UDC 681.518.5

L.S. Fainzilberg

ON PROBABILITIES OF MISSED TARGET AND FALSE ALARM IN MEDICAL DECISION-MAKING SYSTEMS

Leonid Fainzilberg

Information Technologies and Systems of the National Academy of Sciences of Ukraine and the Ministry of Education and Science of Ukraine, National Technical University of Ukraine «Igor Sikorsky Polytechnic Institute»,

fainzilberg@gmail.com

Binary classifiers are widely used in solving a number of applied problems, in particular, problems of medical diagnostics. The effectiveness of such systems is characterized by target miss errors and false alarms. Since the probability of a false alarm error usually increases with a decrease in the probability of an allowance error, and vice versa, the tuning of systems implies a certain compromise between the indicated errors. To build a diagnostic algorithm under a priori uncertainty, the Neumann-Pearson strategy is often used, which involves minimizing the probability of a target miss error with a given constraint on the probability of a false alarm error. To study the question of the formal choice of an acceptable constraint, conditions for the usefulness of a diagnostic algorithm are formulated. On the basis of these conditions, the admissible limits of the probabilities of target miss errors and false alarms are determined. The boundary values of the prevalence of diseases are determined, at which a diagnostic algorithm with known operational characteristics remains useful in terms of reducing average losses. Examples are given to illustrate the practical value of the obtained conditions. The features of the decision-making algorithm based on medical symptoms are considered. It has been shown that if, according to medical statistics, a certain symptom is typical for most sick patients and uncharacteristic for most healthy people, but this does not mean high reliability of diagnostic decisions based on the analysis of this symptom in a particular person. Using the example of the analysis of the symptom «Characteristic skin rashes», it was demonstrated that even if the probability of the absence of a symptom in healthy people is 0,99, the probability of a false alarm may be lower than 0,01. This is explained by the fact that in order to make an substantiated decision for a particular patient, it is necessary to make sure that the indicated symptom is absent during a certain observation interval.

Keywords: missed target, false alarm, diagnostic algorithm usefulness, Neman-Pearson strategy, disease prevalence, medical symptom.

Introduction

Binary classifiers are widely used in computer systems for solving problems of medical diagnostics [1], detection of enemy targets [2], spam filtering [3], identification of unreliable bank customers [4], etc. The efficiency of such systems is characterized by missed target and false alarm errors [5]. With a decrease in the probability of missed target error, the probability of a false alarm error usually increases and vice versa [6]. Therefore, tuning systems involves a certain compromise between these errors.

If a priori probabilities of recognizable classes and losses from errors of the first and second kind are known, then optimal decisions can be made on the basis of a Bayesian strategy that minimizes the average risk of misclassification [7]. In the absence of such information, a number of other strategies are used, in particular, the Neumann-Pearson strategy [8-10], which involves minimizing the probability of a target miss error with a given restriction on the probability of a false alarm error.

Since there are no recommendations in the scientific literature on the choice of such a restriction, it becomes necessary to obtain formal conditions that allow one to set an acceptable constraint.

To study this issue, consider the task of screening, which is to identify sick people in a large group of patients, where there are both sick (class V_1) and conditionally healthy people who do not have the disease in question (class V_2). The diagnostic algorithm (binary classifier), based on the measurement of a physiological indicator $x \in X$ makes one of two decisions:

$$\begin{matrix} \text{person is sick}, & \text{if } x \in X_1, \\ \text{person is helthy}, & \text{if } x \in X_2, \end{matrix} \quad X_1 \cup X_2 = X\ , \ X_1 \cap X_2 = \varnothing\ , \tag{1}$$

where X_1, X_2 are areas of relevant solutions.

It is assumed that when using algorithm (1), two types of errors are possible:

$$E_1, \text{ if } x \in X_2, \text{ but person is ill}\ , \tag{2}$$

$$E_2, \text{ if } x \in X_1, \text{ but person is helthy}\ . \tag{3}$$

Random events (2) and (3) are commonly referred to as missed targets and false alarms, respectively. With known conditional distributions of x values in sick $p(x\,|\,V_1)$ and healthy people $p(x\,|\,V_2)$, the probabilities of these events are determined by the expressions:

$$P(E_1) = \int\limits_{X_2} p(x\,|\,V_1)dx\ ,$$

$$P(E_2) = \int\limits_{X_1} p(x\,|\,V_2)dx\ ,$$

If $p(x\,|\,V_1)$ and $p(x\,|\,V_2)$ unknown, then the probabilities $P(E_1)$ and $P(E_2)$ are estimated experimentally from samples of observations $x \in X$ in people with predetermined diagnoses.

