# Augmenting the Performance of Multi-patient Parameter Monitoring System in LKSVM

**S. Premanand and S. Sugunavathy**

**Abstract** To offer quality wellness care to patients, multi-parameter patient monitors (MPM) need a high accuracy for sensitivity, specificity, and overall classification. Nevertheless, it is likewise important to provide affordable healthcare by providing cheap MPMs using todays handheld computing and communication devices, and low complexity hardware. Support vector machine (SVM) is a vital classification process valuable for the improvement of MPMs for its high exactness and viability in foreseeing the status of patients. It is well known that non-linear kernel SVMs offer better performance, while the linear kernel SVM (LKSVM) are computationally very efficient. This makes the LKSVM particularly attractive for low cost implementations. In this paper, we demonstrate that mapping feature to a higher dimension using locality-constrained linear coding (LLC), added to that the framework by eluding the system reliant features using dimensionality reduction technique called principal component analysis (PCA) to make the framework durable, which improve the execution of MPMs using LKSVM. It was seen that the use of LLC-PCA has helped enhance the sensitivity by 3.27 % from the baseline system.

**Keywords** Multi-parameter patient monitor · Support vector machine · Locality-constrained linear coding · Principal component analysis

S. Premanand (✉)
Department of Electronics and Communication Engineering, Amrita School of Engineering, Ettimadai, Coimbatore, Tamil Nadu, India
e-mail: newgen.anand@gmail.com

S. Sugunavathy
Department of Electronics and Communication Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi, Tamil Nadu, India



R. Bhramaramba and A.C. Sekhar (eds.), *Application of Computational Intelligence to Biology*, Springer Briefs in Forensic and Medical Bioinformatics,
DOI 10.1007/978-981-10-0391-2_2

## 1 Introduction

MPMs [1] make utilization of the indispensable signs of humans like, respiration rate (RR), heart rate (HR), blood pressure (BP) and oxygen saturation (SPO2) for observing the status of patients in escalated care units and inpatient wards. The utilization of this system in intensive care units (ICU) can recognize, identify the weakness in the patients' well being condition and start opportune intercessions to save lives. For the dependable execution of the MPMs, the likelihood of missing cautions (alarms) and in additional false alerts ought to be least, which implies the alert precision(sensitivity) and no-caution (alarm) accuracy (specificity) of the frame work ought to be similarly prominent as could be allowed.

Machine learning (ML) procedures were broadly utilized for the detecting of indispensable disintegration in patients' wellbeing. SVM is an effective classification technique, strategy which has been utilized for cataloguing purposes. SVM is all around perceived for its speculation capacity and effectiveness of arrangement even with higher measurement (dimensional) information. The proficiency of the framework execution, profoundly relies on upon the instance of the part being utilized as a part of the SVM. At the point when managing with linearly non-detachable information, non-linear kernel SVM (NLKSVM) give improved execution at the cost of expanded computational intricacy and capacity necessity.

One class SVMs [2], we model the "normal" condition of the patient using the normative data, and any deviation from the normal behavioral pattern of the vital parameter is a novelty or alarm condition. The performance of the MPM system could be improved using two class SVMs [3], when sufficient examples from the patient deteriorations are available

On the occasion of the monitor, the investigation of the indispensable parameters demonstrates that they are not directly (linear) separate. This affiliation is measured in SVMs by utilizing kernels that measure the closeness between any two cases. In the event that the samples fit in with same class the similitude ought to be augmented and in the event that they go to diverse classes, the closeness ought to be negligible. In this way, the effectiveness of the SVM classifier essentially relies on upon the decision of a fitting part of the kernel. Kernels capacities are assessed by deciding the similitude between data (information) focuses. On the account of linear kernels, the dot product between the two chosen data points focuses is the measure of closeness though in RBF (radial basis function) kernel, the opposite exponential of the Euclidian separation between the two is the measure of comparability.

In case of image classification, SVMs utilizing spatial pyramid matching [4] (SPM) part has been exceedingly effective especially in NLKSVM case. In SPM, a codebook with M sections is connected to quantize every vector and produces the higher dimensional coded vector C relating to the data vector Y that has a much lower measurement. In the event that hard vector quantization (VQ) is utilized, every code C has one and only non-zero component, while for delicate VQ, a little gathering of components can be nonzero. With a big number of images in the preparation information, for complex order issues, the amount of support vectors could be a couple of yards, and the computational prerequisites can be wholly critical.

