

Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine

Davide Anguita¹, Alessandro Ghio, Luca Oneto, Xavier Parra², and Jorge L. Reyes-Ortiz^{1,2}

¹ DITEN - Università degli Studi di Genova, Genoa I-16145, Italy,
{davide.anguita,alessandro.ghio,luca.oneto}@unige.it

² CETpD - Universitat Politècnica de Catalunya, Vilanova i la Geltrú 08800, Spain,
xavier.parra@upc.edu, jorge.luis.reyes@estudiant.upc.edu

Abstract. Activity-Based Computing [1] aims to capture the state of the user and its environment by exploiting heterogeneous sensors in order to provide adaptation to exogenous computing resources. When these sensors are attached to the subject's body, they permit continuous monitoring of numerous physiological signals. This has appealing use in healthcare applications, e.g. the exploitation of Ambient Intelligence (AmI) in daily activity monitoring for elderly people. In this paper, we present a system for human physical Activity Recognition (AR) using smartphone inertial sensors. As these mobile phones are limited in terms of energy and computing power, we propose a novel hardware-friendly approach for multiclass classification. This method adapts the standard Support Vector Machine (SVM) and exploits fixed-point arithmetic for computational cost reduction. A comparison with the traditional SVM shows a significant improvement in terms of computational costs while maintaining similar accuracy, which can contribute to develop more sustainable systems for AmI.

Keywords: Activity Recognition, SVM, Smartphones, Hardware-Friendly

1 Introduction

Since the appearance of the first commercial hand-held mobile phones in 1979, it has been observed an accelerated growth in the mobile phone market which has reached by 2011 near 80% of the world population [2]. This shows that in a very short time, mobile devices will become easily accessible to virtually everybody. Smartphones, which are a new generation of mobile phones, are now offering many other features such as multitasking and the deployment of a variety of sensors, in addition to the basic telephony. Current efforts attempt to incorporate all these features while maintaining similar battery lifespans and device dimensions.

The integration of these mobile devices in our daily life is rapidly growing. It is envisioned that such devices will seamlessly keep track of our activities, learn

from them, and subsequently help us to make better decisions regarding our future actions [3]. This is one key concepts in which AmI relies on. In this paper, we employ smartphones for human Activity Recognition with potential applications in assisted living technologies. We take into account current hardware limitations and propose a new alternative for AR that requires less computational resources to operate.

AR aims to identify the actions carried out by a person given a set of observations of itself and the surrounding environment. Recognition can be accomplished, for example, by exploiting the information retrieved from inertial sensors such as accelerometers [4]. In some smartphones these sensors are embedded by default and we benefit from this to classify a set of physical activities (*standing, walking, laying, walking, walking upstairs* and *walking downstairs*) by processing inertial body signals through a supervised Machine Learning (ML) algorithm for hardware with limited resources.

This paper is structured in the following way: The state of the art regarding previous work is depicted in Section 2. The description of the adopted methodology is presented in Section 3. There, the experimental set up for capturing the data and the mathematical description for the proposed Multiclass Hardware Friendly Support Vector Machine (MC-HF-SVM) are explained. Experimental results and conclusions of this research are described in Sections 4 and 5.

2 Related Work

The development of AR applications using smartphones has several advantages such as easy device portability without the need for additional fixed equipment, and comfort to the user due to the unobtrusive sensing. This contrasts with other established AR approaches which use specific purpose hardware devices such as in [5] or sensor body networks [6]. Although the use of numerous sensors could improve the performance of a recognition algorithm, it is unrealistic to expect that the general public will use them in their daily activities because of the difficulty and the time required to wear them. One drawback of the smartphone-based approach is that energy and services on the mobile phone are shared with other applications and this become critical in devices with limited resources.

ML methods that have been previously employed for recognition include Naive Bayes, SVMs, Threshold-based and Markov chains [6]. In particular, we make use of SVMs for classification as it was also used in [7] and [8]. Although it is not fully clear which method performs better for AR, SVMs have confirmed successful application in several areas including heterogeneous types of recognition such as handwritten characters [9] and speech [10].

In ML, fixed-point arithmetic models have been previously studied [11, 12] initially because devices with floating-point units were unavailable or expensive. The possibility of retaking these approaches for AmI systems that require either low cost devices or to allow load reduction in multitasking mobile devices has nowadays become particularly appealing. Anguita et al. in [13] introduced the concept of a Hardware-Friendly SVM (HF-SVM). This method exploits fixed-

point arithmetic in the feed-forward phase of the SVM classifier, so as to allow the use of this algorithm in hardware-limited devices. In this paper, we extend this model for multiclass classification.