In accordance with the Neiman-Pearson strategy, the regions $X_1 \subset X$ and $X_2 \subset X$ are chosen so that

$$P(E_1) \to \min\ , \tag{4}$$

$$P(E_2) \le \varepsilon\ , \tag{5}$$

where $P(E_1)$, $P(E_2)$ are the known probabilities of events (2) and (3) here ε is a given restriction on the false alarm error probability.

General conditions for the usefulness of a diagnostic algorithm

We assume that the classes V_1 and V_2 are random events for which objectively exist a priori probabilities $P(V_1)$ and $P(V_1) = 1 - P(V_1)$, although unknown to us. According to [11], the Neyman-Pearson strategy is precisely focused on such a practically important case of a priori uncertainty. We will also assume that we do not know the absolute values of losses L_{12} and L_{21}, which cause target misses and false alarms errors, but we can have information about the ratio ω of these losses:

$$\omega = \frac{L_{12}}{L_{21}}. \tag{6}$$

The value ω determines how many times the losses L_{12} caused by the error E_1 exceed the losses L_{21} because of the error E_2. Unlike absolute values L_{12} and L_{21}, such limited information can often be known to the designer of the applied system.

To choose an acceptable value for the restriction ε appearing on the right side of inequality (5), we use the definition of the utility of a diagnostic algorithm (1). Based on pragmatic considerations, a diagnostic algorithm can be considered useful only if the decisions made on its basis regarding classes V_1 and V_2 provide average losses less than those that would be observed without the use of the algorithm [12]. In other words, we will associate the concept of utility with the following definition:

Definition 1. *Algorithm (1) is useful if the strict inequality holds*

$$R < R_0, \tag{7}$$

i.e., the posterior risk R based on the measurement $x \in X$ is less than the prior risk R_0.

Taking into account the introduced notation, we write the a posteriori risk R in the form:

$$R = L_{12}P(V_1)P(E_1) + L_{21}[1 - P(V_1)]P(E_2). \tag{8}$$

Obviously, if we do not use the diagnostic algorithm (1) based on the measurement of the physiological indicator $x \in X$, then the a priori «strategy» of decisions is reduced to choosing one of two possible options:

- allways make a decision in favor V_1 and then the average losses will be equal to $L_{21}(1 - P(V_1))$;
- allways person make a decision in favor V_2 and then the average losses will be $L_{12}P(V_1)$.

It follows that the minimum average losses will be ensured if a priori decisions are made in favor of V_1 when condition holds

$$L_{12}P(V_1) < L_{21}[(1 - P(V_1)], \tag{9}$$

or decisions in favor V_2 when condition holds

$$L_{21}[(1 - P(V_1)] < L_{12}P(V_1). \tag{10}$$

From expressions (9), (10), taking into account notation (6), it follows that the a priori risk R_0 appearing on the right side of (7) can be represented as

$$R_0 = \begin{cases} L_{12}P(V_1), & \text{if } P(V_1)(1+\omega) < 1, \\ L_{21}[(1 - P(V_1)), & \text{if } P(V_1)(1+\omega) > 1. \end{cases} \tag{11}$$

Substituting (8) and (11) into (7) after simple transformations allows to formulate the following statement:

Statement 1. Diagnostic algorithm (1) is useful in the sense of (7) if and only if

$$[(1-P(V_1)]P(E_2) < \omega P(V_1)[1-P(E_1)]\text{, when } P(V_1)(1+\omega) < 1, \tag{12}$$

or

$$[1-P(V_1)][1-P(E_2] > \omega P(V_1)P(E_1)\text{, when } P(V_1)(1+\omega) > 1. \tag{13}$$

Permissible bounds for the probabilities of missed targets and false alarm

It follows from Statement 1 that if condition (12) or condition (13) are not satisfied then algorithm (1) is absolutely useless from the point of view of reducing average losses. This fact allows us to determine the admissible value of the restriction ε that appears on the right side of (5).

