World Bulletin of Social Sciences (WBSS)
Available Online at: https://www.scholarexpress.net
Vol. 3, October- 2021
ISSN: 2749-361X

DETERMINATION OF PRINCIPAL SYMPTOMS IN THE DIVISION OF OBJECTS IN CLASSES

D.X.Tursunmurotov
Teacher of Karshi branch of TUIT named after Muhammad al-Khwarizmi
+998914732838
mr.davrbek@mail.ru

Article history:

Received:	August 11th 2021
Accepted:	September 13th 2021
Published:	October 18th 2021

Abstract:

This paper discusses the problem of calculating the weights of quantitative symptoms on the basis of a given criterion. There is no clear distinction between the terms "symptom weight" and "symptom contribution" in terms of content. Sign weight refers to the degree to which that symptom contributes to the classification of objects.

A computational experiment was conducted using a sample consisting of several quantitative symptoms, as an example, a collected sample for the detection of diabetes mellitus was taken and the results were described in tables.

Keywords: Symptom weight, symptom contribution, quantitative symptoms, artificial intelligence, compactness hypothesis, teacher-assisted learning.

I. INTRODUCTION

Today, as a result of the rapid development of science and technology, the issue of systematization and automation of all industries is becoming increasingly important. There is a growing need for intelligent systems, especially those based on artificial intelligence. There are many branches of artificial intelligence, one of which is machine learning. The main component of machine learning is teacher-assisted learning, which mainly solves the problem of classification of objects, ie the division into classes. Weight plays a key role in determining to what extent an object classification belongs to its class.

Weight calculation methods are aimed at solving comprehension problems with and without a teacher. It is well known that there are no methods of division into all recognized classes. Therefore, conditional and unconditional optimization algorithms can be used in the computational process. In terms of content, the terms "symptom weight" and "symptom contribution" are the same concepts. The essence of the criteria used to calculate the weight and contribution of a symptom is based on verifying the validity of the compactness hypothesis. The calculation of weights can be put in the form of solving conditional and unconditional optimization problems.

II. MAIN PART

One of the important issues is to reduce the number of symptoms in the description of the objects of selection on the basis of a certain criterion. This importance is explained by the need to identify a low-informational and "redundant" symptom (inexplicably linked to other symptoms), to build a subset of minimum symptoms for cases with high cost constraints, or to address the problem of lack of data.

Symptom weights are used for the following purposes:
to calculate the size of the proximity between objects;
- in the selection and sorting of information signs;
- to model the intuitive decision-making process in the search for laws;
- generalized estimates in order to reduce the space of symptoms (latent symptoms) in the calculation [7].

Character selection is an effective way to improve the performance of machine algorithms. Thus, the use of useful cues for machine learning significantly improves efficiency. Instead of treating all symptoms the same, a better result can be achieved if the issue is resolved based on the importance of the symptoms. On the other hand, building a minimum symptom space is a source of additional knowledge on selection. Determining the severity of symptoms is one of the important tasks in carrying out such a process.

The problem of understanding the images placed as standard is considered.

The two do not intersect K_1, K_2 containing representatives of the classes $E_0 = \{S_1, \ldots, S_m\}$ a set of objects is considered to be given. A possible object of choice n та $X = \{x_1, \ldots, x_n\}$ characterized by quantitative symptoms, for convenience we consider class representatives as cases of occurrence and - cases of non-occurrence.

World Bulletin of Social Sciences (WBSS)
Available Online at: https://www.scholarexpress.net
Vol. 3, October- 2021
ISSN: 2749-361X

One of the reasons for considering a two-class problem is that the generalized value of any object is relative, which results from a comparison with objects of the opposite class. Second, any (k> 2) class problem can be solved in a cascading view of two class problems.

A set of selection objects $x_j, j \in I$ we sort it by the non-decreasing (non-increasing) location of the symptom values and $[c_1, c_2], [c_2, c_3]$ intervals. Interval limit - c_2 the determination of the value is based on the assertion that "each interval contains the values of the quantitative characteristics of one class of objects" [1].

Let's say - u_i^1, u_i^2 in line $[c_1, c_2], [c_2, c_3]$ intervals $x_j, j \in I$ of the sign $K_i = 1, 2$ the amount of values belonging to the class $p - E_0$ The objects in the plaid are arranged in a non-decreasing order $r_{j1}, ..., r_{jp}, ..., r_{jm}$ the number of the object with the interval limit in the sequence, i.e. $c_1 = r_{j1}, c_2 = r_{jp}, c_3 = r_{jm}$ [2] in that case

$$\left(\frac{\sum_{i=1}^{2} u_i^1(u_i^1 - 1) + u_i^2(u_i^2 - 1)}{\sum_{i=1}^{2} |K_i| \left(|K_i| - 1 \right)} \right) \left(\frac{\sum_{d=1}^{2} \sum_{i=1}^{2} u_i^d \left(|K_{3-i}| - u_{3-i}^d \right)}{2|K_1||K_2|} \right) \to \max \quad (1.1)$$

Criterion $[c_1, c_2], [c_2, c_3]$ allows you to calculate the optimal limits of the intervals. The expression within the first bracket in the criterion (1.1) is the similarity within the class, the expression in the second bracket is the degree of difference between the classes [1].

