Engineering Applications of Artificial Intelligence 65 (2017) 330–335

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Cognitive validation map for early occupancy detection in environmental sensing

I.M. Kulikovskikh

Information Systems and Technologies Department, Samara National Research University, 34 Moskovskoye sh., 443086 Samara, Russia

ARTICLE INFO

Keywords:
Cognitive ecology
Logistic regression
Environmental sensing
Occupancy detection
Cognitive map

ABSTRACT

Most environmental parameters are clearly indicative of occupants' presence and subtle changes in their behavior. However, this variation in sensor data makes it challenging to create a proper measure of occupancy detection that is both robust and clearly interpretable in an unstable environment. The present study addresses this problem from a cognitive ecology perspective proposing a cognitive validation map. This map is based on the extension of logistic regression that involves two extra parameters — forgetting and guessing factors. The mutual regulation of these factors creates a unique cognitive validation map that adapts the measure to evolving requirements in environmental sensing. The results of computational experiments on the proposed measure demonstrated better occupancy detection under more unstable conditions: on sensor data with fewer observations or more predictors. For this reason, the measure based on a cognitive validation map seems promising in early occupancy detection problems, but may be readily extended to a broader range of practical applications.

1. Introduction

An extensive body of literature exists on the issue of occupancy detection which is based on intellectual control systems (Candanedo and Feldheim, 2016b; Chen et al., 2016; Cocana-Fernandez et al., 2016; Ferreira et al., 2017; Hailemariam et al., 2011; Lam et al., 2009). Different types of heating, lighting, air conditioning (HVAC) systems specifically address the problem of energy consumption and security protection (Candanedo and Feldheim, 2016b; Derksen et al., 2015; Erickson et al., 2010; Roetzel and Tsangrassoulis, 2012; Yang et al., 2012). The efficiency of the solution to this problem greatly depends on the methods of occupancy detection.

These methods may be classified into direct and indirect (Candanedo and Feldheim, 2016b; Huang et al., 2009; Jin, 2016). Direct methods prioritize detection accuracy over occupants' privacy employing vision- and tag-based systems. Indirect methods, in contrast, focus on less intrusive sensors: passive infrared (PIR), pressure sensors, and environmental sensors to measure carbon monoxide (CO), volatile organic compounds, small particulates, particle pollution, light, temperature, humidity, and so on. Thus, environmental sensing introduces privacy-performance trade-off (Candanedo and Feldheim, 2016b; Huang et al., 2009; Jin, 2016) that involves mitigating an invasion of occupants privacy and ensuring the accuracy of occupancy detection.

A critical element of regulating the levels of occupancy comfort and energy performance in environmental sensing is a clear understanding of the complex interaction between occupants and their indoor environment. Most environmental parameters are sensitive enough to reveal occupants' presence and subtle changes in their behavior. But this considerable variation in sensor data makes it challenging to create a proper measure of occupancy detection. The present study is an attempt to address this issue from a cognitive ecology perspective (Marewski and Schooler, 2011; Heft, 2013). Cognitive ecology puts forward an "enactive" approach to data processing (Palacios and Bozinovic, 2003): cognition involves an active transformation of sensor data into meaningful relationships between occupants and their environment. This approach seems promising compared to a passive interpretation of internally represented data, but requires a measure highly adaptive to evolving requirements (Amato et al., 2015; Dragone et al., 2015). Consequently, the objective of present study is to propose a measure for accurate occupancy detection that would be both robust and easily interpretable in an unstable environment.

Logistic regression (LR) seems to be the proper measure due to having lower bias, making fewer assumptions in comparison with other linear classifiers, and deeper understanding the role of predictors (Donnelly and Verkuilen, 2017; Hastie et al., 2013; McCulloch et al., 2009). These advantages help to guarantee remarkable performance in a wide range of practical applications (de Menezes et al., 2017; Donnelly and

E-mail address: kulikovskikh.i@gmail.com.

http://dx.doi.org/10.1016/j.engappai.2017.08.008
Received 9 January 2017; Received in revised form 26 July 2017; Accepted 9 August 2017

Verkuilen, 2017; Hastie et al., 2013; Hosmer and Lemeshow, 2000; Li et al., 2012; Kulikovskikh, 2017). However, they may be easily outweighed due to complete separation of classes (Candanedo and Feldheim, 2016b; Ding and Gentleman, 2005; Firth, 1992a, b; Fort and Lambert-Lacroix, 2005; Gelman et al., 2008; Heinze and Schemper, 2002; Park and Hastie, 2007).

The problem of separation primarily arises in small datasets with several unbalanced and highly predictive features. Moreover, the phenomena of separation may also occur with small to medium-sized datasets when at least one LR parameter is infinite even if the likelihood converges (Candanedo and Feldheim, 2016b; Firth, 1992a, b; Gelman et al., 2008; Heinze and Schemper, 2002). This means that classes can be perfectly separated by a single feature or by a non-trivial linear combination of features. Finally, the problem of separation may arise if the underlying model parameters are low in an absolute value (Heinze and Schemper, 2002).

