The Kullback-Leibler Divergence Class in Decoding the Chest Sound Pattern

Antonio CLIM, Răzvan Daniel ZOTA
The Bucharest University of Economic Studies
antonio.clim@csie.ase.ro, zota@ase.ro

Kullback-Leibler Divergence Class or relative entropy is a special case of broader divergence. It represents a calculation of how one probability distribution diverges from another one, expected probability distribution. Kullback-Leibler divergence has a lot of real-time applications. Even though there is a good progress in the field of medicine, there is a need for a statistical analysis for supporting the emerging requirements. In this paper, we are discussing the application of Kullback-Leibler divergence as a possible method for predicting hypertension by using chest sound recordings and machine learning algorithms. It would have a major outreached benefit in emergency health care systems. Decoding the chest sound pattern has a wide degree in distinguishing different irregularities and wellbeing states of a person in the medicinal field. The proposed method for the estimation of blood pressure is chest sound analysis using a method that creates a record of sounds delivered by the contracting heart, coming about because of valves and related vessels vibration and analyzing it with the help of Kullback-Leibler divergence and machine algorithm. An analysis using the Kullback-Leibler divergence method will allow finding the difference in chest sound recordings which can be evaluated by a machine learning algorithm. The report also proposes the method for analysis of chest sound recordings in Kullback-Leibler divergence class.
Keywords: *Relative entropy, k-nearest neighbors algorithm, Bolster vector machines, Gaussian blend display, Chest sound pattern*

1 Introduction

Assessment of lung sounds is a standard piece of a clinical examination. Auscultation of the chest remains an important clinical device is still presumably the most widely recognized strategy for assessing the health. Decoding the chest sound pattern has a wide scope in detecting various abnormalities and health conditions of an individual in the medical field. The utilization of heart sound as a surrogate of the blood pressure (BP) has been utilized with more accentuation on the identification of pneumonic hypertension patients, considering the recurrence substance, sufficiency, and split the second heart sound and its components, which emerged following the conclusion of the relating heart valves. Estimation of BP utilizing the investigation of heart sound is portrayed by the straightforwardness of the hardware used to acquire information, which after examination permits to accomplish promising outcomes that as of recently were just gotten with methods significantly more intricate and costly gear [13].

The current method in practice for measuring the BP is Doppler echocardiography; which uses ultrasound technology to measure the BP by analyzing the rate of blood flow through the heart. However, continuous monitoring is impossible using this technique. Also this technique cannot be assessed in 34 to 76% of patients with interminable obstructive pneumonic infection, 10 to 20% of patients with hoisted pulmonary alveolar phospholipoproteinosis (PAP), and the most critical in around half of the patients with typical PAP [19]; likewise the Doppler estimations introduce a usual blunder of 30% contrasted with right heart catheterization [8]. So the proposed method for the estimation of BP is chest sound analysis using a method that creates a record of sounds delivered by the contracting heart, coming about because of valves and related vessels vibration and analyzing it with the help of Kullback-Leibler (KL) divergence and machine algorithm. We investigate the different variations in the chest sound to produce an optimum result.

DOI: 10.12948/issn14531305/23.1.2019.05

KL Divergence Class or relative entropy is a particular case of broader divergence. It is a calculation of how one probability distribution diverges from another one, expected probability distribution. We present a basic yet successful approach to utilize KL uniqueness in the arrangement of homemade signs of chest sounds recorded. We demonstrate that machine learning calculation with KL uniqueness as the separation measure, when utilized our component vectors, gives focused arrangement exactness and reliably beats the all the more regularly employed Euclidean k-Nearest Neighbors (k-NN). We likewise create and exhibit the utilization of a KL-based piece to order chest sounds information utilizing bolster vector machines (SVMs). Our KL remove based bit thinks about positively to other entrenched parts, for example, direct and outspread premise work (RBF) portion [16].