Jianchao et al. [5, 6] Proposed sparse coded SPM (ScSPM) supplanting the VQ in the SPM with sparse coding (SC). The upside of the ScSPM is that the NLKSVM backend classifier in the SPM could be supplanted by a LKSVM backend classifier, without bargaining on the execution. The LKSVM lessens the preparation training, quality, and a consistent complexity in testing. SC productively speaks to the information as an in lines combination of an over complete premise set (codebook), in which the quantity of premise is more than that of data measurement of the feature vector [5, 6]. SC's quantization error is a great deal not exactly VQ coding as SPM.

Local coordinate coding (LCC) [7] is an adjustment of SC, which expressly urges the coding to be nearby. Kai et al. [7] clarified hypothetically, that under specific suppositions, locality is more vital than sparsity, as the locality prompts sparsity, while the opposite is not genuine. One primary detriment of the LCC is that it wants to get care of L1-norm optimization issue that is computationally lavish. Wang et al. [8] offered LLC which can be seen as a quick usage of LCC that uses the locality requirement to extend every vector into its neighborhood coordinate framework. What's more, the enhancement issue utilized by LLC has a scientific arrangement. To improve the computational many sided quality further, an approximated LLC technique is proposed in [8] by performing closest neighbor's pursuit to discover the K closest neighbor cells of the VQ centroids to the info vector, and afterward understanding an obliged minimum square fitting issue to discover the code to speak the data vector in the advanced dimensional space.

## 2 Locality-Constrained Linear Coding

Locality-constrained linear coding means a plotting technique (mapping) used to speak to non-linear features to a higher dimensional space to create them linearly distinct [9]. It has been broadly utilized in the application of image classification for its robust execution [8–10]. LLC technique strides can be explained in Fig. 1.

Where X meant for D-dimensional input vectors, B meant for codebook

$$X = [x1, x2, \ldots, xN] \in \mathbb{R}^{DXN}$$

$$B = [b1, b2, \ldots, bM] \in \mathbb{R}^{DXM}$$

$$\min_{C} \sum_{i=1}^{N} ||x_i - Bc_i||^2 + \lambda ||d_i \odot c_i||^2$$

$$s.t\ 1^T c_i = 1, \forall i$$

$$d_i = \exp \frac{dist(x_i, B)}{\sigma}$$

$$dist(x_i, B) = [dist(x_i, b_1), \ldots\ dist(x_i, b_M)]^T$$

$$\min_{C} \sum_{i=1}^{N} ||X_i - \tilde{c}_i B_i||^2$$

$$s.t.,\ 1^T \check{c}_i = 1, \forall \mathrm{i}$$

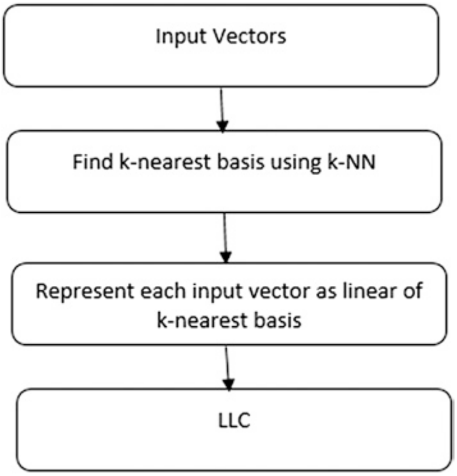



**Fig. 1** Steps in LLC

LLC by fast approximation method lessens the computational complication. At that point we explore PCA with LLC to distinguish the framework-free features for augmenting the execution of the MPMs. We utilize LKSVM as a rear-end classifier.

## 3 Principal Component Analysis

Principal component analysis (PCA) [11–13] is a component change system used to tell apart the contours in the advanced dimensional information and waypoints the similitudes and divergences in the data (Fig. 2).

It is likewise used to diminish the quantity of measurements from the vectors, short of trailing the information. The strides included in the PCA are depicted, Fig. 3. In this work, we utilize PCA for distinguishing the framework-autonomous features over the framework.

## 4 LLC and PCA Method for Robust LKSVM

We take a note of that, with the assistance of LLC, we linearized the features into an advanced dimensional space to make them linearly isolated. Despite the fact LLC enhances the specificity exactness which brings about expansion in general classification precision, affectability precision was not enhanced because of framework reliant information. In this manner, we have to identify and uproot the framework-reliant features to enhance the execution of MPMs. The strategy of LLC-PCA framework for the vigorous MPM framework is appeared in the Fig. 4.