Previous research suggests a number of solutions to the problem of separation such as adopting partial least squares (Ding and Gentleman, 2005; Fort and Lambert-Lacroix, 2005) and iteratively reweighted least squares (Firth, 1992b). Another approach to deal with the separable data consists in penalizing the maximum likelihood (Hastie et al., 2013; Fort and Lambert-Lacroix, 2005). An alternative solution is to apply prior distributions to the likelihood function as suggested in (Hastie et al., 2013; Firth, 1992a; Gelman et al., 2008). In particular, Jeffreys prior distribution (Firth) (Firth, 1992a, b; Hastie et al., 2013; Heinze and Schemper, 2002; Park and Hastie, 2007) that is developed to reduce the bias of maximum likelihood estimates in generalized linear models has been shown to provide an ideal solution to separation. However, in spite of reliable computational results, these estimates are not clearly interpretable as prior information in a regression context (Gelman et al., 2008). Donnelly and Verkuilen (2017) also highlighted the problem of proper interpretability observing complete or a lack of separability in context of floor or ceiling effects (Hastie et al., 2013; McCulloch et al., 2009). The logit transformation permits to move proportions away from the ceiling or floor by adding half a success and half a failure. Even if empirical logit analysis helps to cope with convergence issues, it little addresses the real problem: the estimates of model parameters from logistic regression and empirical logit analysis rest on different assumptions (Donnelly and Verkuilen, 2017). So, the model based on the empirical logit function should be interpreted cautiously.

This study looks at the separability problem from a cognitive point of view. It helps to propose an extension of LR that looks appropriate to handle both stability and interpretability issues. The logic lying behind this extension is discussed at greater length in the next section.

2. Material and methods

2.1. Problem statement

Let $(x_i, y_i)_{i=1}^m$ be independent and identically distributed observations with responses $y_i \in \{0,1\}$. The matrix $X \in R^{m\times n}$ can be presented either as $X = [x_1, \dots, x_n]^T$, with vectors of predictors $x_i \in R^n$, or as $X = [x^1, \dots, x^m]$, with vectors of features $x^j \in R^m$. Let y denote the response vector, i.e. $y = [y_1, \dots, y_n]^T$. Then, for any vector $\theta \in R^n$ of regression coefficients LR models the class conditional probabilities $p(x_i, \theta) = P(y_i = 1|x_i, \theta)$ by

$$\ln\left(\frac{p(x_i,\theta)}{1-p(x_i,\theta)}\right) = \theta^T x_i.$$

Problem 1. Let g denote the link (logit) function that specifies the relationship between the class conditional expectation of the response variable and the underlying linear model $g(E[y_i|x_i]) = \theta^T x_i$. For LR, $E[y_i|x_i] = p(x_i, \theta)$. Consequently,

$$g(p(x_i,\theta)) = \ln\left(\frac{p(x_i,\theta)}{1-p(x_i,\theta)}\right), \tag{1}$$

$$p(x_i,\theta) = \frac{1}{1+\exp(-\theta^T x_i)}. \tag{2}$$

Under the above model (2) the negative log-likelihood expressed as

$$\ln L(\theta) = -\sum_i (y_i \ln(p(x_i,\theta)) + (1-y_i)\ln(1-p(x_i,\theta))) \tag{3}$$

needs to be minimized to solve the following problem

$$\theta = \arg\min_{\theta\in R^n} \ln L(\theta). \tag{4}$$

Due to the problem of clear separation between the classes, (4) fails to converge: a floor or ceiling effect is present (Donnelly and Verkuilen, 2017; Hosmer and Lemeshow, 2000; Huang et al., 2009). The logit function (1) may go to $-\infty$ for 0 successes and ∞ for 0 failures.

Let us now consider the model (2) from a cognitive point of view to directly address the problem of floor or ceiling effects.

2.2. Cognitive validation maps

Macready and Dayton (1977) proposed a model for the assessment of mastery. Mastery means that a test-taker can correct respond to all test items. His or her true response vector contains only ones. A nonmaster, in contrast, does not have the knowledge and skills for passing any item. In this case, his or her true response vector contains only zeros indicating that a test-taker failed all items. This all-or-none concept of mastery covers some aspects of the above-mentioned ceiling or floor problem, but the model of Macready and Dayton makes two underlying assumptions: (1) an observed failure of a master stems from forgetting; (2) an observed success of a nonmaster is attributed to guessing. Let us formalize this model.

Definition 1 (*Macready and Dayton*)**.** Let the probabilities of pure guessing and forgetting for item i are c_i and b_i, respectively. Then, the probability of responding correctly p_i to item i is equal to $1 - b_i$ for a master and c_i for a nonmaster.

