The Method of Applying the Algorithm of Calculating Grades for Finding Similar Diagnostics in Medical Information Systems

M.M.Kamilov, O.K.Khujaev, N.A.Egamberdiyev, A.X. Chorshanbiyev

Abstract: ***This paper focuses on the problem of ranking, where is one of the issues of intelligent data analysis based on the medical information system database. Given common mathematical description of ranking task and analyzed RankBoost, RankSVM, IR-SVM algorithms for solving ranking task and also said its applications. In addition, the neuronal network, which is heuristic method, has been designed to solve the problem of ranking to the feed forward neural network model. And so on six stages calculating ranks algorithm for classification task is applied and are taken results for searching similar diagnosis in database of medical information system and redesigned in three stages. Results checked and compared algorithm accuracy on based database of medical information system. Moreover flowchart of adopted variant of calculating ranks algorithm for solving ranking task is given.***

Keywords: Ranking, RankBoost, RankSVM, IR-SVM, feed forward neural network, algorithm of calculating ranks.

I. INTRODUCTION

Usually, Information about patient which is various laboratory tests, decisions of physicians are stored in relational database by database management system of medical information systems. Sometimes, physicians are required by the results of other physicians' diagnostics. It is a complex process that takes a lot of time to select thousands of patients in the database, choosing the best among them, and selecting optimal diagnostic information.
In this case, it is possible to select optimal solutions by solving the ranking task of machine learning.

II. MATHEMATICAL DESCRIBING OF THE TASK

There are given $I = \{i_1, i_2, \dots, i_j, \dots, i_n\}$ set of objects, which are taken from database of medical information system. Every particular $i_j$object is a result of various laboratory tests of every patient and is described with following parameters: $i_j = \{x_1, x_2, \dots, x_h, \dots, x_m\}$.

For any (i, j) ∈ {1, 2,.., n} pair of indexes of objects, given $k<l$ordering. It is necessary for us to create a function: a: X → R, for a given $k<l$, $a(i_k) < a(i_l)$to complete the condition. Let's call *a* function as a ranking function [13].
In our case, the question is finding *m* the closest object for the newly created i_{n+1} object. In this case, we set the approximation criteria for the other objects relative to the new i_{n+1}object. As a result, the first m objectwhich is taken from the list is the objects we are looking for.

III. EXISTING METHODS ANALYSIS

Ranking task is mostly used to sort documents according to certain criteria, rating links on the Internet, and use large-capacity of solutions such as Google, Yandex, Yahoo, Bing and Bing to search for solutions on their search engines. [9, 10, 11].

So far, many methods and algorithms have been proposed to solve the task of ranking. These are linear regression and classification algorithms based on a streamlined approach, RankNet, FRANK, RankBoost, RankSVM, IR-SVM based on pairing approaches. SoftRank, SVMmap, AdaRank, RankGP, ListNet, ListMLE algorithms based on algorithms and list of validated approach [12].

In addition, solving problem of ranking with feed-forward neural network model gives us good results. The use of this model of neural network in the ranking of electronic documents is considered in [7,8] articles. However main weakness of feed-forward neural network model is large-scale calculations of the training process. Therefore, this method is not very effective [7].

The main aspect of the ranking task is to select the attributes. In our case, attributes are the results of various laboratory tests in the database of the information system of cardiology hospitals, which are numerically expressed.

IV. THE PROPOSED METHOD

In the data mining, the algorithm of calculating grades is used to solve the classification task. The rating calculation algorithm consists of six stages, each stage consists of sequence of specific tasks [1,2]. These steps are as follows:

1. System of support sets
2. The function of proximity
3. Calculating grades by attributes of fixed support set
4. Calculating grades by classes of fixed support set
5. Grade of K_u class by systems of support sets
6. Decisive rule for algorithm

Revised Manuscript Received on December 22, 2018.