Statement 2. *Let the condition* $P(V_1)(1+\omega) < 1$ *be satisfied. Then, for any arbitrarily small probability of missed target error, algorithm (1) is certainly useless in the sense of (7) if the false alarm error probability* $P(E_2)$ *satisfies the condition*

$$P(E_2) \geq \frac{\omega P(V_1)}{1-P(V_1)}. \tag{14}$$

Proof. Taking into account that $P(V_1) \neq 1$, we represent condition (12), which guarantees the usefulness of algorithm (1) for the domain $P(V_1)(1+\omega) < 1$, in the equivalent form

$$P(E_2) < \frac{\omega P(V_1)[1-P(E_1)]}{1-P(V_1)}. \tag{15}$$

Based on (15), we conclude that algorithm (1) is useless if

$$P(E_2) \geq \frac{\omega P(V_1)[1-P(E_1)]}{1-P(V_1)}. \tag{16}$$

Since $P(E_1) \geq 0$, then amplification (16) by substitution $P(E_1) = 0$ leads to estimate the admissible value of the false alarm error probability $P(E_2)$ in the form of condition (14). Assertion 2 is proved.

Based on expression (14), an important consequence may be formulated.

Consequence 1. For $P(V_1)(1+\omega) < 1$ the algorithm (1) that implements the Neyman-Pearson strategy (4), (5) to be useful in terms of reducing average losses, the restriction ε on the right side of (5) must satisfy the condition

$$\varepsilon < \frac{\omega P(V_1)}{1-P(V_1)}. \tag{17}$$

Here it is necessary to give a little explanation. As has already been noted, the Neumann-Pearson strategy is focused on the case when the a priori probability $P(V_1)$ is unknown, for example, the prevalence of a little-studied disease is unknown. Therefore, condition (17) has only an epistemological meaning and it is legitimate to consider it as a condition for the existence of an admissible restriction ε on the probability of a false alarm error $P(E_2)$.

 ISSN 2786-6491

At the same time, based on (17), after simple transformation, it is possible to solve the inverse problem: to determine the boundary value $P_\omega(\varepsilon)$, which depends on the chosen restriction ε and fixed loss ratio ω and determined by the expression:

$$P_\omega(\varepsilon) = \frac{\varepsilon}{\omega + \varepsilon}. \tag{18}$$

Expression (18) is already of practical use. If the prevalence $P(V)$ of the studied disease does not exceed the value (18), i.e. $P(V_1) \le P_\omega(\varepsilon)$, then the Neumann-Pearson strategy will be ineffective.

Figure 1 shows plots of dependence $P_\omega(\varepsilon)$ on ε constructed in accordance with (18) for various values of ω.

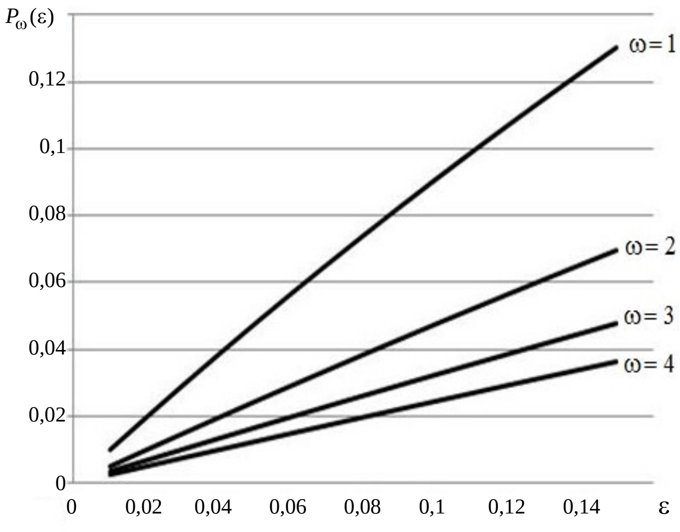

Fig. 1

Example 1. To examine a group of 10 000 people, it is supposed to use a diagnostic test that implements the Neumann-Pearson strategy. The loss ratio (6) is selected $\omega = 3$. If we set a restriction $\varepsilon = 0,07$ on the probability $P(E_2)$ of a false alarm error, then the specificity of the test (the percentage of true negative results) will be equal to $S_p = 93\,\%$. Let's assume that the test has ideal sensitivity (percentage of detection of true patients) $S_E = 100\,\%$.