Then, Birnbaum (1968) and van der Linden (1978) suggested that forgetting may imply guessing. According to their arguments, there is no marked difference between a nonmaster who does not know the answer and a master who has forgotten it. Both can try to guess to succeed. In addition, instead of the step function that considers only mastery and nonmastery, Birnbaum (1968) implemented the tree-parameter logistic model — a latent trait model (Lord and Novick, 1968). With reference to these corrections, let us extend Definition 1.

Definition 2 (*Birnbaum*)**.** Let θ is a latent trait, α_i is a discrimination power of item i, β_i is the difficulty of item i, c_i is the probability of pure guessing. Then, the probability of responding correctly to item i can be given as follows:

$$p_i(\theta) = c_i + \frac{1-c_i}{1+\exp(-\alpha_i(\theta-\beta_i))}. \tag{5}$$

The present study primarily focuses on Definition 2 to address the problem of floor and ceiling effects in LR. For clarity, however, it would be beneficial to distinguish between guessing and forgetting factors. Following the logic of Definition 2, let us extend the logit function (1) and LR model (2) with respect to guessing $c_{guess} \equiv c_g$ and forgetting $c_{forget} \equiv c_f$ factors as

$$g(p(x_i,\theta,c_g,c_f)) = \ln\left(\frac{p(x_i,\theta,c_g,c_f)-c_g}{1-p(x_i,\theta,c_g,c_f)-c_f}\right), \tag{6}$$

$$p(x_i,\theta,c_g,c_f) = c_g + \frac{1-c_g-c_f}{1+\exp(-\theta^T x_i)}. \tag{7}$$

Fig. 1 depicts the extensions of LR subject to a set of indicators $S = \{S_0, S_1, S_2, S_3\}$: S_0 includes no additional factors ($c_g = c_f = 0$); S_1 involves only guessing factor ($c_f = 0$); S_2 points to the presence of only forgetting factor ($c_g = 0$); S_3 means that both extra factors are non-zero and in the extended LR.

Let us now move on to extend Problem 1.

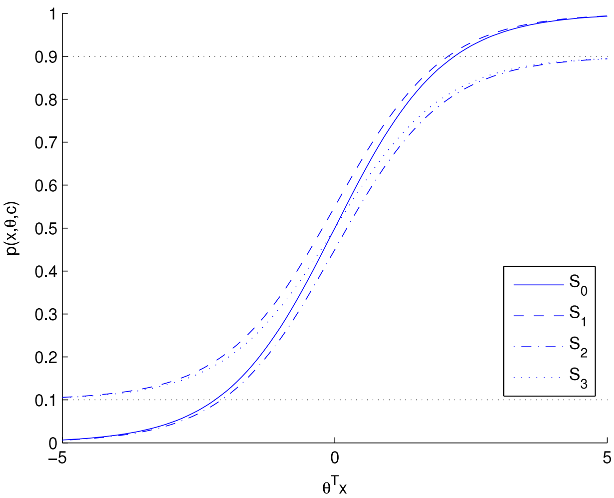

Fig. 1. LR extensions with respect to $(c_g, c_f)^S = (0.1, 0.1)$.

Table 1
A brief description of the datasets.

$\lvert X^m \rvert, X^m = (x_i, y_i)_{i=1}^m$			
m	$y_i \in \{0,1\}$	$y_i = 0$	$y_i = 1$
l	8142	6414	1728
k_1	2664	1693	971
k_2	9751	7703	2048

Problem 2. Let $g(p(x_i, \theta, c_g, c_f))$ and $p(x_i, \theta, c_g, c_f)$ given by (6) and (7), respectively. Under these models, the negative log-likelihood that includes guessing and forgetting factors can be defined as

$$\ln L(\theta, c_g, c_f) = -\sum_i (y_i \ln(p(x_i, \theta, c_g, c_f)) + (1 - y_i)\ln(1 - p(x_i, \theta, c_g, c_f))). \tag{8}$$

Consequently, (4) can be extended to

$$(\theta, c_g, c_f) = \arg \min_{\substack{\theta \in R^n \\ (c_g, c_f) \in [0,1]}} \ln L(\theta, c_g, c_f). \tag{9}$$

From here it follows that a *cognitive validation map* can be introduced as the solution to

$$(c_g, c_f) = \arg \min_{(c_g, c_f) \in [0,1]} \ln L(\theta, c_g, c_f). \tag{10}$$

for each given $\theta \in R^n$.

2.3. Datasets

This section briefly describes the datasets used to detect occupancy through environmental sensing. These datasets are freely available from (Candanedo and Feldheim, 2016a) and thoroughly discussed in (Candanedo and Feldheim, 2016b). The raw datasets were amassed from a number of sensors located in an office room (5.85 m × 3.50 m × 3.53 m) to measure temperature, humidity, light, and CO_2 levels (Candanedo and Feldheim, 2016b). In addition, a digital camera was used to label the datasets with occupancy status: occupied (class 1) or not (class 0). To extend the feature space for better occupancy detection, the authors (Candanedo and Feldheim, 2016b) also included an additional feature — a humidity ratio. It was calculated according to the measured temperature and relative humidity.