M.M.Kamilov, Department of Software Engineering, Tashkent University of Information technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

O.K.Khujaev, Head of department of Information Education technologies, Urgench branch of Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

N.A.Egamberdiyev, Department of Software Engineering, Tashkent University of Information technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

A.X. Chorshanbiyev, Department of Software Engineering, Tashkent University of Information technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

All of these steps need to solve the classification task in data mining. But in our case, it is equally important to compute the first three steps, namely System of support sets, the function of proximity, and Calculating grades by attributes of fixed support set. We analyze the computational processes that are going to take place at these steps:

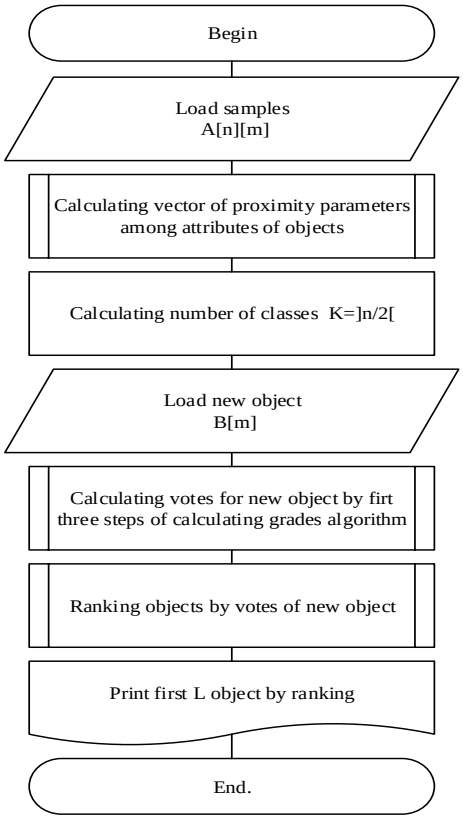
1. System of support sets. At this stage, we load the results of the patient's various laboratory tests from database and create support sets in the matrix form. n is a number of objects and m is an attribute of each object:
2. The function of proximity. At this stage, we calculate the proximity parameter for each attribute on based attributes of support set objects:

$$\varepsilon_j = \frac{1}{n-1}\sum_{i=1}^{n-1}\left|x_{ij} - x_{ij+1}\right| \quad (1)$$

3. Calculating grades by attributes of fixed support set. This step is the main stage of our algorithm, where the votes of each object are compared to the new object. In the algorithm of the calculating of grades, the votes of the objects were taken not only on the basis of the value given to each object of the subject, but also on the grouping of attributes. The grouping coefficient can be k = {1,2, .., n}. However, counting the votes for the total value of k and choosing the best of it requires countless calculations. In this case, the complexity of the algorithm makes more calculations than the neural network. Therefore, choosing the optimal value for k is an important factor in the problem solving process. The experimental results show that the k to the number of clusters is k =] n / 2 [it is desirable to take the form.

After selecting *k*, we calculate the value of the test object for each object in the base collection and create the s[n] vector. And s[n] vector is arranged in descending order, and the first *L* object is the solution of the problem based on the resulting vector.

The flowchart of the algorithm is as follows:



V. EXPERIMENT ANALYSIS

Experimental data is taken from Khorezm branch of Republican exclusive cardiac center. Data is taken from 21 patients for 6 laboratory tests of 90 patients by"iCardio"medical information system. Experimentalresults are as follows:

Initially, we consider the vector of proximity, which is as follows:

Table. 1 Computing results of vector of proximity

e[1]	7.2045	e[8]	4.7193	e[15]	41.3795
e[2]	0.4380	e[9]	16.7750	e[16]	12.6136
e[3]	133.2772	e[10]	55.0840	e[17]	0.41022
e[4]	0.0579	e[11]	21.35	e[18]	0.17954
e[5]	0.0250	e[12]	0.8352	e[19]	0.18068
e[6]	0.4318	e[13]	0.9727	e[20]	1.83522
e[7]	2,75	e[14]	55.4545	e[21]	1.04886

Then a test object is added,

Xolbaev__Pulat__ 102.0, 3.91, 3.8, 1.4, 4.4, 35.8, 16.0, 1.9, 64, 330, 50.0, 6.0, 3.9, 184.0, 66.0, 64.0, 0.7, 0.9, 3.6, 4.6, 2.4

and the evaluations given by the objects are calculated for k = 10.