It follows from expression (18) that at $\varepsilon = 0,07$ and $\omega = 3$ the boundary value of the prevalence is equal to $P_\omega(\varepsilon) = 0,0228$. Then, if the prevalence of the studied disease is $P(V_1) = 0,02$, i.e. $P(V_1) < P_\omega(\varepsilon)$, then a diagnostic test with operational characteristics $S_E = 100\,\%$ and $S_p = 93\,\%$ will be useless. In fact, with the expected 200 ill persons from the group of 10,000 people, all this people will not be detected without using the test. In this case, the a priori risk per person will be

$$R_0 = (L_{12} \cdot 200) / 10000 = (3 \cdot 200) / 10000 = 0,06. \tag{19}$$

If you use a diagnostic test with sensitivity $S_E = 100\,\%$, then all 200 patients will be identified. However, with specificity $S_p = 93\,\%$, out of 9800 healthy individuals, only 9114 people will be recognized as healthy, and the remaining 686 persons will be erroneously assigned to the group of patients V_1. Therefore, the posterior risk due to false positive results will be

$$R = (L_{21} \cdot 686) / 10000 = (1 \cdot 686) / 10000 = 0,0686 . \qquad (20)$$

Based on the comparison (19), (20), we conclude that $R > R_0$, i.e. a diagnostic test leads to an increase in average losses, which means that such test is useless.

If in the conditions of the example we change the value of the prevalence, assuming that $P(V_1) = 0,025$, then in this case $P(V_1) > P_{\omega}(\varepsilon) = 0,0228$. The expected number of patients in the group of 10,000 examined will be already 250 people who will not be identified without using the test. Therefore, in this case, the a priori risk will be

$$R_0 = (L_{12} \cdot 200) / 10000 = (3 \cdot 250) / 10000 = 0,075 . \qquad (21)$$

When using a diagnostic test with sensitivity $S_E = 100\,\%$, all 250 patients will be identified. The expected number of healthy people is 9750 people. With specificity $S_p = 93\,\%$, only 9068 truly healthy people will be assigned to the group V_2 of healthy people, and the remaining 682 people will be erroneously recognized as sick. Therefore, the posterior risk will be

$$R = (L_{21} \cdot 682) / 10000 = (1 \cdot 682) / 10000 = 0,0682 . \qquad (22)$$

Based on the comparison (21), (22), we conclude that now $R < R_0$, i.e. a diagnostic test is useful.

In a number of scientific publications, in particular in [13], decision-making algorithms are studied that satisfy the criterion of minimum false alarm probability $P(E_2)$. To fulfill this condition, instead of (4), (5), an alternative strategy for choosing solution areas $X_1 \subset X$ and $X_2 \subset X$ may be considered in the form:

$$P(E_2) \to \min , \qquad (23)$$

$$P(E_1) \le \delta , \qquad (24)$$

where δ is a given restriction on the probability E_1 of missed target error.

Note that screening tasks are characterized by low values of prevalence $P(V_1)$ [14]. Therefore, for relatively small values ω, the condition $P(V_1)(1+\omega) < 1$ will be satisfied, for which Consequence 1 is formulated.

At the same time, when diagnosing especially dangerous diseases, the ratio ω of losses should take on very large values, for example $\omega = 400$, when diagnosing oncological diseases [15]. In this case, even at low values of the prevalence $P(V_1)$, not the condition $P(V_1)(1+\omega) < 1$, but the condition $P(V_1)(1+\omega) > 1$ is already satisfied.

Therefore, to determine the value δ for which the diagnostic algorithm (1) which implements strategy (23), (24), wll be useful in the sense of (7), we use condition (13) and prove the statement.

Statement 3. *Let the condition $P(V_1)(1+\omega) > 1$ be satisfied. Then, for any arbitrarily small probability of false alarm error $P(E_2)$, algorithm (1) is obviously useless in the sense of (7) if the missed target error probability $P(E_1)$ satisfies the condition*

$$P(E_1) \ge \frac{1 - P(V_1)}{\omega P(V_1)} . \qquad (25)$$

ISSN 2786-6491

Proof. Since $P(V_1) \neq 0$ then condition (13), which guarantees the usefulness of the test for the region $P(V_1)(1+\omega) > 1$, can be represented in the equivalent form

$$P(E_1) < \frac{[1-P(V_1)[1-P(E_2)]}{\omega P(V_1)}. \tag{26}$$

From (26) it follows that algorithm (1) is useless if

$$P(E_1) \geq \frac{[1-P(V_1)[1-P(E_2)]}{\omega P(V_1)}. \tag{27}$$

Since $P(E_2) \geq 0$, then amplification (27) by substitution $P(E_2) = 0$ leads to an estimate of the admissible value of the probability of false alarm error $P(E_1)$ in the form of condition (25). Statement 3 is proved.