The collected datasets were used for both training and testing. The training dataset was recorded when the door in the office room was mostly closed. Two testing datasets involve both conditions: a closed door and an open door.

Let us now turn to the formal description of the datasets. Consider observations $X^m = (x_i, y_i)_{i=1}^m$ with responses $y_i \in \{0,1\}$. According to the obtained measurements, m is equal to: $l = 8142$ for the training dataset; $k_1 = 2664$ for the closed-door test dataset; $k_2 = 9751$ for the opened-door test dataset. To summarize, Table 1 presents the relevant information on these observations.

Following the provided denotations, let us define the vector of predictors as $x_i \in R^n$, while the vector of features as $x^j \in R^m$, where $n = 5 : x^j = \{Temperature, Humidity, Light, CO_2, HumidityRatio\}$.

Table 2
Prediction error subject to $m = \{1000, 500, 200\}$, $d = \{10, 5, 1\}$.

$(c_{guess}, c_{forget})^{l_1, l_2}$	X^{l_1}	X^{l_2}
$m = 1000$, $d = 10$		
$(0,0)$	0.0	0.0554
$(0, 0.002)^{l_1}$, $(0.06, 0.002)^{l_2}$	0.0148	**0.0312**
$m = 1000$, $d = 5$		
$(0,0)$	0.0169	0.0185
$(0, 0.08)^{l_1}$, $(0, 0.08)^{l_2}$	0.0164	**0.0181**
$m = 1000$, $d = 1$		
$(0,0)$	0.0291	0.0322
$(0.02, 0)^{l_1}$, $(0.02, 0)^{l_2}$	0.0259	**0.0272**
$m = 500$, $d = 1$		
$(0,0)$	0.0253	0.0414
$(0.012, 0)^{l_1}$, $(0.02, 0.002)^{l_2}$	0.0235	**0.0311**
$m = 200$, $d = 1$		
$(0,0)$	0.0201	0.057
$(0.002, 0)^{l_1}$, $(0.04, 0)^{l_2}$	0.02	**0.0377**

Table 3
Prediction accuracy subject to $m = \{1000, 500, 200\}$, $d = \{10, 5, 1\}$.

$(c_{guess}, c_{forget})^{l_1, l_2}$	X^{k_1}	X^{k_2}
$m = 1000$, $d = 10$		
$(0,0)$	0.8624	0.6627
$(0, 0.002)^{l_1}$, $(0.06, 0.002)^{l_2}$	**0.9052**	**0.8298**
$m = 1000$, $d = 5$		
$(0,0)$	0.8793	0.8205
$(0.0, 0.08)^{l_1}$, $(0, 0.08)^{l_2}$	**0.889**	**0.8342**
$m = 1000$, $d = 1$		
$(0,0)$	0.8826	0.9307
$(0.02, 0)^{l_1}$, $(0.02, 0)^{l_2}$	0.7115	0.9198
$m = 500$, $d = 1$		
$(0,0)$	0.8768	0.8293
$(0.012, 0)^{l_1}$, $(0.02, 0.002)^{l_2}$	0.7469	**0.9232**
$m = 200$, $d = 1$		
$(0,0)$	0.7618	0.8039
$(0.002, 0)^{l_1}$, $(0.04, 0)^{l_2}$	**0.7841**	**0.8285**

Table 4
Prediction accuracy subject to $m = \{100, 200, 500\}$, $d = 10$.

$(c_{guess}, c_{forget})^{l_1, l_2}$	X^{k_1}	X^{k_2}
$m = 100$		
$(0,0)$	0.8695	0.6531
$(0, 0.016)^{l_1}$, $(0, 0.16)^{l_2}$	**0.8844**	**0.8579**
$m = 200$		
$(0,0)$	0.8773	0.7806
$(0, 0.04)^{l_1}$, $(0, 0)^{l_2}$	**0.9757**	**0.8643**
$m = 500$		
$(0,0)$	0.8697	0.6679
$(0, 0.06)^{l_1}$, $(0, 0.06)^{l_2}$	**0.8902**	**0.8511**

Table 5
Prediction estimates subject to $m = \{1000, 500, 200\}$, $d = 1$.

$(c_{guess}, c_{forget})^{l_1}$	θ_1	θ_1	θ_2	θ_3	θ_4	θ_5
$m = 1000$						
$(0, 0)$	−5.1427	−0.0419	2.9546	3.9692	2.2677	−2.8657
$(0.02, 0)$	−5.6822	−0.2144	−0.5306	3.3354	1.8976	−0.6816
$m = 500$						
$(0, 0)$	−5.1169	2.4699	14.577	4.7341	2.7537	−14.599
$(0.012, 0)$	−5.7385	−0.1183	−0.7268	3.6182	2.4594	−0.8095
$m = 200$						
$(0, 0)$	−24.704	−20.138	−0.565	16.827	20.689	−5.539
$(0.002, 0)$	−14.156	−9.5	0.9064	10.02	8.439	−1.477

3. Results

This section describes the results of computational experiments conducted to test the stability and interpretability of the proposed measure on the described datasets.