352716 combination

1-closet object isEginmatov_Jumageldi 113.0, 4.06, 5.4, 1.4, 4.4, 35.8, 7.0, 1.3, 66.0, 325.0, 29.6, 5.8, 4.5, 164.0, 91.0, 44.0, 1.1, 1.0, 3.5, 4.3, 1.9 with 184756.0 votes;

2-closet object is Yakubova_Xoljon 109.0, 3.91, 4.1, 1.4, 4.4, 35.8, 6.0, 1.7, 68.0, 302.0, 23.0, 5.7, 4.7, 161.0, 105.0, 35.0, 1.0, 1.2, 3.6, 4.2, 2.1, with 184756.0 votes;

3 - closet object isNajmiddinov_Erkin 142.0, 4.7, 7.2, 1.4, 4.4, 35.8, 7.0, 1.1, 92.0, 398.0, 14.8, 5.1, 3.7, 126.0, 58.0, 54.0, 0.9, 1.0, 3.2, 3.6, 2.2 19 with 92378.0 votes;

4 closet object is Axmedov_Baxiyor 141.0, 4.38, 5.3, 1.4, 4.4, 35.8, 7.0, 1.8, 69.0, 335.0, 40.0, 5.3, 3.7, 140.0, 57.0, 57.0, 1.3, 0.9, 3.2, 3.0, 1.9with 92378.0 votes

The total combinations for the total number of votes is equal to 352716, and first rank received is 48-objet, and the object is collected 187756 votes, so this object is the most closest object to our test object. In this way, we can find the most closest *L* objects.

VI. CONCLUSIONS

In conclusion, it can be said that nowadays, information systems have been developed and used to improve the efficiency of the work process not only in cardiology, but also in almost all medical institutions. Many medical information systems are integrated into the intelligent analysis modules and the system is being expanded. The problem of racing is one of the important issues in this process. In this process, the use of the above-mentioned modified version of the algorithm calculationof grades and the algorithm will give a good result in increasing the reliability of system performance.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

REFERENCES

1. JuravlevYu.I. Izbrannienauchnietrudi. –M.: Magistr, 1998. –420 s.
2. JuravlevYu.I., KamilovM.M., TulyaganovSh.Ye. Algoritmivichisleniyaotsenokiixprimenenie. Tashkent: Fan, 1974. - 119s.
3. Kamilov M.M., Hudayberdiev M.X., Khamroev A.Sh. Methods ofComputing Epsilon Thresholds in the Estimates' Calculation's Algorithms. International Conference "Problems of Cybernetics and Informatics" (PCI'2012),Volume III. September 12-14, 2012. – Baku, Azerbaijan. – Pp. 133-135.
4. XamroevA.Sh.Algoritmviboraoptimalnogometodavichisleniyaznacheni yye-porogovvalgoritm axvichisleniyaotsenok. Mejdunarodniynauchno-texnicheskiy jurnal. Ximicheskayatexnologiya. Kontroliupravlenie. – Tashkent, 2012. – № 3. – S. 78-82.
5. XudayberdievM.X., XamroevA.X., MamievaD.Z. Ob'ektlar haqidagio'quvvanazorattanlanmalarinishakllantirishdabaholarnihisobla shalgoritmi. Respublika ilmiy-texnik anjumani: "Axborot va telekommunikatsiya texnologiyalari muammolari". - Toshkent, 2016 yil 10-11 mart. – 185-187 bb.
6. Xudayberdiev M.X., Xamroev A.Sh. O vzaimosvyazi parametrov v modeley algoritmov vichisleniya otsenok. Intellektualnie sistemi (INTELS'-2014): Desyatiy mejdunarodniy simpozium. 30 iyunya - 4 iyulya –Moskva, 2014. – S. 49-52.
7. Xo'jaev O.Q., Abdullaev F.O., Raxmanova M.R. "Metodi opredelenie vesa dokumentov v elektronnix resursax" // NAUKA I MIR Mejdunarodniy nauchniy jurnal, № 3 (19), 2015, Tom 1 ISSN 2308-4804.
8. Xo'jaev O.Q. Abdullaev F.O., Artikov M.E., Bobojanov B. "Primenenie neyronnix seteyv poiskovix mashinax tipa elektronnogo pravitelstva"// NAUKA I MIR mejdunarodniy nauchniy jurnal, № 3 (19), 2015, Tom 1 ISSN 2308-4804.
9. Tie-Yan Liu (2009), Learning to Rank for Information Retrieval, Foundations and Trends in Information Retrieval: Vol. 3: No 3, ss. 225-331, ISBN 978-1-60198-244-5, DOI 1561/1500000016.