

Expected Bayes Error Rate in Supervised Classification of Spatial Gaussian Data

Kęstutis DUČINSKAS, Lijana STABINGIENĖ

Department of Statistics, Klaipėda University

H. Manto 84, LT-92294 Klaipėda, Lithuania

e-mail: kestutis.ducinkas@ku.lt, lijana.stabingiene@gmail.com

Received: March 2010; accepted: December 2010

Abstract. In the usual statistical approach of spatial classification, it is assumed that the feature observations are independent conditionally on class labels (conditional independence). Discarding this popular assumption, we consider the problem of statistical classification by using multivariate stationary Gaussian Random Field (GRF) for modeling the conditional distribution given class labels of feature observations. The classes are specified by multivariate regression model for means and by common factorized covariance function. In the two-class case and for the class labels modeled by Random Field (RF) based on 0–1 divergence, the formula of the Expected Bayes Error Rate (EBER) is derived. The effect of training sample size on the EBER and the influence of statistical parameters to the values of EBER are numerically evaluated in the case when the spatial framework of data is the subset of the 2-dimensional rectangular lattice with unit spacing.

Keywords: Bayes discriminant function, Gaussian random fields, spatial correlation, divergence.

1. Introduction, Main Concepts

Spatial supervised classification is a problem of classifying locations (sites, pixels) into several categories by learning the feature observations and the adjacency relationships with training sample. The classification of pixel into one of the classes is fundamental problem in image pattern analysis (see, e.g., Mardia, 1988). Switzer (1980) was the first to treat classification of spatial data. Mardia (1984) extended this research by including spatial discrimination methods in forming the classification maps. The application of spatial contextual (or supervised) classification methods in geospatial data mining is considered by Shekhar *et al.* (2004). It should be noted, that widely applicable methods of exemplar-based image completion (see, e.g., Wu *et al.*, 2010) are also closely tied with the methods of spatial discrimination.

It is usually assumed that feature observations are conditionally independent given class labels (conditional independence) (see, e.g., Cressie, Section 7.4). This approach with normally distributed features and the labels following the Markov RF model is widely used for remote sensing image classification (see, e.g., Nishii, 2003).

Computer-intensive methods, including simulated annealing and MCMC methods, can be used for the solving of spatial classification problems, but the implementation

is often difficult because of computational complexity. So the derivation closed form expressions of error rates as performance measures of the classification procedures are of great importance.

The exact error rates due to Bayes classification rule are derived by Nishii and Eguchi (2006), under the assumptions of conditional independence. Problems of spatial classification of Gaussian feature observations by discarding the assumption of conditional independence were considered for fixed (nonrandom) class labels in training sample (Dučinskas, 2009) and for class labels following Markov RF model (Stabingienė and Dučinskas, 2009).

In the present paper we performed the generalization of the aforementioned paper results to the multivariate feature case and different approach to modeling of class labels. The stationary multivariate Gaussian Random Fields (GRF) model for features and logistic type discrete RF model based on 0–1 divergence are considered. Conditional Bayes error rate (CBER) and EBER are derived for the two-class case. Numerical analysis of derived expected error rate is carried out in the case of spatial framework of data being the subset of 2-dimensional rectangular lattice with unit spacing and anisotropic exponential spatial correlation model for feature observations. Free software system *R* is used for calculations.

In the following we will give a brief descriptions of feature and label models.

Suppose that spatial data consist of observed values of feature variable which is modeled by p -variate RF

$$\{Z(s): s \in D \subset \mathbb{R}^2\}.$$

Each location in area D is assumed to belong to one of two classes Ω_1, Ω_2 . A class label or simply the label for location $s \in D$ is denoted by $Y(s)$, and is treated as random variable over a label set $L = \{1, 2\}$.

Let the model of feature observation $Z(s)$ in class Ω_l (i.e., $Y(s) = l$) be

$$Z(s) = B_l'x(s) + \varepsilon(s), \quad (1)$$

where $x(s)$ is a $q \times 1$ vector of non random regressors and B_l is a $q \times p$ matrix of parameters, $l = 1, 2$. It is required that $B_1 \neq B_2$. The error term in (1) is generated by p -variate zero-mean stationary GRF $\{\varepsilon(s): s \in D\}$ with covariance function defined by the following model for all $s, u \in D$

$$\text{cov}\{\varepsilon(s), \varepsilon(u)\} = r(s - u)\Sigma, \quad (2)$$

where $r(s - u)$ is the spatial correlation function and Σ is feature variance – covariance matrix.