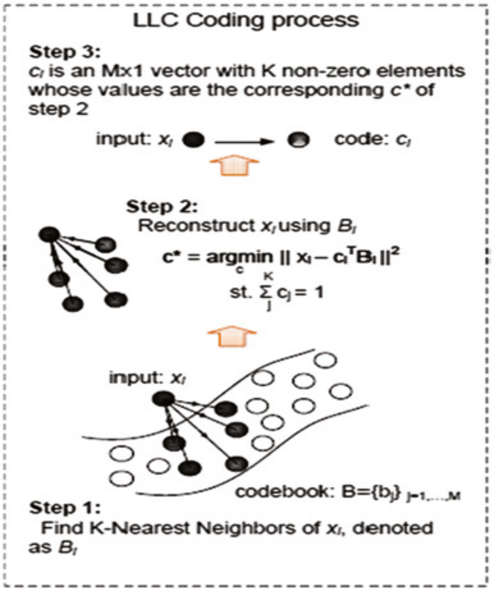



**Fig. 2** LLC coding technique

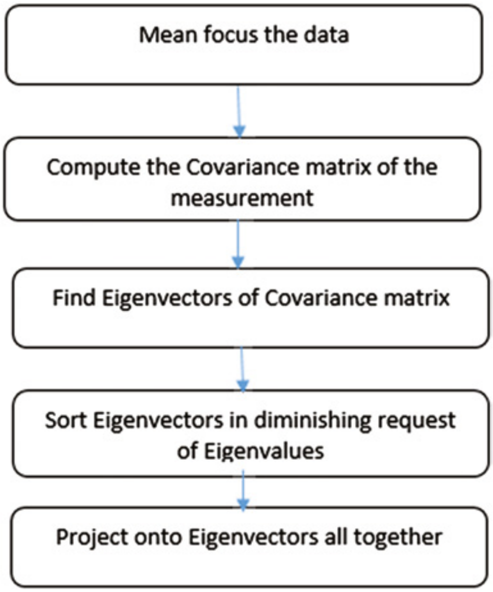



**Fig. 3** Steps in principle component analysis

From our examinations with this technique, we similarly base, few features don't have valuable data subsequent to mapping to a sophisticated dimensional space.

SVMs [14, 15] have produced as a prevalent means to deal with ML, for grouping and regression approach, exhibiting condition of-craftsmanship execution in differing applications and offering an alluring distinct option for counterfeit neural

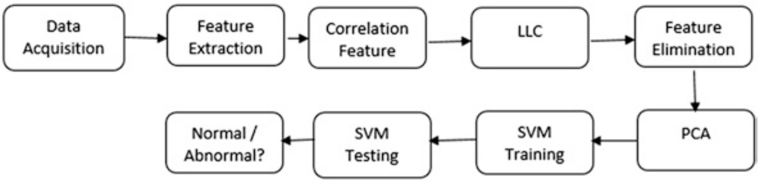



**Fig. 4** Proposed methodology support vector machine

system and master-based methodologies. It builds a hyperplane that isolates two groups indicated in Fig. 5. At the same time, the SVM approach tries to accomplish most extreme partition between the categories. Isolating the classes with an extensive edge minimizes a bound on the normal speculation mistake. A 'minimum generalization error', implies that when new instances touch the base for classification, the likeliness of getting a wrong confidence in the forecast in view of educated classifier ought to be unimportant.

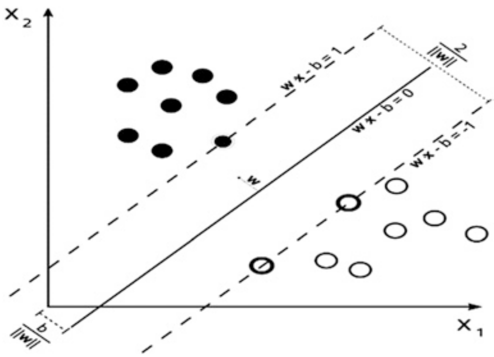
Vapnik and Vapnik [16] has demonstrated that if the preparation vectors (train section) are isolated without mistakes by an ideal hyperplane, the normal blunder rate on a test is restricted by the proportion of the desire of the support vectors (SV) to the quantity of the preparing vectors. Since this proportion is liberated of the measurement of the issue, if one can acquire a humble arrangement of SV, great speculation is ensured. Hyperplane boosts the edge and it can be acquired by deciding the separation between bounding planes to the cause separately (b1, b2) and subtracting the separation (b2–b1), to augment the edge, SVM can be planned as minimization issue and communicated just as equally,

$$\frac{2}{||w||^2} \rightarrow \min_{x,\gamma} \frac{||w||^2}{2}$$

Subject to the constraint:

$$d_i\left[\left(w^T \cdot x_i\right)\right] + \gamma \geq 1$$

**Fig. 5** LKSVM process



The ideal arrangement can be acquired by Lagrangian dual technique, and SVM learning formulation will be,

$$\max_{\propto\geq 0}\ \min_{w,\gamma}\left\{\frac{1}{2}||w||^2-\sum_{i=1}^{n}\alpha_i\left[d_i\left(w^T\cdot x_i-\gamma\right)-1\right]\right\}$$