The dataset X^m was divided into the training subset X^{l_1} and the validation subset X^{l_2} using 5-fold cross validation. To increase a chance of identifying the separation problem, the experiments suggested varying the limited number of observations. As for small to moderate sample sizes the resampling estimates are better than the asymptotic estimates, the bootstrap method was adopted to provide reliable results. In addition, to vary a number of predictors $x^j = f(x^j)$ for different pairs $\{x^p, x^q\}$, where $\{p, q\} \subset j$, the following feature extraction function was employed:

$$f(x^p, x^q, d) = \prod_{m=1}^{d} \prod_{n=0}^{m} (x^p)^{m-n} (x^q)^n.$$

Example 1. Let $m = \{1000, 500, 200\}$ and $d = \{10, 5, 1\}$. Table 2 presents the values of $(c_{guess}, c_{forget})^{l_1, l_2}$ and prediction errors based on cognitive validation maps (10) for both training (l_1) and validation (l_2) subsets. The values of prediction accuracy on the test datasets (k_1, k_2) for given m, d, and the designed maps $(c_{guess}, c_{forget})^{l_1, l_2}$ are shown in Table 3. From the presented results the following conclusion may be drawn: even if the cognitive maps ensured better results on the validation subset for each m and d, the proposed measure seems more effective for fewer observations and more predictors. To support this point, let us consider the next example.

Example 2. Let $m = \{100, 200, 500\}$ and $d = 10$. Table 4 demonstrates the values of $(c_{guess}, c_{forget})^{l_1, l_2}$ and prediction accuracy on the test datasets (k_1, k_2). As can be seen, the validation maps helped to achieve better results, but there is a clear difference in the impact of the proposed measure on the datasets (k_1, k_2) if $m = \{100, 500\}$. This may be explained by the conditions on which were recorded these datasets compared to the training dataset (see Section 2.3).

Example 3. Let $m = \{1000, 500, 200\}$ and $d = 1$. To look into the issue of interpretability, the predicted estimates are given (see Table 5). If we look at the presented results, we can see that the designed cognitive validation maps provided less biased estimates in case of fewer observations.

Example 4. The values of prediction errors for $m = \{50, 100, 200\}$, $d = 1$ and $m = \{100, 200, 500\}$, $d = 10$ are presented in Tables 6, 7, respectively. For clarity, the corresponding cognitive validation maps for the training subset (left) and the validation subset (right) are illustrated in Figs. 2–7. These maps show the solution to the problem (10) with regard to $c_g \in [0, 0.15]$ and $c_f \in [0, 0.2]$. As can be seen, the proposed measure made more marked improvement on the prediction results for the validation subset even if the optimal values for (c_g, c_f) help to achieve better accuracy for both subsets. Furthermore, Figs. 2–7 demonstrate that the impact of (c_g, c_f) on the measure seems to depend on the number of observations and predictors. The proposed measure produced more meaningful results in case of fewer observations or more predictors when the corresponding cognitive validation maps are more dynamic. Finally, the proposed measure may provide equal or worse results (see Table 7) on the validation dataset, but still guarantee better performance on the test datasets (see Table 4). This phenomena can also be attributed to the difference in the conditions on which the datasets were collected (see Example 2).

Table 6
Prediction errors subject to $m = \{50, 100, 200\}$, $d = 1$.

$(c_{guess}, c_{forget})^{l_1, l_2}$	X^{l_1}	X^{l_2}
$m = 50$		
$(0, 0)$	0.0076	0.0738
$(0.004, 0)^{l_1}$, $(0, 0.002)^{l_2}$	0.0065	**0.0142**
$m = 100$		
$(0, 0)$	0.0159	0.123
$(0.002, 0)^{l_1}$, $(0.02, 0)^{l_2}$	0.0158	**0.0329**
$m = 200$		
$(0, 0)$	0.0073	0.0428
$(0.002, 0)^{l_1}$, $(0.016, 0)^{l_2}$	0.0067	**0.024**

Table 7
Prediction errors subject to $m = \{100, 200, 500\}$, $d = 10$.