Set $Y = (Y(s_1), \dots, Y(s_n))'$ and $Z = (Z(s_1), \dots, Z(s_n))'$ and call them training label vector and training feature matrix, respectively. Thus, the $n \times (p + 1)$ matrix $T = (Y, Z)$ constitutes the training sample.

Denote by $S_n = \{s_i \in D; i = 1, \dots, n\}$ the set of locations where training sample T is taken, and call it the Set of Training Locations (STL). It is also referred to as spatial framework for training sample (see Shekhar *et al.*, 2002).

Suppose that the realizations of random training variables $Y = y$ and $Z = z$ correspond to the realization of training sample $T = t$. Then the distribution of training feature matrix Z for given $Y = y$ is matrix-variate normal distribution, i.e.,

$$Z | Y = y \sim N_{n \times p}(X_y B, R \otimes \Sigma), \tag{3}$$

where X_y is the $n \times 2q$ design matrix, $B' = (B'_1, B'_2)$ is the $p \times 2q$ matrix of mean parameters.

Here R denotes the spatial correlation matrix for feature observations at S_n with elements equal to the corresponding values of spatial correlation function r specified in (2).

The $n \times 2q$ design matrix X_y is formed in the following way: the first q columns contain regressors for feature observations from Ω_1 , and the second q columns contain regressors for feature observations from Ω_2 . In the following we specify the classification problem.

2. Bayes Discriminant Function and Associated Error Rates

In the present paper, we consider the problem of the classification of the feature observation $Z_0 = Z(s_0), s_0 \in D$ with an unobserved class label in the case of given training sample. The label for location s_0 is denoted by Y_0 .

Denote by r_0 the vector of spatial correlations between Z_0 and Z and set

$$Z^+ = \begin{pmatrix} Z \\ Z_0 \end{pmatrix}, \quad R^+ = \begin{pmatrix} R & r_0 \\ r'_0 & 1 \end{pmatrix}, \quad x_0 = x(s_0), \quad \alpha_0 = R^{-1}r_0.$$

It follows from (1)–(3), that for $l = 1, 2$,

$$Z^+ | Y = y, \quad Y_0 = l \sim V_{(n+1) \times p}(X_y^l B, R^+ \otimes \Sigma), \tag{4}$$

where

$$X_y^l = \begin{pmatrix} X_y \\ x_0^l \end{pmatrix}, \quad x_0^l = (\delta_{1l}I_q \otimes \delta_{2l}I_q)x_0.$$

Here δ_{ij} is the Kronecker delta symbol, and I_q is the identity matrix of order q . Then from (4) it follows that the conditional distribution of Z_0 given $T = t$ is Gaussian, i.e.,

$$Z_0 | T = t, Y_0 = l \sim N_p(\mu_{lt}^0, \Sigma_{ot}). \tag{5}$$

Here the conditional means μ_{lt}^0 are

$$\mu_{lt}^0 = E(Z_0 | T = t; Y_0 = l) = B'_l x_0 + (z - X_y B)' \alpha_0, \quad l = 1, 2. \tag{6}$$

Conditional covariance matrix Σ_{0t} is specified as follows

$$\Sigma_{0t} = \text{Var}(Z_0 | T = t; Y_0 = l) = k\Sigma, \quad (7)$$

with $k = 1 - r'_0\alpha_0$.

The marginal squared Mahalanobis distance between populations for feature observation taken at location $s = s_0$ is

$$\Delta_0^2 = (\mu_1^0 - \mu_2^0)' \Sigma^{-1} (\mu_1^0 - \mu_2^0), \quad (8)$$

where $\mu_l^0 = B_l'x_0$, $l = 1, 2$.

The squared Mahalanobis distance between conditional distributions of Z_0 for given $T = t$ is specified by

$$\Delta_{0n}^2 = (\mu_{1t}^0 - \mu_{2t}^0)' \Sigma_{0t}^{-1} (\mu_{1t}^0 - \mu_{2t}^0). \quad (9)$$

Then using (6), (7) in (8), (9) yields

$$\Delta_{0n}^2 = \Delta_0^2/k.$$

It is obvious, that Δ_{0n} depends on training sample only through S_n .

Throughout the present paper we suppose that STL S_n is fixed, but the labels are distributed randomly on it.

So S_n is partitioned into union of two disjoint subsets, i.e.,

$$S_n = S^{(1)} \cup S^{(2)},$$

where $S^{(l)}$ is the random subset of S_n that contains N_l locations with labels equal l , $l = 1, 2$. Since $N_1 + N_2 = n$, it is sufficient to consider only N_1 distribution.