Based on expression (25), such consequence may be formulated.

Consequence 2. *In order for* $P(V_1)(1+\omega) > 1$ *algorithm (1), which implements the Neyman-Pearson strategy (23), (24), to be useful in terms of reducing average losses, the restriction* δ *on the right-hand side of (24) must satisfy the condition*

$$\delta \leq \frac{1-P(V_1)}{\omega P(V_1)} \tag{28}$$

Condition (28) after simple transformations allows us to estimate the admissible value of the prevalence

$$P_\omega(\delta) = \frac{1}{\delta\omega + 1}, \tag{29}$$

in which a diagnostic test remains useful. If the prevalence of the studied disease exceeds the value (29), i.e. $P(V_1) > P_\omega(\delta)$, then the Neumann-Pearson strategy (23), (24) turns out to be useless in terms of reducing the a priori risk. The condition $P(V_1) > P_\omega(\delta)$ is due to the fact that the value $P_\omega(\delta)$ decreases as the restriction δ increases (in contrast to the function $P_\omega(\varepsilon)$, which increases with the constraint increases ε).

Figure 2 shows plots of dependence $P_\omega(\delta)$ on δ, plotted in accordance with (29) for various values of ω.

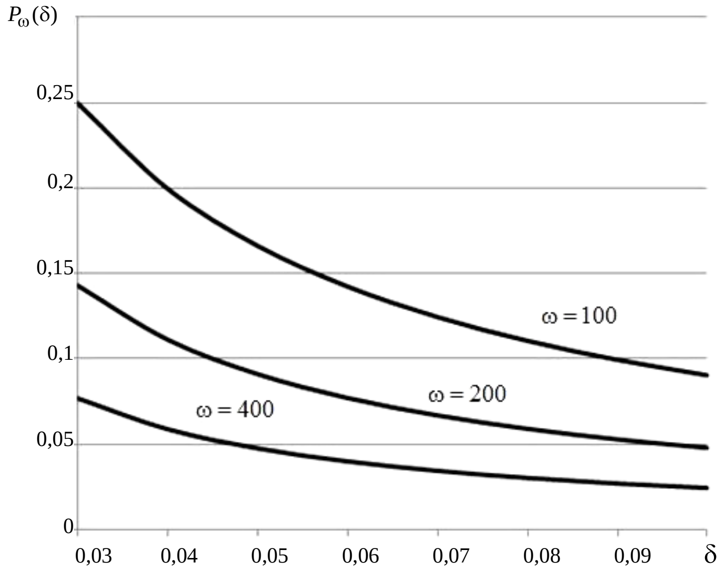

Fig. 2.

Example 2. To examine a group of 10 000 people, it is supposed to use a diagnostic test that implements the strategy (23), (24). Due to the high risk of the disease $\omega = 400.$ Given $\delta = 0,07,$ we will assume that $P(E_1) = 0,07.$ In this case, the sensitivity of the test is $S_E = 93\,\%$. We assume that the test has ideal specificity $S_P = 100\,\%$, i.e. everyone who is healthy is recognized as healthy.

It follows from expression (29) that at $\delta = 0,07$ and $\omega = 400$ the boundary value of the prevalence is equal to $P_\omega(\delta) = 0,0345.$ Let us show that if the prevalence of the studied disease is $P(V_1) = 0,04,$ i.e. $P(V_1) > P_\omega(\delta),$ then a diagnostic test with operational characteristics $S_P = 100\,\%$ and $S_E = 93\,\%$ will be useless.