$(c_{guess}, c_{forget})^{l_1, l_2}$	X^{l_1}	X^{l_2}
$m = 50$		
$(0, 0)$	0	0.074
$(0, 0.016)^{l_1}$, $(0, 0.16)^{l_2}$	0.0175	**0.0657**
$m = 100$		
$(0, 0)$	0	0.009
$(0.04, 0)^{l_1}$, $(0, 0)^{l_2}$	0.0154	0.009
$m = 200$		
$(0, 0)$	0.0	0.0
$(0, 0.06)^{l_1}$, $(0, 0.06)^{l_2}$	0.0168	0.0184

It is worth noting that the most time demanding step in the measure training is the optimization of (c_g, c_f) involving the bootstrap method and 5-fold cross validation for the chosen values of the extra parameters c_g and c_f. This procedure shares a close similarity with the regularization method. However, considering the fact that the cognitive validation maps may directly deal with floor and ceiling effects, the proposed measure seems the effective solution to the posed problem.

4. Limitations and future directions

There are two main limitations that need to be addressed regarding this research. The proposed measure was tested: (1) on the same sensor data and (2) only under varying the number of observations and predictors. Thus, the following directions for future research may be suggested. First, it would be reliable to significantly extend a number of datasets to support the findings of the present study. Second, it seems

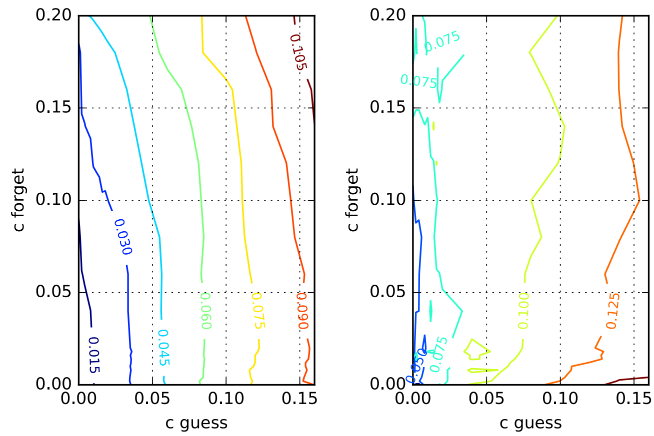

Fig. 2. The cognitive validation map subject to $m = 50$, $d = 1$.

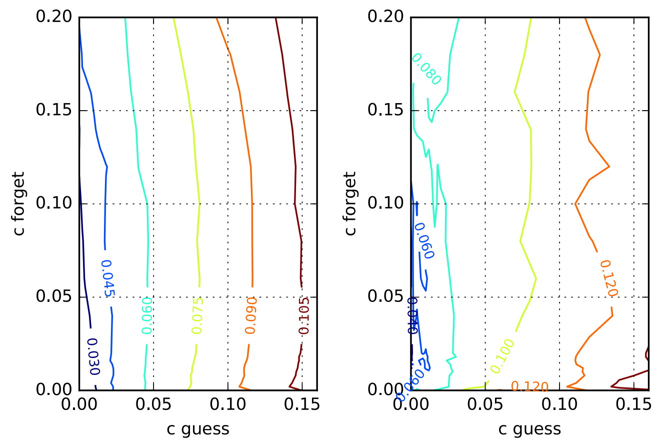

Fig. 3. The cognitive validation map subject to $m = 100$, $d = 1$.

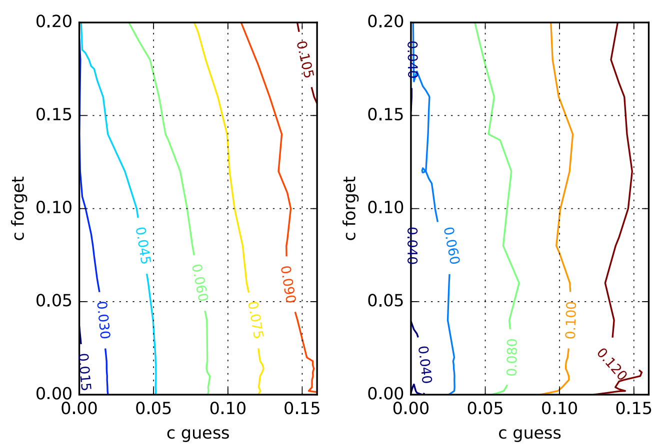

Fig. 4. The cognitive validation map subject to $m = 200$, $d = 1$.

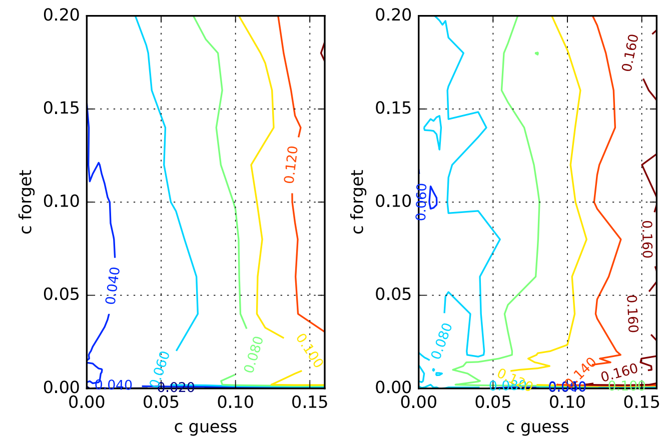

Fig. 5. The cognitive validation map subject to $m = 100$, $d = 10$.