2 Methods

The proposed method for the estimation of BP is chest sound analysis using a technique that creates a record of sounds delivered by the contracting heart, coming about because of valves and related vessels vibration and analyzing it with the help of KL divergence and machine algorithm. A programmed lung sound examination framework primarily advances through the ventures of pre-handling, highlight extraction and order. In pre-handling step, gathered flag is set up for the ensuing handling by decreasing the heart sound impact, separating or test rate change and adequacy standardization. Highlight extraction includes inferring a smaller portrayal of an extensive arrangement of information without losing discernable data. Then again, order allocates diverse signs to their relating gatherings [11]. A deliberate what's more, broad examination of different highlights and classifiers utilized for the system based lung sound acknowledgement can be found in a current paper. Wavelet-based highlights turned out to be the most prevalent among these since they center around non-stationary of lung sounds. Contrasted with this, be that as it may, cepstral highlights alongside Gaussian blend display (GMM) classifier have demonstrated better characterization comes about. Despite the fact that Gaussian blend display (GMM) and mel-recurrence cepstral coefficients (MFCCs) yielded better acknowledgement exactness on account of crackle, it was when MFCCs were utilized alongside artificial neural network (ANN) system, that nearly better outcomes for wheeze and ordinary lung sounds were enrolled [8].
Deliberate gathering of lung sound examples with a solid ground truth is a critical piece of this research. In our study, we consider three types of lung sounds: normal, wheeze and crackle [13]. The database incorporates three sorts of lung sounds: typical, crackle (both fine and coarse crackle) and wheeze. As the heart sound is blended with the recorded lung sound, the lung sound division is a fundamental advance before the examples are demonstrated and described. In this work, the impact of heart sound has been decreased from the gathered information by receiving the strategy proposed [9].

2.1 Data Sets

Were utilized two informational collections as a part of this examination [11] [20]. It was based on the sound objects as these were recorded. The principal set contained singular developments individual movements (IM) information which included 20 subjects, Analog-to-digital conversion (ADC) of 14 bits, four electromyography (EMG) channels, and eleven hand open/close, genius/supination, wrist flexion/expansion, side grasp, pointer or record expansion, subtle hold, concur or thumb up, and rest or no development classes). Another set considered contained individual and concurrent developments of simultaneous movements (SM) information which included 17 subjects, 8 EMG channels, ADC of 16 bits, and 27 hand open/close, star/supination, wrist flexion/expansion and all their conceivable mixes classes.
Dispensable Ag/AgCl (Ø = 1 cm) terminals in a bipolar setup (2 cm between terminal separation) were utilized as a part of the two sets [11]. There is an equal distribution of the bipoles around the most proximal third in the lower arm, with the principal channel set

DOI: 10.12948/issn14531305/23.1.2019.05

along the extensor carpi ulnaris. Subjects were situated serenely with their elbow flexed at 90 degrees and lower arm bolstered, leaving just the hand to move openly. The informational collections, alongside points of interest on socioeconomics and securing equipment, are accessible online as a feature of BioPatRec [20].

2.2 Characteristics of chest sounds

Lung sounds can be comprehensively classified into ordinary, irregular, and extrinsic. Regular lung sound alludes to the respiratory sound of a solid subject, barely perceptible without a stethoscope. Despite what might be expected, strange sound demonstrates the nonappearance or diminishing of ordinary sound while unusual sound applies to the superposition of extra stable with average sound. Unusual sounds can again be constant or spasmodic. Among the few notable lung sound occasions, often run over in the therapeutic field viz. wheeze, crackle, hack, rhonchus, cackle, stridor et cetera, two most regular unusual sounds wheeze (persistent) and crackle (spasmodic) are considered for point by point ponder in this paper. Regularly, either wheeze or crackle is found to be a trademark highlight of the vast majority of alternate extrinsic aspiratory sounds specified before. Acoustically, wheezes describe hack sounds, even the asthmatic ones. Rhonchus is a low-pitched wheeze, cackle is a short wheeze gone before by a crackle, and stridor noticeably is low-recurrence wheeze [4].