Denote the distribution of discrete random variable N_1 by

$$\{\pi_j = P(N_1 = j), j = 0, 1, \dots, n\}. \quad (10)$$

These probabilities are sometimes called prior probabilities for labels.

We call $\xi(y) = \{S^{(1)}, S^{(2)}\}$ the Spatial Labels Design (SLD) corresponding to the training labels vector realization $Y = y$. It is obvious, that there is in one-to-one correspondence between $\xi(Y)$ and Y . Suppose that if $Y = y$, then $N_i = n_i$, $i = 1, 2$, where $n_1 + n_2 = n$.

Let $J(l, k)$ be a nonnegative divergence between two classes Ω_l and Ω_k , for $k, l = 1, 2$, satisfying $J(l, l) = 0$.

Denote the conditional distribution of Y_0 given $Y = y$ by

$$\pi_l(y) = P(Y_0 = l | Y = y), \quad l = 1, 2.$$

In the present paper, we extend the approach used by Nishii and Eguchi (2006) for modeling the label distribution on $S_n \cup s_0$.

We assume that the distribution of label Y_0 conditionally by on $Y = y$ is specified by labels in S_n (not only by labels in some neighborhood of s_0) and by divergence.

With an insignificant loss of generality, we will consider the case when the conditional distribution of label Y_0 does not directly depend on S_n and s_0 .

Let $\Delta(l)$ denote the average of divergences between location s_0 with label l and labels of locations in S_n as $\Delta(l) = \sum_{k=1}^2 n_l J(l, k)/n$.

ASSUMPTION 1. The conditional distribution of Y_0 given $Y = y$ (with $N_1 = n_1$) is specified by the following equation

$$\pi_l(y) = \exp \left\{ -\rho \Delta(l) \right\} / \sum_{k=1}^2 \exp \left\{ -\rho \Delta(k) \right\}, \quad l = 1, 2, \tag{11}$$

where ρ is a non-negative constant called a clustering parameter.

Note that ρ gives the degree of spatial dependency of the RF. If $\rho = 0$, then the classes are spatially independent.

Under the assumption, that the classes are completely specified the Bayes Discriminant Function (BDF; Fukunaga, 1990) minimizing the probability of misclassification, is formed by the logarithm of ratio of conditional densities described above. We shall call that situation the case of complete parametric certainty.

Then BDF for classification of Z_0 given $T = t$ is

$$W_t(Z_0) = \left(Z_0 - \frac{1}{2}(\mu_{1t}^0 + \mu_{2t}^0) \right)' \Sigma_{0t}^{-1} (\mu_{1t}^0 - \mu_{2t}^0) + \gamma(y),$$

where $\gamma(y) = \ln(\pi_1(y)/\pi_2(y))$.

DEFINITION 1. The Conditional Bayes Error Rate (CBER) is defined as the probability, conditional on $T = t$, that random observation Z_0 is misclassified by BDF $W_t(Z_0)$ and is denoted by $P_0(t)$.

Let $\Phi(\cdot)$ be the standard normal distribution function.

Lemma 1. Suppose that conditional distribution of Z_0 is specified in (5)–(7) and conditional distribution of Y_0 satisfies Assumption 1. Then conditional Bayes error rate of classifying Z_0 by BDF $W_t(Z_0)$ is

$$P_0(t) = \sum_{l=1}^2 \pi_l(y) \Phi(Q_l(t)), \tag{12}$$

where $Q_l(t) = -\Delta_{0n}/2 + (-1)^l \gamma(y)/\Delta_{0n}$.

Proof of Lemma 1 follows directly from Definition 1 and from properties of multivariate Gaussian distribution.

DEFINITION 2. The Expected Bayes Error Rate (EBER) for $W_t(Z_0)$ is defined as $P_{0n} = E_T(P_0(T))$, where E_T denotes the expectation with respect to the distribution of T .

Without essential loss of generality, but for greater interpretability, we restrict our considerations to the special type of quasi-distance.

ASSUMPTION 2. Quasi-distance $J(l, k)$ is the 0–1 distance defined by $J(l, k) = 1 - \delta_{lk}$, where δ_{lk} stands for Kronecker's delta.

Note that spatial models with assumed type of divergence is frequently used in image segmentation (Besag, 1986).