The SVM algorithm was originally proposed only for binary classification problems but it has been adapted using different schemes for multiclass problems such as in [9]. In particular, we have chosen the One-Vs-All (OVA) method as its accuracy is comparable to other classification methods as demonstrated by Rifkin and Klautau in [14], and because its learned model uses less memory when compared for instance to the One-Vs-One (OVO) method. This is advantageous when used in limited resources hardware devices.

3 Methodology

3.1 Experimental Setup

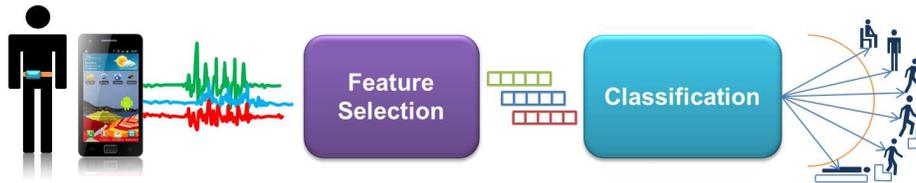


Fig. 1. Activity Recognition process pipeline.

The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed the six activities previously mentioned wearing the smartphone on the waist. The experiments have been video-recorded to facilitate the data labeling. The obtained database has been randomly partitioned into two sets, where 70% of the patterns has been used for training purposes and 30% as test data: the training set is then used to train a multiclass SVM classifier which is described in the following section. A Samsung Galaxy S2 smartphone has been exploited for the experiments, as it contains an accelerometer and a gyroscope for measuring 3-axial linear acceleration and angular velocity respectively at a constant rate of 50Hz, which is sufficient for capturing human body motion.

For AR purposes, we have developed a smartphone application based on the Google Android Operating System. The recognition process starts with the acquisition of the sensor signals, which are subsequently pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap. From each window, a vector of 17 features is obtained by calculating variables from the accelerometer signals in the time and frequency domain (e.g. mean, standard deviation, signal magnitude area, entropy, signal-pair correlation, etc.). Fast Fourier Transform is used for finding the signal frequency

components. Finally, these patterns are used as input of the trained SVM Classifier for the recognition of the activities. The entire AR process pipeline is as shown in Figure 1.

3.2 The Multiclass HF-SVM model

Consider a dataset consisting of l patterns where each one is a pair of the type $(\mathbf{x}_i, y_i) \forall i \in [1, \dots, l]$, $\mathbf{x}_i \in \mathfrak{R}^m$, and $y_i = \pm 1$. A standard binary SVM can be learned by solving a Convex Constrained Quadratic Programming (CCQP) minimization problem which is given by the following formulation [13]:

$$\min_{\boldsymbol{\alpha}} \quad \frac{1}{2} \boldsymbol{\alpha}^T Q \boldsymbol{\alpha} - \mathbf{r}^T \boldsymbol{\alpha} \quad (1)$$

$$0 \leq \alpha_i \leq C \quad \forall i \in [1, \dots, l], \quad (2)$$

$$\mathbf{y}^T \boldsymbol{\alpha} = 0, \quad (3)$$

where C is the regularization parameter, $r_i = 1 \forall i$ and Q is the symmetric positive semidefinite $l \times l$ kernel matrix where $q_{ij} = y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$.

After solving this CCQP problem, the $\alpha_i \forall i \in [1, \dots, l]$ values can be found and used to predict the class of any new pattern using the Feed-Forward Phase (FFP) formulation of the SVM:

$$f(\mathbf{x}) = \sum_{i=1}^l y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b, \quad (4)$$

where b is the bias term and is obtained by using the method proposed in [15]. Clearly, this output is not valid for use in a fixed-point arithmetic approach as these α_i values are by default real numbers ranging between zero and C . Hence a normalization procedure is proposed that will not affect the sign of the classifier output but only its magnitude, maintaining the performance of the SVM as it is known that the class is only determined by the feed-forward function sign. The HF-SVM described in [13] proposes a new vector $\boldsymbol{\beta}$ and it is defined as:

$$\beta_i = \alpha_i \frac{2^k - 1}{C}, \quad (5)$$

where k is the number of bits and $\beta_i \in \mathbb{N}^0$. Also the bias term b of the FFP formulation is removed with the purpose of having a complete integer parameter prediction. The modified formulation is:

$$\min_{\boldsymbol{\beta}} \quad \frac{1}{2} \boldsymbol{\beta}^T Q \boldsymbol{\beta} - \mathbf{s}^T \boldsymbol{\beta} \quad (6)$$

$$0 \leq \beta_i \leq \frac{2^k - 1}{C} \quad \forall i \in [1, \dots, l], \quad (7)$$

where $s_i = (2^k - 1)/C \quad \forall i \in [1, \dots, l]$. Note that the cost function keeps unchanged but Eq. 3 disappears. This is consistent because we use an RBF kernel such as the Laplacian that has infinite VC dimension and the bias b becomes unnecessary [16].