Really in the group of 10,000 people, the expected number of sick people will be 400 person and healthy people will be 9600 persons. We assume that the value $\omega = 400$ is determined by the absolute values of losses $L_{12} = 400$ and $L_{21} = 1.$ It follows that in order to minimize the a priori risk, all persons must be recognized as sick (otherwise, the a priori losses will be higher). In this case

$$R_0 = (L_{21} \cdot 9600)/10000 = (1 \cdot 9600)/10000 = 0,96. \tag{30}$$

When using the diagnostic test that has specificity $S_E = 100\,\%$, all 9600 healthy people will be considered healthy. With sensitivity $S_E = 93\,\%$, 372 people from 400 patients will be recognized as sick, and the remaining 28 people will be erroneously classified as healthy. In this case, the average losses will be

$$R = (L_{12} \cdot 28)/10000 = (400 \cdot 28)/10000 = 1,12. \tag{31}$$

Based on the comparison (30), (31), we conclude that $R > R_0$, i.e. a diagnostic test leads to an increase in average losses and so the test is useless.

Let us change the value of the prevalence under the conditions of the example, assuming that $P(V_1) = 0,04$, i.e. $P(V_1) < P_\omega(\delta) = 0,0345.$ It is easy to verify that in this case

$$R_0 = (L_{21} \cdot 9700)/10000 = (1 \cdot 9700)/10000 = 0,97, \tag{32}$$

and

$$R = (L_{12} \cdot 21)/10000 = (400 \cdot 21)/10000 = 0,84. \tag{33}$$

Based on the comparison (32), (33), we conclude that now $R < R_0$, i.e. a diagnostic test is useful.

Features of the decision-making algorithm based on medical symptoms

Diagnosis of diseases is often based on the analysis of symptoms, the presence or absence of which makes it possible to make the following decision whether the patient belongs to the class V_1 or V_2. Methodological aspects of such diagnostic features are widely discussed in the scientific literature [16, 17]. Here we will be interested only in the question connected with the missed target and false alarms errors E_1 and E_2 that arise in the analysis of symptoms $x \in \{x^+, x^-\}$.

Consider the simplest symptom decision scheme:

$$\begin{aligned} &\text{person is sick}, \quad \text{if } x = x^+, \\ &\text{person is helthy, if } x = x^-, \end{aligned} \tag{34}$$

where x^+ , x^- are respectively the presence or absence of a symptom x in a particular patient.

It would seem that if, according to medical statistics, it is known that the symptom x is typical for most sick patients and uncharacteristic for most healthy people, then the rule (34) ensures high efficiency of diagnostic solutions, allowing small values of error probabilities E_1 and E_2. In fact, this is not always the case!

In the general case, information about the environment, including information about a symptom x, can be represented as a field

$$x = x(z, t) \tag{35}$$

at a point z in space at a time moment t. It follows from (35) that decision-making according to the rule (34) is reduced to determining the presence x^+ or absence x^- of a symptom in some body area x of a particular patient for the time points $t \in [T_1, T_2]$ of the observation interval.

From (35), taking into account (34), it follows that for a particular patient, a diagnostic solution requires checking the logical conditions:

$$\text{person is sick, if } \exists z, t : z \in Z,\ t \in [T_1, T_2],\ \ x(z, t) = x^+, \tag{36}$$

$$\text{person is helthy, if } \forall z, t : z \in Z,\ \ t \in [T_1, T_2],\ x(z, t) = x^-. \tag{37}$$

Let us consider the features of making such decisions on the example of the symptom «Characteristic skin rashes», which is an indicator of many diseases [18]. Suppose that, according to medical statistics, the probability of having this symptom in a typical sick patient is $P(x^+ \mid V_1) = 0,99$, and the probability of not having a symptom in healthy people is $P(x^- \mid V_2) = 0,99$.

According to (36), in order to make a decision in favor of the class V_1 (person is sick), it is sufficient to detect the symptom «Characteristic skin rashes» at some point in time $t_0 \in [T_1, T_2]$ on a local area $z_0 \in Z$ of the patient's body. We can assume that for a particular person the probability of a random event $x(z_0, t_0) = x^+$ is equal $P(x^+ \mid V_1) = 0,99$ if this person is sick. In this case, the probability of missed target is equal $P(E_1) = 0,01$.

At the same time, according to (37), the reliability of making the opposite decision in favor of the class V_2 (person is helthy) characterizes the probability of a complex event

$$A = \bigcup_{i,\, j} [x(z_i, t_j) = x^-], \tag{38}$$

consisting in the fact that at each moment of observation $t_i \in [T_1, T_2]$, the symptom x is not detected in all local areas z_i of the region Z. Assuming the independence of elementary events that form a complex event (38), the probability A can be estimated by the value

$$P(A) = \prod_{i,\, j} P[x(z_i, t_j) = x^-]. \tag{39}$$

From (39) it follows that with $P[x(z_i, t_j) = x^-] = 0,99$ the false alarm probability $P(E_1) > 0,01$.