2.3 Signal gathering and processing

BioPatRec recording schedules guided the subjects to play out every development three times with resting periods in the middle. The taught withdrawal time, and in addition, the resting time, was 3 s. The underlying and last 15% of every compression was disposed of as this ordinarily relates to a postponed reaction and expectant unwinding by the subject, while the staying focal 70% still jam bits of the dynamic compression. Time windows of 200 ms were removed from the linked withdrawal information utilizing 50 ms time increase. Highlights were then removed from each time window and conveyed in sets utilized for preparing (40%), approval (20%), and testing (40%) of the classifiers. The testing sets were never observed by the classifier amid preparing or approval. A 10-crease cross-approval was performed by randomizing the highlight vectors between the three sets previously preparing, what's more, trying [5].

2.4 Machine learning algorithms

In machine learning issues that include taking in a "state-of-nature" or perhaps an unending circulation from a limited number of information tests in a high-dimensional component space with each element having various conceivable esteems, a huge measure of preparing information is required to guarantee that there are a few examples with each mix of qualities. With a settled number of preparing tests, the prescient power diminishes as the dimensionality increments, and this is known as the Hughes impact or Hughes marvel [14].

The arrangement many-sided quality assessing calculations CCEAs (classification complexity estimating algorithms) were intended to return arrangement many-sided quality gauges CCE (classification complexity estimates) for every development independently (person result) and arrived at the midpoint of overall developments (normally comes about). Singular outcomes give data that encourages the selection of events to be incorporated into guaranteed myoelectric pattern recognition (MPR) issue by recognizing clashing classes. Usual result considers the entire component space, including all developments, and can in this manner be utilized to assess and contrast highlight sets utilized with construct the component space [14].

2.5 Probabilistic Approach

Since probabilistic methodologies have diverse details, we have picked one approach which adjusts along a typical execution stage for these three systems. It is utilized as a meta-student, with k-Nearest Neighbors (k-NN) in every one of its emphases as the base calculation in this examination. The likelihood conveyance of k-NN on the yield marks is proportional to a Naive Bayes classifier utilizing

DOI: 10.12948/issn14531305/23.1.2019.05

non-parametric thickness estimation with variable window sizes. Likewise, the weighted expansion of these likelihood values over the cycles in the structure has been translated as a Bayesian joining of likelihood esteems. This weighted whole of the likelihood values [2].

2.6 Separability index

Separability index (SI) was executed as presented by Bunderson et al. [18]; that is, the normal of the separations between all developments and their generally clashing neighbor. It is to be a large portion of the KL removes, coming about in the below condition:

$$SI = \sum_{i=1}^{K} \left(\min_{j=1,\ldots i-1, i+1,\ldots,K} \frac{1}{2} \sqrt{\left(\mu_i - \mu_j\right)^T S_i^{-1} \left(\mu_i - \mu_j\right)} \right)$$

where,

K is the number of classes,

μ_x and S_x are mean vectors and covariance matrices for class x, respectively, according to Nilsson et al. [11].

This definition just considers the covariance of the target development (Si), and not that of the contrasting development (that is, Sj). We thought about this specific definition as potential confinement, so we presented extra separation definitions. The separation definitions were utilized under the presumption of ordinariness as KL separate was characterized under a similar hypothesis [3].

2.7 Nearest neighbor separability

Nearest neighbor separability (NNS) was roused by the calculation with a similar name characterized by Singh [17]. It depends on the predominance of closest neighbors, in highlight space, having a place with a similar class (development) as a target information point. The commitments of the closest neighbors are weighted by their vicinity to the objective point, and the outcome is standardized to be an incentive between 0 and 1. Let

$$b\ (p_t, p_i) = \begin{cases} 1, & if\ p_t, p_i\ \in C \\ 0, & if\ p_t\ \in C, p_i\ \notin C \end{cases}$$

where,

p_t is the target point,

p_i is p_t's i-th nearest neighbour

C is a class

Dominance [11] can be defined as,

$$d_t = \left(\sum_{i=1}^{K} \frac{1}{i} \right)^{-1} \sum_{i=1}^{K} \frac{b(p_t, p_i)}{i}$$

Numerous bunching strategies utilize separate measures to decide the likeness or divergence between any combine of items. The most usually utilized separation measure is known as Euclidian Distance [1], but the least-squares estimators could be here the preferred linear unbiased estimators, when very high accuracy is needed, despite its higher compute. Diverse measures are utilized for various sorts of traits, especially when anyone assumes the vectors are independent and distributed. For traits of sort double, ostensible and blended write properties, the remove between focuses can be characterized as takes after [10].

