STATISTICAL DIAGNOSTICS OF METASTATIC INVOLVEMENT OF REGIONAL LYMPH NODES

M.S. Abramovich, A.Yu. Kharin, N.I. Belyakova
Belarussian State University
N.N. Alexandrov Research Institute of Oncology and Medical Radiology
Minsk, BELARUS
e-mail: Abramovichms@bsu.by

Abstract

The results diagnostics of metastatic involment of regional lymph nodes using robust discriminant analysis and support vector machines are considered. The method of statistical classification with indicating patients that require more detailed diagnostics is proposed and analysed.

1 Introduction

Melanoma of the skin is one of most aggressive human malignant tumors. Every 10-20 years the morbidity rate among European population is increasing rapidly (twice at list with annual increase from 3% to 7%) [3]. In Europe melanoma takes 17th place in men and 8th place in women in rating of most frequently diagnosed cancer types [1].

Hystorically the melanoma is considered as a disease with variable and often unfavourable course. Unlike many other human malignant tumors, the melanoma strikes predominantly young people, the melanoma has high metastatic potential and resistance towards chemo- and radiotherapy. Although melanoma forms only 4% of malignant skin cancer cases, it is responsible for 79% of deaths in this group of patients.

Nevertheless local melanoma is not a disease with synonymous poor prognosis. In case of early diagnosis complete cure may be achievedd in 95% of cases.

During last decades the investigators are attempting to determine the factors, responsible for the course and prognosis of the disease. The significance of the factors is taken into consideration in determination of the stage of the disease, the importance of other factors is rejected during time.

Nowadays the problem of lymph node dissection combined with excision of primary focus of tumor remains disputable. The selection of group of patients with high risk of metastases in lymph nodes is very actual for the decrease of rate of lymph node dissection (often groundless). Although this group of patients is a very limited part of population, in a majority of cases right diagnostic decision is out of difficulties and requires working out special diagnostic programms.

In clinical oncology parametric discriminant analysis [2] and support vector machines [4] are widely used for malignant neoplasms diagnostic in the case of presence of a priori information on a training sample.

2 Results of diagnostics using discriminant analysis and support vector machines

An attempt of working out indications to prophilactic lymphadenectomy in patients with clinically intact peripheral lymph nodes. Retrospective investigation of frequency of lymph node metastases depending on clinico-anatomical factors, characterizing primary tumor and organism of patient is carried out.

From 1970 to 2002, 982 melanoma patients treated at the State Institution "N.N. Alexandrov Research Institute of Oncology and Medical Radiology" (Minsk, Belarus). Class Ω_1 consisted of 205 patients with histological evidence of metastases in regional lymph nodes. Class Ω_2 included 777 patients with no regional lymph node metastases on histological examination.

We attempted to establish indications for lymphadenectomy in patients whose regional lymph nodes were not changed. A retrospective study was carried out to evaluate the rate of metastatic involvement of regional lymph nodes with regard to the followingn clinico-anatomical features characterizing the primary tumor and the patient's body: age (AG), background (B), disease duration before the start of the treatment (DDST), tumor growth (TG), tumor square (TS), degree of tumor pigmentation (DTP), Clark invasion level (CIV), Breslow tumor thickness (BTH), ulceration (U), tumor-infiltrating lymphocytes (TIL), histological subtype (HS), growth phase (GP), satellites (ST), vascular invasion (VI), anatomic distribution of melanoma (ADM).

Quadratic discriminant functions were used to include infomation on difference in covariances between features of the classes; they increase the dianostics accuracy in comparison with the linear decision rules [2].

Clinico-anatomical features DDST, BTH and TS have an abnormal observations (outliers). As it was mentioned in [2], the presence of such observations (which have to be used to construct diagnosis rules) causes a loss in the accuracy of the decision rules. The reason for that is the calculation of discriminant coefficients employing mean values of the features and their covariance matrixes which non-robust in the presence of the outliers. Therefore robust decision rules based on the robust Huber M-estimators [2] of the mean vectors and covariance matrices were used. The accuracy of the decision rules was werified by reclassification of the training samples.

From analyzed features two informative sets are formed: 1) AG, B, DDST, TS, CIV, BTH; 2) AG, DDST, TS, DTP, CIV, BTH.

For the first informative set quadratic decision rule the rate of true classification of patients from class Ω_1 is 65.22%, the rate of true classification of patients from class Ω_2 is 63.88%, the rate of true classification for the both classes is 64.14%.

For thes second informative set the rate of true classification of patients from class Ω_1 is 62.07%, the rate of true classification of patients from class Ω_2 is 62.70%, the rate of true classification for the both classes is 62.58%.

It should be noted that the usage of the classic estimators for the mean and the covariance matrix in decision rules results in low diagnostic performance: the performance of the decision rule for the first set of features is lower on 4.5%, the performance of the decision rule for the second set of features is lower on 3.6%.