Then applying the Assumption 2 to the (11), we have

$$\begin{aligned}\pi_1(y) &= 1/(1 + \exp\{-\rho(n_1 - n_2)/n\}), \\ \pi_2(y) &= 1/(1 + \exp\{\rho(n_1 - n_2)/n\}).\end{aligned}\quad (13)$$

Note that for fixed n , above probabilities depend on y only through n_1 . So we can introduce the new notations

$$\pi_l^*(n_1) = \pi_l(y), \quad \gamma^*(n_1) = \gamma(y).\quad (14)$$

Since the random variable N_1 is a function of Y , thus (10) and (14) imply that the joint distribution $\{P(Y = y, Y_0 = l)\} = \{P(Y_0 = l | Y = y)P(Y = y)\}$ have the following form

$$\{P(N_1 = n_1, Y_0 = l) = \pi_l^*(n_1) \cdot \pi_{n_1}, \quad n_1 = 0, \dots, n; \quad l = 1, 2\}.\quad (15)$$

Lemma 2. Under the conditions of Lemma 1 and Assumption 2 and for given prior probabilities $\{\pi_j\}$, the EBER for $W_t(Z_0)$ is

$$\begin{aligned}P_{0n} &= \sum_{j=0}^n \sum_{l=1}^2 \pi_j \Phi(-\Delta_{0n}/2 + (-1)^l \rho(2j/n - 1)/\Delta_{0n}) \\ &\quad / (1 + \exp\{(-1)^l \rho(2j/n - 1)\}).\end{aligned}$$

Proof. Applying Assumption 2, formula (15) and notations (14) to (12) we have the following expression for the CBER

$$P_0(t) = \sum_{l=1}^2 \pi_l^*(n_1) \Phi(-\Delta_{0n}/2 + (-1)^l \gamma^*(n_1)/\Delta_{0n}).\quad (16)$$

From (15) we see that CBER depends on $T = t$ only through $N_1 = n_1$. So we can replace the averaging of $P_0(T)$ with respect to the distribution of T by the averaging it with respect to the distribution of N_1 specified in (10). Then the assertion of Lemma 2 is obtained by using (13)–(15) in (16).

The derived closed form expression can be effectively used as performance measure for BDF and as optimality criterion for spatial sampling design.

3. Numerical Examples and Conclusions

Here we shall evaluate the effect of training sample size on the EBER and shall analyze numerically the dependence of EBER on some statistical parameters of classes.

For the numerical illustration we consider the univariate case of model (1)–(4) with constant means and isotropic exponential covariance function given by

$$C(h) = \sigma^2 \exp\{-|h|/\alpha\},$$

where σ^2 is a variance and α is a range parameter.

Suppose that D is 2-dimensional rectangular lattice with unit spacing and $S_0 = (0, 0)$. Without essential loss of generality, we consider the case with $n = 2M$, where M is a fixed natural number. Then, EBER P_{0n} is given by

$$P_{0n} = \sum_{j=0}^{2M} \sum_{l=1}^2 \pi_j \Phi(-\Delta_{0n}/2 + (-1)^l \rho(j/M - 1)/\Delta_{0n}) / (1 + \exp\{(-1)^l \rho(j/M - 1)\}) \tag{17}$$

EXAMPLE 1. In the first example, we consider how the training sample size affects the values of EBER. The effect of the sample size to EBER is explored using first-order ($M = 2$), second-order ($M = 4$) and third-order neighborhoods ($M = 6$) to s_0 . For greater interpretability we set $\pi_M = 1$. In this case the number of locations in S_n with label 1 is equal to the number of locations in S_n that are labeled 2 with probability one. Then EBER is expressed by

$$P_{0n} = \Phi(-\Delta_{0n}/2).$$

Note that in considered case, EBER does not depend on the clustering parameter ρ .

The cases of the SLD satisfying the conditions of the example1 are illustrated in Fig. 1.

Denote by $BE(I)$ the EBER for the STL forming the l -st order neighborhood of s_0 , for $l = 1, 2, 3$. The comparison of two STL with different sizes is done by the efficiency index defined by the ratio $E(ij) = BE(i)/BE(j)$.

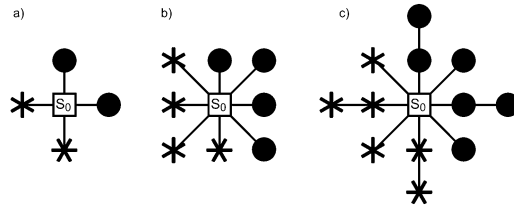


Fig. 1. Different spatial labels designs with points of $S^{(1)}$ and $S^{(2)}$ signed as \bullet and $*$, respectively. Cases (a), (b) and (c) indicate first-order, second-order and third-order neighborhoods of s_0 , respectively.