2.8 Kullback-Leibler Divergence

The raw data storage will require a considerable amount of memory, so the analysis of each sound variation is used to create a parametric, statistical model, KL Divergence. The probability density function for any selected variable can be defined as,

DOI: 10.12948/issn14531305/23.1.2019.05

$$p(x) = \sum_{k=1}^{K} C_k \frac{1}{\sqrt{|2\pi\ \varepsilon_k|}} \exp\left(-\frac{1}{2}(x-\mu_k)^T \sum\nolimits_k^{-1}(x-\mu_k)\right)$$

where,
K is the number of mixtures
$\mu_k, \varepsilon_k, c_k$ are the mean, covariance matrix, and weight of the k-th divergence respectively.
For, K = 1,

$$\varepsilon_{ML} = \frac{1}{M}\sum_{n=1}^{M}(x_n - \mu_{ML})\ (x_n - \mu_{ML})^T$$

For, K > 0, the k-implies calculation took after by the desire boost calculation.
As specified, we utilize the symmetrized KL disparity between the Gaussian blends as separation measure between the two chest sound varieties. The KL divergence is an asymmetric information theoretic measure of the separation between two likelihood density functions [15].
The KL divergence is given by,

$$d_{KL}(p_1, p_2) = \int p_1(x) \log\frac{p_1(x)}{p_2(x)}\, dx$$

For discrete arbitrary factors$d_{KL}(p_1, p_2)$, is the punishment of outlining a code that portrays information with appropriation $p_2(x)$ with most brief conceivable length however rather utilise it to encode information with distribution $p_1(x)$. In the event that both are close, the punishment will be little and the other way around. For two multivariate Gaussian appropriations, $p_1(x)$ and $p_2(x)$, the KL dissimilarity is given in shut shape by,

$$d_{KL}(p_1, p_2) = \frac{1}{2}\left[\log\left(\frac{|\varepsilon_1|}{|\varepsilon_2|}\right) + tr\ (\varepsilon_1^{-1}\varepsilon_2) + (\mu_1 - \mu_2)^T \varepsilon_1^{-1}(\mu_1 - \mu_2) - D\right]$$

where,
D is the dimensionality of x .
For Gaussian mixtures, a closed form expression for $d_{KL}(p_1, p_2)$does not exist, and it must be estimated.
Considering the two KL divergences under a single integral,

$$d_{KL}(p_1, p_2) = \int d'_{sKL}(p_1(x), p_2(x))\, dx$$

where,

$$d'_{sKL}(p_1(x), p_2(x)) = \left(p_1(x) - p_2(x)\right) \log\frac{p_1(x)}{p_2(x)}$$

Since the value of $d'_{sKL}$is considered to be large, there is a need of x to be present for making the difference and ratio between $p_1(x), p_2(x)$ to be large. So we are considering the distance as the square of L2.