Conclusion

The article shows that the diagnostic algorithm is useful in terms of reducing average losses if and only if the probabilities of missed target and false alarm errors satisfy conditions (12), (13). On the basis of these conditions, restrictions on the permissible limits of the probabilities of missed target and false alarms errors of the algorithms that implement the Neyman-Pearson strategies (4), (5) and (21), (22) are determined. Boundary values have been established for the prevalence of diseases for which a diagnostic algorithm with known operational characteristics remains useful. The reasons error are established due to which the decision-making algorithm based on a medical symptom can have a high probability of a false alarm.

Л.С. Файнзільберг

ПРО ЙМОВІРНОСТІ ПОМИЛОК ПРОПУСКУ ЦІЛІ ТА ХИБНОЇ ТРИВОГИ В МЕДИЧНИХ СИСТЕМАХ ПРИЙНЯТТЯ РІШЕНЬ

Файнзільберг Леонід Соломонович,

Information Technologies and Systems of the National Academy of Sciences of Ukraine and the Ministry of Education and Science of Ukraine, National Technical University of Ukraine «Igor Sikorsky Polytechnic Institute»

fainzilberg@gmail.com

Бінарні класифікатори широко використовують для вирішення низки прикладних задач, зокрема, задач медичної діагностики. Ефективність таких систем характеризують помилки пропущення цілі та хибної тривоги. Оскільки зі зменшенням ймовірності помилки пропуску цілі збільшується ймовірність помилки хибної тривоги та навпаки, налаштування систем передбачає певний компроміс між зазначеними помилками. Для побудови алгоритму діагностики в умовах апріорної невизначеності часто використовують стратегію Неймана-Пірсона, яка передбачає мінімізацію ймовірності помилки пропуску цілі при заданому обмеженні на ймовірність помилки хибної тривоги. Для дослідження питання формального вибору прийнятних обмежень сформульовані умови корисності діагностичного алгоритму. На основі цих умов визначено допустимі межі ймовірностей помилок пропуску цілі та хибної тривоги. Визначено граничні значення розповсюдженості захворювань, для яких діагностичний алгоритм з відомими операційними характеристиками залишається корисним з погляду зменшення середніх втрат. Наведено приклади, що ілюструють практичну цінність одержаних умов. Розглянуто особливості алгоритму прийняття рішень на основі медичних симптомів. Показано, якщо за даними медичної статистики деякий симптом характерний для більшості хворих пацієнтів і нехарактерний для більшості здорових людей, то звідси ще не випливає висока достовірність діагностичних рішень, заснованих на аналізі цього симптому для конкретної людини. На прикладі аналізу симптому «Характерні шкірні висипання» продемонстровано, що навіть якщо ймовірність відсутності симптому у здорових людей становить 0,99 ймовірність помилкової тривоги може бути нижчою, ніж 0,01. Це пояснюється тим, що для прийняття обґрунтованого рішення для конкретного пацієнта необхідно переконатися, що зазначений симптом відсутній протягом певного інтервалу спостереження.

Ключові слова: пропуск мети, помилкова тривога, корисність діагностичного алгоритму, стратегія Немана-Пірсона, превeconвеленс захворювання, медичний симптом.