Table 1

Values of EBER and values of efficiency indexes for three neighborhoods and various values of range parameter α

	α					
	0.5	1	1.5	2	2.5	3
$BE(1)$	0.30260	0.27007	0.23612	0.20707	0.18255	0.16170
$BE(2)$	0.30237	0.26946	0.23562	0.20671	0.18229	0.16151
$BE(3)$	0.30236	0.26930	0.23512	0.20588	0.18120	0.16021
$E(21)$	0.99924	0.99774	0.99788	0.99826	0.99858	0.99882
$E(13)$	0.99921	0.99715	0.99576	0.99425	0.99260	0.99079
$E(23)$	0.99997	0.99941	0.99788	0.99599	0.99402	0.99195

The values of EBER specified in (17) and the values of efficiency indexes are calculated for different values of range parameter α but for fixed $\Delta_0 = 1$. They are presented in Table 1.

Figures in Table 1 confirm quite logical conclusion that EBER decreases with the increasing of sample size. Analyzing the rows of Table 1 with values of efficiency indexes, we can conclude that decreasing rate of EBER is higher for larger values of range parameter α .

EXAMPLE 2. In the second example we explore the influence of statistical parameters to the values of EBER. Here we restrict our attention on the STL forming the second-order neighborhood of s_0 (i.e., $n = 8$).

Note that the second-order neighborhood to $s_0 = (0, 0)$ is the set

$$S_8 = \{(0, 1), (1, 1), (1, 0), (1, -1), (0, -1), (-1, -1), (-1, 0), (-1, 1)\}.$$

Assume that prior probabilities are

$$\pi_2 = \pi_6 = 0.05, \quad \pi_3 = \pi_5 = 0.15, \quad \pi_4 = 0.6 \quad \text{and} \quad \pi_j = 0$$

for $j = 0, 1, 7, 8$.

Then only SLD corresponding to $N_1 = 2, 3, \dots, 6$ are admissible.

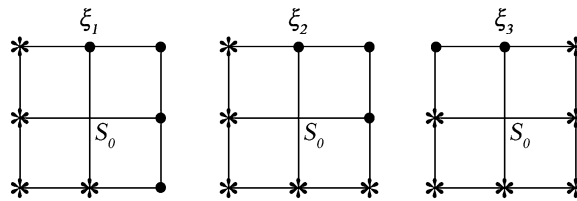


Fig. 2. Three different SLD with $S^{(1)}$ and $S^{(2)}$ points \bullet and $*$, signed as respectively.

Table 2
Values of P_{0n} for $\Delta_0 = 0.2$ and various α and ρ

ρ	α					
	0.5	1	1.5	2	2.5	3
0	0.45877	0.45110	0.44271	0.43503	0.42805	0.42166
0.4	0.44698	0.44095	0.43394	0.42725	0.42100	0.41518
0.8	0.42014	0.41655	0.41191	0.40712	0.40240	0.39783
1.2	0.38851	0.38656	0.38372	0.38055	0.37725	0.37392
1.6	0.35633	0.35537	0.35374	0.35173	0.34950	0.34715
2	0.32529	0.32490	0.32405	0.32285	0.32140	0.31979

The following three SLD

$$\begin{aligned} \xi_1 &= \{S^{(1)} = \{(0, 1), (1, 1), (1, 0), (1, -1)\}, \\ &\quad S^{(2)} = \{(0, -1), (-1, -1), (-1, 0), (-1, 1)\}\}, \\ \xi_2 &= \{S^{(1)} = \{(0, 1), (1, 1), (1, 0)\}, \\ &\quad S^{(2)} = \{(1, -1), (0, -1), (-1, -1), (-1, 0), (-1, 1)\}\}, \\ \xi_3 &= \{S^{(1)} = \{(-1, 1), (0, 1)\}, \\ &\quad S^{(2)} = \{(1, 1), (1, 0), (1, -1), (0, -1), (-1, -1), (-1, 0)\}\}, \end{aligned}$$

can be considered as examples of admissible SLD, because SLD ξ_1 corresponds the situation $N_1 = 4$, SLD ξ_2 corresponds the situation $N_1 = 3$, and SLD ξ_3 corresponds the situation $N_1 = 2$. These SLD are illustrated in Fig. 2.