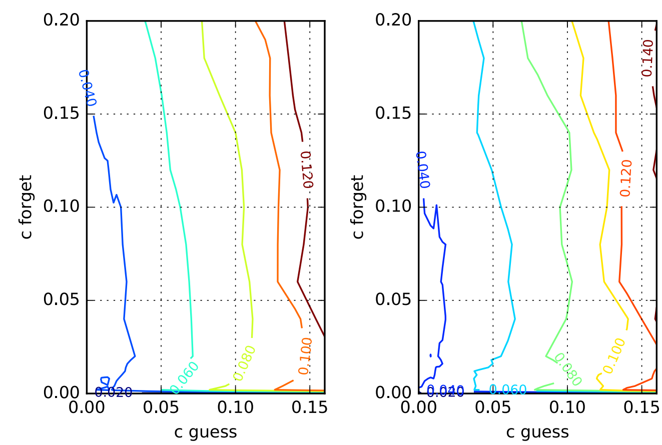

Fig. 6. The cognitive validation map subject to $m = 200$, $d = 10$.

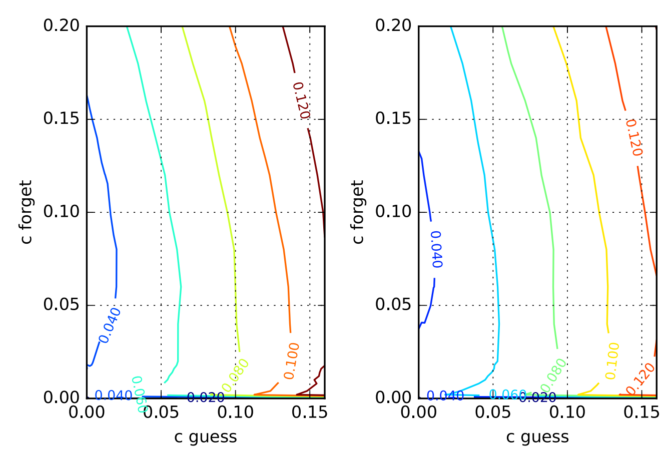

Fig. 7. The cognitive validation map subject to $m = 500$, $d = 10$.

interesting to model different evolving requirements such as varying the level of outliers in sensor data or preserving numerical stability in intelligent real-time systems.

5. Conclusions

The present research was aimed at proposing the reliable measure for accurate occupancy detection in the presence of variation in sensor data. For this purpose, a cognitive validation map based on the extension of logistic regression was proposed. The LR model was extended with regard to guessing and forgetting factors adopted from a latent trait theory. The results of computational experiments proved the validity of proposed measure and revealed its benefits in case of fewer observations or more predictors. This makes the measure based on a cognitive validation map more suitable for early occupancy detection, but still can be applied to a broader range of practical problems.

Acknowledgments

This work was supported by the Ministry of Education and Science of the Russian Federation, grant 074-U01. The author would like to thank the reviewers for the valuable comments and suggestions and to express the gratitude to Dr. Sergej Prokhorov for a constructive discussion on further research.