DOI: 10.12948/issn14531305/23.1.2019.05

$$d_{L2}(p_1, p_2)^2 = \int d'_{L2sqr}(p_1(x), p_2(x))\, dx$$

where,

$$d'_{L2sqr}(p_1(x), p_2(x)) = (p_1(x), p_2(x))^2$$

Tentatively, utilizing the L2 remove between Gaussian blend models does not function admirably for kind grouping. In unpublished closest neighbor investigates the ISMIR 2004 kind order preparing the set, they got 42% precision utilizing the L2 remove contrasted with 65% utilizing the symmetrized KL difference (in the analyses, closest neighbor melodies by an indistinguishable craftsman from the question tune were disregarded). From this no doubt the accomplishment of the symmetrized KL disparity in music data recovery is critically connected to it asymptotically going towards limitlessness when one of and goes towards zero, i.e., it profoundly punishes contrasts. This is bolstered by the perception in that exclusive a minority of sounds MFCCs, as a matter of fact, segregate it [8].
The repeatability record (RI) measures how much person classes shifts between various occurrences of the chest sound utilizing KL separate. The three redundancies amid the chronicle session were the events that were assessed. The final product is the normal KL separates between the primary redundancy and the accompanying ones for all developments [15].
With Gaussian blend models, the covariance lattices are regularly thought to be the corner to corner for computational effortlessness. In the equations aforementioned, it was demonstrated that rather than a Gaussian blend show where each Gaussian part has askew covariance framework, a solitary Gaussian with full covariance lattice could be utilized without giving up segregation execution. This streamlines both preparing and assessment since the shut shape articulations in the equations can be utilized. In the event that the converse of the covariance lattices is pre-computed, this can be assessed efficiently since the following term just requires the inclining components to be figured. For the symmetric variant, the log terms even cross out, along these lines not notwithstanding requiring the determinants to be pre-computed [8].

3 Classifiers and Topologies

We have used three normal classifiers for the MPR as a part of this contemplate which are multi-layer perceptron (MLP), linear discriminant investigation (LDA), and bolster vector machine (SVM) [8]. A quadratic part work was utilized for SVM. These classifiers were employed as actualized in BioPatRec where code accessible online, where LDA and SVM were actualized utilizing Matlab's factual tool compartment. MLP and SVM are intrinsically equipped for concurrent order when furnished with the component vectors of blended (concurrent) yields, henceforth alluded as "Blend" yield designs; that is, there is one yield for each particular development, and mixes of events create the relating blend of yields to be turned on [8]. LDA's return is processed by dominant part voting, which implies it cannot create synchronous grouping by making a blended yield. Be that as it may, classifiers which include LDA can at present be utilized for concurrent order using the mark control set procedure, where the classifier is built having a similar number of yields as the aggregate number of classes.

4 Experimental Analysis

In this segment, we exhibit the analyses that further examine the conduct of the MFCC–Gaussian– KL approach. The essential presumption behind every one of the examinations is that this approach is a timbral remove measure and that accordingly it is expected to perform well at sound grouping. In all trials, we consequently perceive how the chest sound acknowledgement execution is influenced by different changes and twists. To play

DOI: 10.12948/issn14531305/23.1.2019.05

out the tests, we take various media documents that are created with sample chest sounds and change them in multiple approaches to indicate diverse MFCC properties explicitly. To orchestrate wave signals from the media documents, a product synthesizer [8] has been utilized. Different samples were employed for each trial. Can be considered the samples of an average person, and of persons with different types of breathing difficulties including hypertension. We utilize the approach with a solitary Gaussian with full covariance network, since this would be the evident decision in handy applications because of the reasonable computational focal points. The sum total of what tests have been performed with various diverse MFCC requests to perceive how it influences the outcomes.

5 Results

The utilization of KL disparity to therapeutic and wellbeing information is exceptionally testing. The datasets, as a rule, are vast, complex, heterogeneous, progressive, and change in quality. Rahman [12] set the great difficulties of Clinical Decision Support into three broad classifications: Improve the adequacy of Clinical Decision Bolster intercessions, Create new Clinical Decision Support mediations and Disseminate existing Clinical Decision Support information and mediations. However, Rahman [12] distinguishing proof covers minimal about information pre-handling. Without quality information, these three points cannot be figured it out. The lung sound classification performance for proposed statistical features is shown in Table 1.