$\alpha_i$ denotes Lagrangian Multiplier. The issue can be illuminated as, by using KKT condition (Karush-Kuhn-Tucker)

$$\max_{\propto_i}\left\{\sum_{i=1}^{n}\propto_i-\frac{1}{2}\sum_{i,j}\propto_i\propto_j\ d_i d_j k\left(x_i,x_j\right)\right\}$$

Subject to:

$$\propto_i\geq 0,\sum_{i=1}^{n}\propto_i\ d_i=0$$

The decision function can be expressed equally, $\gamma$ can be computed by compelling a point on the edge for which $\alpha_i\neq 0$.

$$f(x)=sign\left(w^T x-\gamma\right)=sign(\sum_{i=1}^{n}\alpha_i d_i x_i^T x-\gamma)$$

## 5 Experiments and Results

We utilized MIMIC-II databank [17, 18] for every one of our tests in this study. MIMIC II databank contains baseline parameters gathered from 413 patients. The database also comprises of a screen produced cautions alongside the doctor explained documents. Among the information from 413 patients, we chose the information from 401 patients for this test based on the tone of the information from the patient. We split the information from 401 patients, as 14,54,010 samples into training from 300 patients and 11,00,510 samples as test sets from 101 patients by using a random permutation method. Subsequently, the entire training data were shuffled randomly and a subset of 50,000 samples with their corresponding labels was selected as the training data for one trial of the experiment, and a subset of 20,000 samples from the test set was selected as the test data for the corresponding trial. Another set of training and test data is selected from the remaining data, to generate the data for another trial. The method is iterated to get tested for seven free trials, and the answers acquired from these seven tests are found the middle value to get the last reply. The size of the data was reduced just for computational considerations, to avoid dealing with large kernel matrices in solving the model parameters. We used LIBSVM toolbox [19] for all our experiments with SVM.

Foremost, we developed an MPM consuming the baseline parameters. Based on previous research results [20], we then used correlation features, geometric mean of two vital parameters required at a time, in addition to the four vital parameters, causing the total number of features to ten. It was determined that the role of correlation features helped enhance the overall classification accuracy to 14.8 %. We preferred this system as our baseline system for comparing the potency of utilizing an LLC to map the features to a higher dimension to improve the operation of the MPM, using the LKSVM backend classifier.

To cause the LLC coding, we first generated a VQ codebook using Linde-Buzo-Gray (LBG) [21, 22] clustering algorithm till about 2048 clusters, only the training information was employed for generating the codebook. We experiment with different k-NN and codebook sizes for better classification performance. In our experiments, we note that for 1024 clusters was found to be the optimum value for the codebook size (M), 19 for the number of nearest neighbors, K.

In this work, features are linearized into advanced dimensional space by utilizing LLC and that has no information with it are wiped out, henceforth it improves no alarm (specificity) condition which enhances overall accuracy for classifier but in the case of sensitivity condition the system is not enhanced because of reliable features. Thus we go for dimensionality reduction technique PCA to recognize and get rid of the framework reliant-features. We tested LLC-PCA technique for distinctive codebook sizes and k-closest neighbors for more reliant working. We exactly found, for 1024 cluster codebook size, with a choice of 180 dimensions using dimensionality technique, the execution has been enhanced when contrasted with the standard framework and LLC for the strong MPM model. For a system should be in robust condition and should be reliable for all kind of classification it not only depends on overall classification but also satisfies both sensitivity and specificity condition also.

Table 1 compares the performance of different MPM systems. It may be noted that the sensitivity increased by 3.27 %, specificity of 0.46 %, and the overall classification accuracy of 0.35 % with the use of an LLC-PCA methodology. We enhanced the affectability precision (sensitivity) for the MPM system framework model to the detriment of specificity exactness. It might be noticed that a superior affectability precision is coveted in basic social insurance applications even to the detriment of a lower specificity exactness.

**Table 1** Comparability of results for enhanced LKSVM for different features

| Input data | D | OA | SEN | SPE |
|---|---|---|---|---|
| Vital parameter | 4 | 77.14 | 1.55 | 100 |
| Vital + Corr. | 10 | 91.94 | 77.09 | 96.43 |
| With LLC | 1024 | 97.67 | 92.57 | 99.21 |
| LLC-PCA | 180 | 98.02 | 95.84 | 99.67 |

*D* Dimension of the feature
*OA* Overall Accuracy
*SEN* Sensitivity
*SPE* Specificity

## 6 Conclusions

It is really important to deliver high sensitivity, specificity, and overall classification accuracy, for MPM to provide quality health care. However, it is also important to provide affordable healthcare, by being able to make the MPM system algorithm using low cost computing and communication devices, and low complexity hardware.