For the conditions described in the beginning of Example 2, the values of EBER for various clustering parameters and range parameters are calculated. Figures for the case with $\Delta_0 = 0.2$ are presented in Table 2 and figures for the case with $\Delta_0 = 1$ are presented in Table 3.

Table 2 shows that exact error rate for weakly separated classes ($\Delta_0 = 0.2$) is monotonically decreasing in α for fixed values of ρ . Also we can deduce that observations of the features with “stronger” spatial dependence can be classified more correctly.

The similar trends in dependence of the EBER on range parameter α and clustering parameter ρ for strongly separated classes ($\Delta_0 = 1.0$) can be seen in Table 3.

Table 3
 Values of P_{0n} for $\Delta_0 = 1$ and various α and ρ

ρ	α					
	0.5	1	1.5	2	2.5	3
0	0.30237	0.26946	0.23562	0.20671	0.18229	0.16151
0.4	0.30178	0.26900	0.23525	0.20641	0.18204	0.16129
0.8	0.30005	0.26761	0.23415	0.20551	0.18129	0.16065
1.2	0.29728	0.26539	0.23237	0.20405	0.18006	0.15961
1.6	0.29362	0.26242	0.22999	0.20209	0.17842	0.15822
2	0.28923	0.25884	0.22711	0.19971	0.17642	0.15651

Numerical analysis, performed for small training samples, shows that the greater dependency of class labels and stronger spatial correlation between feature observations ensures the smaller EBER. This conclusion can be verified directly by differentiation EBER given in formula (17). So we can expect the similar dependencies for other spatial correlation models for features and more complicated label distribution.

References

- Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of Royal Statistical Society B*, 48(3), 259–302.
- Cressie, N. (1993). *Statistics for Spatial Data*. Revised edition. Wiley, New York.
- Dučinskas, K. (2009). Approximation of the expected error rate in classification of the Gaussian random field observations. *Statistics and Probability Letters*, 79, 138–144.
- Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*. Second edn., Academic Press, New York.
- Mardia, K. V. (1984). Spatial discrimination and classification maps. *Comm. Statist. Theory Methods*, 13(18), 2181–2197.
- Mardia, K. V. (1988). Multidimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis*, 24, 265–284.
- Nishii, R. (2003). A Markov random field-based approach to decision level fusion for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 41(10), 2316–2319.
- Nishii, R., Eguchi, S. (2006). Image classification based on Markov random field model with Jeffreys divergence. *Journal of Multivariate Analysis*, 97(9), 1997–2008.
- Shekhar, S., Schrater, P.R., Vatsavai, R.R., Wu, W., Chawla, S. (2002). Spatial contextual classification and prediction models for mining geospatial data. *IEEE Transactions on Multimedia*, 4(2), 174–188.
- Stabingienė L., Dučinskas, K. (2009). Exact error rates of the supervised classification based on Markov random fields. In: *ASMDA Proceedings*, pp. 120–123.
- Switzer, P. (1980). Extensions of linear discriminant analysis for statistical classification of remotely sensed satellite imagery. *Mathematical Geology*, 12(4), 367–376.
- Wu, J., Ruan, Q., An, G. (2010). Exemplar-based image completion model employing PDE corrections. *Informatica*, 21(2), 259–276.

K. Dučinskas graduated from the Vilnius University in 1976 in applied mathematics, where received doctor degree in 1983. He is a head of Statistics Department and a professor of Klaipėda University. Present research interests include discriminant analysis of spatially correlated data, statistical pattern recognition, geostatistics and Markov random fields, spatial data mining.

L. Stabingienė graduated from the Klaipėda University in 2007 in system research. She is an assistant in Statistics Department of Klaipėda University. Present research interests include classification of spatially correlated data, geostatistics, spatial data mining.

Vidutinė Bajeso klaidos tikimybė klasifikuojant erdvinius Gauso duomenis

Kęstutis DUČINSKAS, Lijana STABINGIENĖ

Straipsnyje sprendžiamas stacionaraus Gauso atsitiktinio lauko stebinio klasifikavimo uždavinys, esant statistiškai priklausomai mokymo imčiai. Mokymo imties klasių žymės tenkina atsitiktinio lauko modelį. Išvesta originali vidutinės Bajeso klaidos tikimybės formulė, atsisakant sąlyginės nepriklausomybės prielaidos. Skaitinė gautų formulių analizė atlikta su erdviniais Gausiniais duomenimis, stebimais taisyklingoje dvimatėje gardelėje. Gauti rezultatai atskleidžia klasifikavimo klaidų priklausomybės pobūdį nuo statistinių parametrų ir mokymo imties didumo.