Table 1. Lung sound classification performance for proposed statistical features [13]

Parameter	Feature	Crackle	Wheeze	Normal	Sensitivity	Specificity	Overall	Time
Mean	LPCC	99.83	95.66	95.00	97.91	95.00	96.80	3.59
	PLPCC	100.00	93.50	91.00	97.25	91.00	94.83	3.70
	LFCC	99.50	95.00	98.30	97.41	98.33	97.61	3.69
	MFCC	99.67	94.50	97.50	97.41	97.50	97.20	3.43
	IMFCC	100.00	89.16	87.67	95.41	87.66	92.27	3.99
Standard Deviation	LPCC	94.00	73.00	73.83	88.58	73.83	80.27	7.02
	PLPCC	98.50	70.60	72.16	84.66	72.16	80.44	31.12
	LFCC	96.60	77.16	83.16	90.25	83.16	85.00	35.31
	MFCC	95.83	76.50	68.30	87.67	68.33	80.22	33.43
	IMFCC	97.66	77.30	71.16	90.00	71.17	82.05	34.99

Can be led tries different things with study proposed, include and also KL disparity approach. Affectability, specificity and general exactness were likewise figured. In all investigations, we let multivariate Gaussian disseminations display the MFCCs from every melody and utilized the symmetrized KL disparity between the Gaussian appropriations as separation measures. Entirely, our outcomes in this way just talk about the MFCCs with this specific separation measure and not of the MFCCs all alone. In any case, we see no conspicuous reasons that different classifiers would perform profoundly extraordinary. In the principal explore, we saw that when keeping as meagre as four coefficients while barring the 0th cepstral coefficient, chest sound grouping exactness was over 80%. We along these lines presume that MFCCs principally catch the otherworldly envelope while experiencing a polyphonic blend of sounds from one source.

The sound-related discernment particular data caught by MFCCs are reasonable for the acknowledgement reason. Then again, benchmark IMFCCs give the most exceedingly terrible execution among the cepstral highlights. The reason behind this might be that determination is more in high-recurrence than that of

DOI: 10.12948/issn14531305/23.1.2019.05

low recurrence in reversed Mel scale. Because of the higher-dividing of channels at bringing down frequencies, this element neglects to catch the discriminative data of the three classes. Also, at times, enhanced information quality is itself the objective of the examination, for the most part, to improve forms in a generation database and the planning of choice help. Numerous different scientists say a few various issues of clinical choice help identified with clinical information furthermore information pre-preparing. The ones most significant to this proposal were a high volume of data, information refresh, and conflicting information portrayal, number of factors, missing/inadequate information, and class unevenness. As there are various issues with restorative information mining, this looks into was restricted to tending to the issues with missing worth, issues with class awkwardness and the impact of class irregularity for include choice. The approach of getting a typical arrangement of p-values at the substance level, as opposed to at the property level, was approved with the multimodal individual acknowledgement issue in the grouping set, and the saliency expectation issue in the relapse setting. The acquired outcomes showed that quintile strategies for consolidating p-values, for example, the Standard Normal Function and the Non-similarity Aggregation strategies gave the best alignment come about, and can be considered to receive a no-SQL system for data combination.

6 Discussion

Acoustic signs delivered by the lungs amid motivation and lapse render helpful data concerning the lung status. In this examination, the competency of the measurable parameters of cepstral includes in perceiving three composes of lung sounds is examined. Regardless of subjective components like sound-related affectability or the doctor's freshness, obtaining and using phantom data identified with lung sounds and in this way isolating diverse lung sounds can be done efficiently. Regular techniques for characterization that utilization the KL divergence strategy or pattern MFCC based method are, most likely, acceptable. Be that as it may cepstral highlights yield great outcomes contrasted with wavelet as well as, the measurable highlights from cepstral coefficient devour significantly less computational overhead in contrast with the gauge cepstral highlights. The proposed highlights are additionally improved, and their execution within sight of three extraordinary added substance clamors are additionally contemplated independently. We have discovered that cepstral highlights are superior to anything wavelet-based highlights as far as grouping exactness. This examination thought about calculations that gauge the many-sided grouping quality of MPR. Two such calculations, Distinctness Index (SI) and Nearest Neighbors Separability (NNS), were found to yield high connection with grouping precision. The utility of these calculations for MPR was exhibited with the high grouping precision generated by the capabilities chose utilizing these two calculations. SI was assessed using diverse separation definitions, from which best execution was accomplished employing an adjusted variant of the KL divergence, which likewise considers the covariance of the neighboring class.