REFERENCES

1. Maxim D., Niebo R., Utel M.J. Screening tests: a review with examples. *Inhalation Toxicology*. 2014. Vol. 26. N 13. P. 811–828. DOI : http://doi.org/10.3109/08958378.2014.955932.
2. Driggers R., Pollak E., Grimming R., Velazquez E., Short R., Gerald Holst G., Orges Furxhi O. Detection of small targets in the infrared: an infrared search and track tutorial. *Applied Optics*. 2021. Vol. 60. N 16. P. 4762–4777. DOI : https://doi.org/10.1364/AO.424767.
3. Hammoud D., Maamri R., Sahnoun Z. Machine learning in an agent: a generic model and intelligent agent based on inductive decision learning. *Journal of Artificial Intelligence*. 2011. Vol. 4. N 1. P. 29–44. DOI : https://doi.org/10.3923/jai.2011.29.44.
4. Pouramirarsalani A., Khalilian M., Nikravanshalmani A. Fraud detection in E–banking by using the hybrid feature selection and evolutionary algorithms. *International Journal of Computer Science and Network Security*. 2017. Vol. 17. N 8. P. 271–279. http://paper.ijcsns.org/07_book/201708/20170836.pdf.
5. Schwark J., Sandry J., MacDonald J. et al. False feedback increases detection of low-prevalence targets in visual search. *Atten Percept Psychophys*. 2012. Vol.74. P. 1583–1589. DOI : https://doi.org/10.3758/s13414-012-0354-4.
6. Lin Y-S., Heinke D., Humphreys G.W. Modeling visual search using three-parameter probability functions in a hierarchical Bayesian framework. *Attention, Perception and Psychophysics*. 2015. Vol. 77. P. 985–1010. DOI : https://doi.org/10.3758/s13414-014-0825-x.
7. Petrus Mikkola P., Todorovic M., Jarvi J., Rinke P., Kaski S. Projective preferential bayesian optimization. *Proceedings of the 37th International Conference on Machine Learning* (Vienna, Austria), PMLR 108. 2020. P. 1–10. https://www.researchgate.net/publication/339164741.
8. Scott C., Nowak R. A Neyman–Pearson approach to statistical learning. *IEEE Trans. on Inform. Theory. 200*. Vol. 51. N 11. P. 3806–3819. DOI : https://doi.org/10.1109/TIT.2005.856955.
9. Tong X., Feng Y., Jessicali J. Neyman-Pearson classification algorithms and NP receiver operating characteristics. *Science advances*. 2018. Vol. 4. N 2. P. 1–10. DOI : https://doi.org/10.1126/ sciadv.aao1659.
10. Cannon A., Howse J., Hush D., Scovel C. Learning with the Neyman-Pearson and min-max criteria. *Los Alamos National Laboratory. 2002* [Online]. Available: http://www.c3.lanl.gov/ kelly/ml/pubs/ 2002minmax/paper.pdf.
11. Schlesinger M.I., Hlavac V. ten lectures on statistical and structural pattern recognition. 2014, SPRINGER-SCIENCE+BUSINESS MEDIA, B.V., 522 p. DOI : https://doi.org/10.1007/ 978-94-017-3217-8.
12. Fainzilberg L.S. Conditions of utility of diagnostic tests from the point of view of the statistical theory of decision making. *Journal of Automation and Information Sciences*. 2003. Vol. 35. N 4. P. 63–73. DOI : https://doi.org/10.1615/JAutomatInfScien.v35.i4.30.
13. Wang J. Minimizing the false alarm rate in systems with transient abnormality. *Naval Research Logistics*. Vol. 63(4). P. 320–334. DOI : https://doi.org/10.1002/nav.21690.
14. Sabir A.A, Balarabe S, Sani A.A, Isezuo S.A, Bello K.S, Jimoh A.O. Prevalence of diabetes mellitus and its risk factors among the suburban population of Northwest Nigeria. *Sahel Medical Journal*. 2017. Vol. 20. N 4. P. 168–72. DOI : https://doi.org/10.4103/smj.smj_47_16.
15. Sridharraj K. Venkatesh S, Rao S.A., Hathiram N. Brain tumor recognition and categorization in MRI images utilizing optimal deep belief network. 2022 *IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, Ballari, India. 2022. P. 1–5. DOI : https://doi.org/10.1109/ICDCECE53908.2022.9793211.
16. Malterud K., Guassora A.D. Graungaard, A.H. Understanding medical symptoms: a conceptual review and analysis. *Theoretical Medicine and Bioethics*. 2015. Vol. 36. N 6. P. 411–424. DOI : https://doi.org/10.1007/s11017-015-9347-3.
17. Eriksen, T.E., Risør, M.B. What is called symptom? *Medicine health care and philosopy*. 2014. Vol. 1, . P. 89–102. DOI : https://doi.org/10.1007/s11019-013-9501-5.
18. Kang J.H. Febrile illness with skin rashes. *Infections and Chemotherapy*. 2015. Vol. 47. N 3. P. 155–166. DOI : https://doi.org/10.3947/ic.2015.47.3.155.

Submitted 30.01.2023