It is surely understood that LKSVM is well known for its computational proficiency, and its suitability to ease execution. In this paper, we explored the use of the LLC technique to linearize the component vectors to a higher dimensional space and after that we chose the framework free elements utilizing dimensional reduction technique PCA to upgrade the execution of the MPM with a LKSVM backend classifier.

## References


1. National Patient Safety Association (2007) Safer care for acutely ill patients: learning from serious accidents. Tech Rep (NPSA)
2. Clifton L, Clifton, DA, Watkinson PJ, Tarassenko L (2011) Identification of patient deterioration in vital-sign data using one-class support vector machines. In: Federated conference on computer science and information systems (Fed-CSIS) (Szczecin, Poland). IEEE, pp 125–131
3. Khalid S, Clifton DA, Clifton L, Tarassenko L (2012) A two-class approach to the detection of physiological deterioration in patient vital signs, with clinical label refinement. In: IEEE transactions on information technology in biomedicine, vol 16, no 6, pp 1231–1238
4. Lazebnik S, Schmid C, Ponce J (2006) Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: IEEE computer society conference on computer vision and pattern recognition
5. Jianchao Y, Yu K, Gong Y, Huang T (2009) Linear spatial pyramid matching using sparse coding for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, pp 1794–1801
6. Lee H, Battle A, Raina R, Ng AY (2007) Efficient sparse coding algorithms. In: Advances in neural information processing systems, vol 19, p 801
7. Kai Y, Zhang T, Gong Y (2009) Nonlinear learning using local coordinate coding. NIPS 9:1
8. Wang J, Yang J, Yu K, Lv F, Huang T, Gong Y (2010) Locality-constrained linear coding for image classification. In: IEEE conference on computer vision and pattern recognition (CVPR), pp 3360–3367
9. Taniguchi K, Han XH, Iwamoto Y, Sasatani S, Chen YW (2012) Image super-resolution based on locality-constrained linear coding. In: 21st IEEE international conference on pattern recognition (ICPR), pp 1948–1951
10. Zhang P, Wee CY, Niethammer M, Shen D, Yap PT (2013) Large deformation image classification using generalized locality-constrained linear coding. In: Medical image computing and computer-assisted intervention (MICCAI 2013), pp 292–299
11. Hyvärinen L (1970) Principal component analysis, mathematical modeling for industrial processes. Springer, Berlin, pp 82–104
12. Moore B (1981) Principle component analysis in linear systems: controllability, observability, and model reduction. In: Automatic Control, IEEE Transactions, pp 17–32, Feb 1981
13. Kim KI (2002) Face recognition using kernel principle component analysis. IEEE Sig Process Lett 9(2):40–42

14. Soman KP, Loganathan R, Ajay V (2009) Machine learning with SVM and other kernel methods. Prentice Hall India Learning Private Ltd., New Delhi
15. Scholkopf B, Smola AJ (2001) Learningwith kernels—support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge
16. Vapnik VN, Vapnik V (1998) Statistical learning theory, vol 2. Wiley, New York
17. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCH, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) Physio Bank, Physio Toolkit and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 101(23):215–220
18. Lee J, Scott D, Villarroel M, Clifford G, Saeed M, Mark R (2011) Open-access MIMIC-II database for intensive care research. In: 33rd annual international conference of the IEEE EMBS. Boston, Massachusetts, USA, pp 8315–8318, Sept 2011
19. Chang C-C, Lin C (2011) LIBSVM: A library for support vector machines. ACM T Intell Syst Technol 2(3):27:1–27:27
20. Tarassenko L, Clifton DA, Pinsky MR, Hravnak MT, Woods JR, Watkinson PJ (2011) Centile-based early warning scores derived from statistical distributions of vital signs. Resuscitation 82(8):1013–1018
21. Gray RM (1984) Vector quantization. IEEE ASSP Magazine, pp 4–29, Apr 1984
22. Linde Y, Buzo A, Gray RM (1980) An algorithm for vector quantizer design. In: IEEE Transactions on Communications, pp 702–710, Jan 1980

http://www.springer.com/978-981-10-0390-5

Application of Computational Intelligence to Biology
Bhramaramba, R.; Sekhar, A.C. (Eds.)
2016, IX, 102 p. 60 illus., 42 illus. in color., Softcover
ISBN: 978-981-10-0390-5