While breaking down the chest sound recording from different sources, as it were two of the three sources were frequently perceived. The number of situations where all instruments were perceived expanded drastically at the point when the sounds were analyzing thus rather than all the while, proposing that the reason is either the log-step when processing the MFCCs, or the marvel that the likelihood thickness elements of an aggregate of arbitrary factors is the convolution of the individual likelihood thickness capacities. From this unmistakably the achievement of the MFCC–Gaussian– KL approach n sort and craftsman order is exceptionally conceivable due as it were to the sound analysis and recognition. This is bolstered that demonstrated that for emblematic sound, the chest sound is extremely vital to kind grouping. We theorize that in type characterization tests, perceiving the two most remarkable sources is sufficient to accomplish worthy execution.

DOI: 10.12948/issn14531305/23.1.2019.05

We have examined the properties of a recorded chest sound measure, in view of the KL, separate between Gaussian models of MFCC highlights. Our investigations appear that the MFCC– Gaussian– KL measure of separation between sounds perceives the analysis which means the variations in the chest sound (since we analyzed three variations of chest sound) does not corrupt acknowledgement precision. However, the chest sound variations of a person with hypertension showed the tendency to stifle the weaker sides and thus clarify the analysis. Moreover, extraordinary acknowledge of the sounds altogether decreases acknowledgement execution. Our comes about propose that the utilization of source partition strategies in a blend with officially existing comparability measures in the sound analysis may prompt expanded order execution.

Abbreviations

ADC	Analog-to-digital converter
ANN	Artificial (manufactured) neural network
CCE	Classification complexity estimates
CCEA	Classification complexity estimating algorithm
EMG	Electromyography
GMM	Gaussian blend display classifier
IM	Individual movements
KL	Kullback-Leibler
k-NN	Euclidean k-Nearest Neighbors
LDA	Linear discriminant analysis
MFCC	Mel-frequency cepstrum coefficient
MLP	Multi-layer perceptron
MPR	Myoelectric pattern recognition
NNS	Nearest neighbour separability
PAP	Pulmonary alveolar phospholipoproteinosis
RBF	Radial Basis Function
RI	Repeatability index
SI	Separability index or distinctness index
SM	Simultaneous movements
SVM	Support (bolster) vector machine

References

[1] Aha, D., Kibler, D., & Albert, M. (1991). “Instance-based learning algorithms.” *Machine Learning*, 6(1). http://dx.doi.org/10.1007/bf00153759

[2] Balasubramanian, V. (2011). “Conformal predictions in multimedia pattern recognition.”, *ACM Sigmultimedia Records*, 3(1). http://dx.doi.org/10.1145/2069196.2069202

[3] Benatti, S., Casamassima, F., Milosevic, B., Farella, E., Schonle, P., & Fateh, S. et al. (2015). “A Versatile Embedded Platform for EMG Acquisition and Gesture Recognition.”, *IEEE Transactions On Biomedical Circuits And Systems*, 9(5). http://dx.doi.org/10.1109/tbcas.2015.2476555

[4] Chowdhury, S., & Majumder, A. (1982). “Frequency analysis of adventitious lung sounds.”, *Journal Of Biomedical Engineering*, 4(4). http://dx.doi.org/10.1016/0141-5425(82)90048-6

[5] Coccagna, G., Pollini, A., & Provini, F. (2006). “Cardiovascular disorders and obstructive sleep apnoea syndrome.”, *Clinical And Experimental Hypertension*.

[6] Dougherty, R., Freiling, C., & Zeger, K. (2006). “Six New Non-Shannon Information Inequalities.” *Information Theory*, 2006 *IEEE International Symposium* on, 233 – 236. https://doi.org/10.1109/ISIT.2006.261840 . 10.1109/ISIT.2006.261840

[7] Harremoës P. (2015). “Lattices with non-Shannon Inequalities”. Cs.IT. http://dx.doi.org/1502.04336v1

[8] Jensen, J., Christensen, M., Ellis, D., & Jensen, S. (2009). “Quantitative Analysis of a Common Audio Similarity Measure.”, *IEEE Transactions On Audio, Speech, And Language Processing*, 17(4). http://dx.doi.org/10.1109/tasl.2008.2012314