$w_k -$ (1.1) of the criterion $x_k, k \in I$ the optimal value for a sign is k, a quantitative value that reflects the degree to which the objects in the instructional sample are classified, i.e., the weight of that sign.

As a computational experiment, a sample formed based on the data collected was used to predict whether the patient had diabetes mellitus. The sample consisted of 14 quantitative signs from 390 objects (patients) and was selected from kaggle.com.

Distribution of symptoms by weight and intervals

Table 1.1.

Аломат	**c_2– қиймати**	**Оралиқ интерваллари**	**w_i қиймати**
Glucose	313	48.000...118.000...385.000	0.640504
Gender	26	0.000...0.000...1.000	0.526500
Age	269	19.000...56.000...92.000	0.550718
systolic_bp	298	90.000...150.000...250.000	0.534895
chol_hdl	297	1.500...5.500...19.300	0.516436
Waist	294	26.000...4.000...56.000	0.512133
hdl_chol	109	12.000...40.000...120.000	0.501512
Cholesterol	297	78.000...232.000...443.000	0.300625
Weight	265	99.000...189.000...325.000	0.298322
waist_hip_ratio	303	0.680...0.940...1.140	0.291917
Bmi	261	15.200...30.500...55800	0.287985
Hip	296	30.000...47.000...64.000	0.284731
diastolic_bp	235	48.000...87.000...124.000	0.266321
Height	85	52.000...63.000...76.000	0.263353

1.1 жадвалда кўриниб турибдики, glucose, gender, age, systolic_bp, chol_hdl, waist, hdl_chol аломатларда K_1 and K_2 the division into classes was relatively strong quantitative symptoms were calculated according to the intervals (1.1) according to formula K_1 and K_2 an ordered view of the contributions to the separation of classes is given in Table 1.2.

World Bulletin of Social Sciences (WBSS)
Available Online at: https://www.scholarexpress.net
Vol. 3, October- 2021
ISSN: 2749-361X

Contributions of symptoms to classification.

т/н	Аломат	Аломатнинг синфга ажратишдаги ҳиссаси
1	Glucose	0.640504
2	Gender	0.526500
3	Age	0.550718
4	systolic_bp	0.534895
5	chol_hdl	0.516436
6	Waist	0.512133
7	hdl_chol	0.501512
8	Cholesterol	0.300625
9	Weight	0.298322
10	waist_hip_ratio	0.291917
11	Bmi	0.287985
12	Hip	0.284731
13	diastolic_bp	0.266321
14	Height	0.263353

The analyzes in Table 1.2 show that the contribution of glucose, gender, age, systolic _b p, chol _ hdl, waist, hdl_ chol symptoms to the class separation of selection objects is relatively large.

CONCLUSION

We have seen above that determining the weight of a character, i.e. its contribution, is one of the main parameters that indicate the degree of class distinction in the selection of an object. Symptoms are also one of the main processes performed in the study of weight selection objects. Including generalized indicators and their application to man in cases of application of the method of generalized assessment in different areas of activity. One of the most pressing problems in the study of multidimensional, complex phenomena and processes is the construction of summary, generalized estimates on a set of indicators that describe these processes. On the basis of generalized indicators can be used in the management of enterprises and organizations, assessing their financial stability, assessing the psychological state of the individual, determining the degree of disease of the patient and solving other similar issues.

REFERENCES

1. Ignatev N.A., Madraximov Sh.F., Xurramov A.X. Intervalnyy method of calculation of general values // Trudy konferentsii Aktualnyeproblemy prikladnoy matematiki i informatsionnyx tehnologii - Al Khorezmi 2009, Tashkent, September 18-21, 2009, Tom 1, pp. 154-155.
2. Ignatev N.A., Matlatipov G.R. Avtomaticheskoe vydvijenie gipotez v sisteme usileniya intellekta // Vestn. TashGU. - 1998.- № 3. - C. 22 - 25.
3. Kruglov V. V., Dli M. I., Golunov R. Yu. Nechetkaya logika i iskusstvennые neyronnye seti: Uchebnoe posobie - M .: Izdatelstvo Fiziko-matematicheskoy literatury, 2001. - 224 s. [4]. Jackson P. Introduction to expert systems.- Izdatelskiy dom «Williams», 2001.
4. Gavrilov A.V.Sistemy iskusstvennogo intellekta.-Novosibirsk: NGTU, 2004.-59p. [6]. Andon F.I., Balabanov A.S. Vyyavlenie znaniy i izyskaniya v bazax dannyx: podkhody, modeli, metody i sistemy (obzor) // Problemsof programming. 2000. № 1-2, p. 513-526.
5. Vapnik V.N. Restoration of dependence on empirical data. - M .: Nauka, 1979. - 447 p.