Sicree R, King H. "Global prevalence of diabetes estimates for the year 2000 and projections for 2030", Diabetes Care, 2004, Vol 27, pp 1047-1053

[3] *International Diabetes Federation. Diabetes Atlas*, World Diabetes Foundation, 2006

[4] Satman I., "The update criteria and the reasons of them in diagnosis and follow up of diabetes mellitus",A Türkiye Klinikleri Journal of Internal Medical Sciences, 2007,Vol 3,pp 1-15

[5] Temurtas, H., Yumusak, N., & Temurtas, F. (2009), "A comparative study on diabetes disease diagnosis using neural networks", Expert Systems with Applications, 36, 8610–8615.

[6] Mohamed, E. I., Linderm, R., Perriello, G., Di Daniele, N., Poppl, S. J., & De Lorenzo, A. (2002), "Predicting type 2 diabetes using an electronic nose-base artificial neural network analysis", Diabetes Nutrition & Metabolism, 15(4), 215–221.

[7] Saraoglu, H.M., & Kocan, M. (2010), "A study on non-invasive detection of blood glucose concentration from human palm perspiration by using artificial neural networks", Expert Systems, 27(3), 156-165.

[8] *Diabetes Care*, American Diabetes Association, January 2008, Vol 31,pp 55-60

[9] *Diabetes Care*, American Diabetes Association, January 2010, Vol 33,pp 11-13

[10] *Diabetes Voice*, December 2007, Vol 52, pp 33-34

[11] Kamei, T., T. Tsuda, Y. Mibu, S. Kitagawa, H. Wada, K. Naitoh and K. Nakashima (1998a), "Novel instrumentation for determination of ethanol concentrations in human perspiration by gas chromatography and a good interrelationship between ethanol concentrations in sweat and blood", Analytica Chimica Acta, 365 (1–3), 259–266.

[12] Kamei, T., T. Tsuda, S. Kitagawa, K. Naitoh, K. Nakashima and T. Ohhashi (1998b), "Physical stimuli and emotional stress-induced sweat secretions in the human palm and forehead", Analytica Chimica Acta, 365, 319–326

[13] Saraoglu, H. M. ve Koçan M., "Determination of Blood Glucose Level Based Breath Analysis by a Quartz Crystal Microbalance Sensor Array", IEEE Sensors Journal, Special Issue on the Future of Sensors and Instrumentation for Human Breath Analysis, 10, 104-109, 2010

[14] Masumura, M., S. Shinomiya, K. Nagata, T. Kato, Y. Matsuzaki, S. Nagaoka, R. Inoue, J. Perspiration Res., 3, 43-45, 1996

[15] Kuno, Y., *Human Perspiration*, C.C. Thomas Publisher, Springfield,IL, 1956.

[16] Ohhashi T., M. Uono (eds.), "*Palmer perspiration – measurement and clinical applications (in Japanese)*", Life Medicom Co, Ltd., Nagoya, 1993.

[17] Groscurth, P., "*Anatomy of sweat glands*". In: Kreyden OP, Bo¨ni R, Burg G, editors. Hyperhidrosis and botulinum toxin in dermatology. Basel: Karger, 1–9, 2002.

[18] Kamei, T., T. Tsuda, S. Kitagawa, K. Naitoh, K. Nakashima, T. Ohhashi, "Physical stimuli and emotional stress-induced sweat secretions in the human palm and forehead", Analytica Chimica Acta, 365, 319-326, 1998 b

[19] Huestis, M.A., J.M. Oyler , E.J. Cone , A.T. Wstadik , D. Schoendorfer , R.E. Joseph JR., "Sweat testing for cocaine, codeine and metabolites by gas chromatography–mass spectrometry, Journal of Chromatography ", 733, 247–264, 1999.

[20] Chang, B.W., S.J. Yeh, P.P. Tsai, H.C. Chang, "Monitoring perspiration from palms of hypohidrosis patients with a stopped-flow conductometric mini-system",Clinica Chimica Acta, 348, 107– 111, 2004.

[21] Asahina, M., A. Suzuki, M. Mori, T. Kanesaka, T. Hattori, "Emotional sweating response in a patient with bilateral amygdala damage", International Journal of Psychophysiology, 47, 87–93, 2003.

[22] Saraoglu, H. M. ve Koçan M., "A Study on Non-Invasive Detection of Blood Glucose Concentration from Human Palm Perspiration by Using Artificial Neural Networks", Expert Systems, 27(3), 156-165, 2010

[23] Vapnik, V. , *Statistical Learning Theory*, Wiley,1998

[24] Brownz M.P.S., Grundyz W.N., Linz D., Cristianini N., Sugnet C., Ares, JR. M., Hausslerz D., "Support Vector Machine Classification of Microarray Gene Expression Data", UCSC-CRL-99-09

[25] Burges, C. J. C. (1998). "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowledge Discovery,* 2(2):121–167.

[26] V.N. Vapnik, "An Overview of Statistical Learning Theory*", IEEE Trans. Neural Networks 10 (5)* (1999).

[27] Saraoglu, H.M., Yildirim, M., Ozbeyaz, A., TEMURTAS, F., (2011), "Kandaki Glikoz ve HbA1c Diyabet Tanı Değerlerinin Avuç İçi Terleme Ölçümleri Kullanılarak DVM ile Sınıflandırması", In Proceedings of Tip Teknolojileri Ulusal Kongresi (Tip Tekno 11), October 13-16, Antalya

[28]](2012),Univesty of Strathclyde website. [Online]. Available:http://www.strath.ac.uk/media/departments/eee/cesip/cesipseminar/Jinchang_Ren_seminar. Pdf

[29]](2012), A-SVM based server for rice website. [Online]. Available:www.imtech.res.in/raghava/rbpred/algorithm.html.