[9] Kandaswamy, A., Kumar, C., Ramanathan, R., Jayaraman, S., & Malmurugan, N. (2004). “Neural classification of lung sounds using wavelet coefficients.”, *Computers In Biology And Medicine*, 34(6), 523-537. http://dx.doi.org/10.1016/s0010-4825(03)00092-1

DOI: 10.12948/issn14531305/23.1.2019.05

[10]Nigrin, A. (1993). "Neural networks for pattern recognition". *Cambridge, Mass: Bradford Books/MIT Press*. https://doi.org/10.1016/0893-6080(95)90002-0
[11]Nilsson, N., Håkansson, B., & Ortiz-Catalan, M. (2017). "Classification complexity in myoelectric pattern recognition." *Journal Of Neuroengineering And Rehabilitation*, 14(1). http://dx.doi.org/10.1186/s12984-017-0283-5
[12]Rahman, M. (2014). "Machine Learning Based Data Pre-processing for the Purpose of Medical Data Mining and Decision Support. My learning."
[13]Sengupta, N., Sahidullah, M., & Saha, G. (2016). "Lung sound classification using cepstral-based statistical features.", *Computers In Biology And Medicine*, 75. http://dx.doi.org/10.1016/j.compbiomed.2016.05.013
[14]Shrestha, D., & Solomatine, D. (2006). "Machine learning approaches for estimation of prediction interval for the model output.", *Neural Networks*, 19(2). http://dx.doi.org/10.1016/j.neunet.2006.01.012
[15]Wasito, I., & Mirkin, B. (2005). "Nearest neighbour approach in the least-squares data imputation algorithms." *Information Sciences*, 169(1-2). http://dx.doi.org/10.1016/j.ins.2004.02.014
[16] D. Manivannan, N.K.Manikandan and Ravikumar S, "A Highly Efficient Audio-Visual Recognition System Based On Multiband Feature", *International Journal of Civil Engineering and Technology*, 8(10), 2017, pp. 315–324. [Accessed 12 Dec 2017] http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=8&IType=10
[17] Singh S. "Multiresolution estimates of classification complexity.", IEEE Trans Pattern Anal Mach Intell. 2003; 25:1534–1539. Doi: 10.1109/TPAMI.2003.1251146.
[18] Bunderson NE, Kuiken TA. "Quantification of feature space changes with experience during electromyogram pattern recognition control.", IEEE Trans Neural Syst Rehabil Eng. 2012; 20:239–246. Doi: 10.1109/TNSRE.2011.2182525.
[19] Rui Guedes, Henrique Cyrne, Ana Castro "Heart Sound Analysis for Blood Pressure Estimation", *Encyclopedia of Information Science and Technology*, 4h edition. DOI: 10.4018/978-1-5225-2255-3.ch087
[20] BioPatRec.Available: https://github.com/biopatrec/biopatrec/wiki. [Accessed: 17 Mar 2018]

Antonio CLIM has graduated the Faculty of Food Science and Engineering at the University of Galati in 1996. In 2017 he has been admitted as Ph.D. candidate at the Bucharest University of Economic Studies in the field of Economic Informatics. His research interests include Business Informatics, Computer Networks and Smart Cities. Currently he is Teaching Assistant at Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies. His work focuses on the analysis of machine-learning applications for health services.

DOI: 10.12948/issn14531305/23.1.2019.05

Răzvan Daniel ZOTA has graduated the Faculty of Mathematics – Computer Science Section at the University of Bucharest in 1992. In 2000 he has received the Ph.D. title from the Bucharest University of Economic Studies in the field of Cybernetics and Economic Informatics. His research interests include Business Informatics, Computer Networks and Smart Cities. Currently he is full Professor of Economic Informatics within the Department of Computer Science in Economics at Faculty of Cybernetics, Statistics and Economic Informatics from the Academy of Economic Studies. He is the author of more than 10 books and over 120 articles published in journals and conference proceedings in the field of business informatics, computer networking and e-business.

DOI: 10.12948/issn14531305/23.1.2019.05