Standard procedures of the discriminant analysis

Thus, satisfactory values of error probabilities for the discriminant function can not be reached.

3 Classification with detection of observations that require more detailed diagnostics

Because some extra diagnostics and prophylactic treatment is available, we construct a decision rule with 3 admissible decisions about an observation $x = x(\omega) \in \mathbf{R}^N$ on the patient $\omega \in \Omega = \Omega_1 \cup \Omega_2$. The decisions $d = 1$ and $d = 2$ mean that $\omega \in \Omega_1$ and $\omega \in \Omega_2$ respectively. The decision $d = 0$ corresponds to the situation, where ω needs more detailed diagnostics and prophylactic treatment.

Denote by $\lambda(x)$ the discriminant function constructed for the problem of distinguishing between Ω_1 and Ω_2. Introduce the decision rule:

$$d = d(x) = 2 \cdot \mathbf{1}_{(B,+\infty)}(\lambda(x)) + \mathbf{1}_{(-\infty,A)}(\lambda(x)), \tag{1}$$

where x is the observed value of the vector of features, $A < 0$, $B > 0$ are parameters. For the choice of A, B, the following two criteria are proposed.

3.1 Risk-oriented approach

Introduce the notation: $\pi_1 \in (0,1)$ is the prior probability of the random event $\{\omega \in \Omega_1\}$; $P_{ij}(A,B)$ is the probability of the decision $d = j$ provided the observation comes from Ω_i; $w_{ij} \geq 0$ is the cost of the correspondent decision. The values A, B are the solutions of the risk (expected losses) minimization problem:

$$\begin{gathered} R(A,B) = \pi_1 \left(w_{12}P_{12}(A,B) + w_{10}P_{10}(A,B)\right) + \\ +(1-\pi_1)\left(w_{21}P_{21}(A,B) + w_{20}P_{20}(A,B)\right) \to \min_{A<0,B>0}. \end{gathered} \tag{2}$$

Consider the case where observations from the class Ω_i have the Gaussian probability distribution:

$$\mathcal{L}\{x(\omega)\} = \mathcal{N}_N(\mu_i, \Sigma),\ \omega \in \Omega_i, i = 1,2. \tag{3}$$

Denote: $\Phi(\cdot)$ is the distribution function of $\mathcal{N}_1(0,1)$;

$$\Delta = \sqrt{(\mu_2-\mu_1)^T\Sigma^{-1}(\mu_2-\mu_1)},\ b = \Sigma^{-1}(\mu_2-\mu_1),$$

$$H = \frac{1}{2}\left(\mu_2^T\Sigma^{-1}\mu_2 - \mu_1^T\Sigma^{-1}\mu_1\right),\ m_1 = b^T\mu_1 - H.$$

Theorem 1 *If the model (3) is valid, then the minimization problem (2) is equivalent to the pair of separate minimization problems:*

$$R(A) = (1-\pi_1)(w_{21}-w_{20})\Phi\left(\frac{A-m_1}{\Delta} - \Delta\right) - \pi_1 w_{10}\Phi\left(\frac{A-m_1}{\Delta}\right) \to \min_{A<0},$$

$$R(B) = \pi_1(w_{10}-w_{12})\Phi\left(\frac{B-m_1}{\Delta}\right) + (1-\pi_1)w_{20}\Phi\left(\frac{B-m_1}{\Delta} - \Delta\right) \to \min_{B>0}.$$

3.2 Frequency-oriented approach

Let the maximal possible values α, β for the error probabilities be given:

$$P_{12}(A, B) = \alpha, \ P_{21}(A, B) = \beta. \tag{4}$$

Solving (4) w.r.t. A, B, we get the values to be used in the decision rule (1).

Denote by $\Phi^{-1}(\gamma)$ the quantile of the level $\gamma \in (0, 1)$ for the standard normal probability distribution $\mathcal{N}_1(0, 1)$.

Theorem 2 *If the model (3) holds, then the solution of (4) is*

$$A = m_1 + \Delta \cdot \Phi^{-1}(\beta) + \Delta^2, \ B = m_1 + \Delta \cdot \Phi^{-1}(1 - \alpha).$$

The numerical results of using the decision rule (1) and Theorems 1, 2 for the decribed data set will be given in the talk.

References

[1] Ferlay J., Bray F., Pisani P., Parkin DM. (2001). CLOBOCAN 2000: cancer incidence, mortality and prevalence worldwide, version 1.0. *IARC cancerbase No.5.* Lyon, IARC Press.

[2] Kharin Yu. S.(1996). *Robustness in statistical pattern recognition.* Dordrecht; Boston; London: Kluwer Academic Publishers.

[3] Osterlind A. (1992). Epidemiology on malignant melanoma in Europe. *Acta Oncology*, N **31(8)**, pp. 903-908. John Wiley and sons.

[4] Vapnik V. (1998). *Statistical learning theory.* John Wiley and sons.