References

Amato, G., et al., 2015. Robotic ubiquitous cognitive ecology for smart homes. J. Intell. Robot. Syst. 80 (Suppl 1), 57–81.
Birnbaum, A., 1968. Some latent trait models and their use in inferring an examinee's ability. In: Lord, F.M., Novick, M.R. (Eds.), Statistical Theories of Mental Test Scores. Addison-Wesley, Reading, Massachusetts.
Candanedo, L.M., Feldheim, V., 2016. Occupancy Detection. UCI Machine Learning Repository. https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+ [dataset].
Candanedo, L.M., Feldheim, V., 2016. Accurate occupancy detection of an office room from light, temperature, humidity and CO_2 measurements using statistical learning models. Energy Build. 112, 28–39.
Chen, Z., Masood, M.K., Soh, Y.C., 2016. A fusion framework for occupancy estimation in office buildings based on environmental sensor data. Energy Build. 133, 790–798.
Cocana-Fernandez, A., Sanchez, L., Ranilla, J., 2016. Leveraging a predictive model of the workload for intelligent slot allocation schemes in energy-efficient HPC clusters. Eng. Appl. Artif. Intell. 48, 95–105.
de Menezes, F.S., et al., 2017. Data classification with binary response through the Boosting algorithm and logistic regression. Expert Syst. Appl. 69, 62–73.
Derksen, C., et al., 2015. Structure and classification of unified energy agents as a base for the systematic development of future energy grids. Eng. Appl. Artif. Intell. 41, 310–324.
Ding, B., Gentleman, C.M., 2005. Classification using generalized partial least squares. J. Comput. Graph. Statist. 14 (2), 280–298.
Donnelly, S., Verkuilen, J., 2017. Empirical logit analysis is not logistic regression. J. Mem. Lang. 94, 28–42.
Dragone, M., et al., 2015. A cognitive robotic ecology approach to self-configuring and evolving AAL systems. Eng. Appl. Artif. Intell. 45, 269–280.
Erickson, V.L., Carreira-Perpinan, M.A., Cerpa, A.E., OBSERVE: Occupancy-based system for efficient reduction of HVAC energy. In: Proceeding of the 10th IEEE International Conference on Information Processing in Sensor Networks (IPSN), Stockholm, Sweden, April 12–16, 2010.
Ferreira, E.F.M., Rego, P.H.M., Neto, J.V.F., 2017. Numerical stability improvements of state-value function approximations based on RLS learning for online HDP-DLQR control system design. Eng. Appl. Artif. Intell. 63, 1–19.
Firth, D., 1992a. Bias reduction, the Jeffreys prior and GLIM. In: Fahrmeir, L., Francis, B., Gilchrist, R., Tutz, G. (Eds.), Advances in GLIM and Statistical Modelling. Springer-Verlag, New York.
Firth, D., 1992b. Generalized linear models and Jeffreys priors: An iterative weighted least-squares approach. In: Dodge, Y., Whittake, J. (Eds.), Computational Statistics. Physica-Verlag, Vienna.
Fort, G., Lambert-Lacroix, S., 2005. Classification using partial least squares with penalized logistic regression. Bioinformatics 21 (7), 1104–1111.
Gelman, A., Jakulin, A., Pittau, M.G., Su, Y.-S., 2008. A weakly informative default prior distribution for logistic and other regression models. Ann. Appl. Stat. 2 (4), 1360–1383.
Hailemariam, E., et al., Real-time occupancy detection using decision trees with multiple sensor types. In: Proceeding of the 2011 Symposium on Simulation for Architecture and Urban Design, Boston, MA, USA, April 3–7, 2011.
Hastie, T., Tibshirani, R., Friedman, J., 2013. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. In: Springer Series in Statistics.
Heft, H., 2013. An ecological approach to psychology. Rev. Gen. Psychol. 17 (2), 162–167.
Heinze, G., Schemper, M., 2002. A solution to the problem of separation in logistic regression. Stat. Med. 21, 2409–2419.
Hosmer Jr., D.W., Lemeshow, S., 2000. Applied Logistic Regression, second ed. John Wiley & Sons, New York.
Huang, K.L., Kanhere, S.S., Hu, W., Towards privacy-sensitive participatory sensing. In: Proceedings of IEEE International Conference on Pervasive Computing and Communications, Galveston, TX, USA, March 9–13, 2009.
Jin, M., 2016. Occupancy detection via environmental sensing. IEEE Trans. Autom. Sci. Eng. PP (99), 1–13.
Kulikovskikh, I.M., 2017. Anomaly detection in the space of ecological features to improve accuracy of identifying activity in buildings. Comput. Opt. 41 (1), 126–133.
Lam, K.P., et al., 2009. Occupancy detection through an extensive environmental sensor network in an open-plan office building. IBPSA Build. Simul. 145, 1452–1459.
Li, L., et al., 2012. Coupling of logistic regression analysis and local search methods for characterization of water distribution system contaminant source. Eng. Appl. Artif. Intell. 25 (2), 309–316.
Lord, F.M., Novick, M.R., 1968. Statistical Theories of Mental Test Scores. Addison-Wesley, Reading, Massachusetts.
Macready, G.B., Dayton, C.M., 1977. The use of probabilistic models in the assessment of mastery. J. Educ. Stat. 2, 99–120.
Marewski, J.N., Schooler, L.J., 2011. Cognitive niches: an ecological model of strategy selection. Psychol. Rev. 118 (3), 393–437.
McCulloch, C.E., Searle, S.R., Neuhaus, J.M., 2009. Generalized, Linear, and Mixed Models, second ed. John Wiley, New York.
Palacios, A.G., Bozinovic, F., 2003. An "enactive" approach to integrative and comparative biology: thoughts on the table. Biol. Res. 36 (1), 101–105.
Park, M.Y., Hastie, T., 2007. L1-regularization path algorithm for generalized linear models. J. R. Stat. Soc. Ser. B Stat. Methodol. 69 (4), 659–677.
Roetzel, A., Tsangrassoulis, A., 2012. Impact of climate change on comfort and energy performance in offices. Build. Environ. 57, 349–361.
van der Linden, W.J., 1978. Forgetting, guessing, and mastery: The MacReady and Dayton models revisited and compared with a latent trait approach. J. Educ. Stat. 3 (4), 305–317.
Yang, Z., et al., A multi-sensor based occupancy estimation model for supporting demand driven HVAC operations. In: Proceedings of Symposium on Simulation for Architecture and Urban Design, Orlando, FL, USA, March 26–29, 2012.