Lastly, to hold true the assumption of having a FFP with only integer values, the kernel $K(\cdot, \cdot)$ and the input vector \mathbf{x} are also represented with u and v bits respectively [13]:

$$0 \leq K(\mathbf{x}_i, \mathbf{x}) \leq 1 - 2^{-u} \quad \forall i \in [1, \dots, l], \quad (8)$$

$$0 \leq x_i \leq 1 - 2^{-v} \quad \forall i \in [1, \dots, m]. \quad (9)$$

The modified FFP formulation with the β vector is:

$$f(\mathbf{x}) = \sum_{i=1}^l y_i \beta_i K(\mathbf{x}_i, \mathbf{x}). \quad (10)$$

We opted for a Laplacian kernel, instead of the more conventional Gaussian kernel, as it is more convenient for hardware limited devices because it can be easily computed using shifters:

$$K(\mathbf{x}_i, \mathbf{x}_j) = 2^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|_1}, \quad (11)$$

where $\gamma > 0$ is the kernel hyperparameter and the norm is expressed as $\|\mathbf{x}\|_1 = \sum_{i=1}^m |x_i|$.

The complete learning process for each SVM consists of performing grid search model selection of the C and γ hyperparameters that converge with the minimum validation error. A k -fold cross validation with $k = 10$ is employed for each hyperparameter pair.

The output of the FFP varies depending on each learned SVM model as these are not normalized. Our extension of this binary problem for the multiclass case employs the OVA method in which each class c is compared against the other classes. This evidently requires a method to allow comparability between the output of each SVM. For this reason, we have opted to compute probability estimates for each SVM $p_c(\mathbf{x})$ and choose the one with the highest probability as the actual class c^* of each test pattern.

We have developed the following approach using the J.Platt's method for estimating probability estimates [17]. The training dataset and the learned SVM model are employed to fit the output of the FFP $f(\mathbf{x})$ with a sigmoid function of the form:

$$p(\mathbf{x}) = \frac{1}{1 + e^{(Af(\mathbf{x})+B)}}, \quad (12)$$

in which $p(\mathbf{x})$ is the probability estimate, and A and B are function parameters which are properly fitted on the available learning samples.

Taking into account that we have the fixed-point arithmetic restriction, the sigmoid function cannot be directly applied on $f(\mathbf{x})$. To solve this, we have

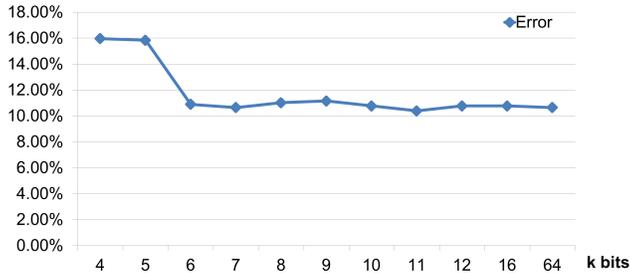


Fig. 2. Comparison between the MC-SVM and the MC-HF-SVM: Classification test error for the MC-HF-SVM with different values of k against the MC-SVM which is represented with $k = 64$ bits.

designed a method based on Look-Up-Tables (LUTs). By defining a fixed number of bits t , it is possible to map the probability estimates $p(\mathbf{x})$ given $f(\mathbf{x})$ without requiring floating-point arithmetic. It has been observed that $t = 8$ is suitable for this application and it only requires a LUT with 256 elements.

4 Experimental results

For evaluating the performance of the MC-HF-SVM, a set of experiments were carried out using the AR dataset described in this paper. They consisted of learning SVM models with different number of bits k for β estimation and then comparing their performance in terms of test data error against the standard floating-point Multiclass SVM (MC-SVM). The results of this comparison are depicted in Figure 2.

The experiment shows that for this dataset $k = 6$ bits are sufficient for achieving a performance comparable with the MC-SVM approach that uses 64-bit floating-point arithmetic. The test error remains stable (around 1% variation) for k values from 64 to 6 bits, but it increases noticeably to 15% when it reaches 5 bits. Moreover, it is also seen from the graph that some values of k produced smaller errors than the one obtained with the MC-SVM. This finding coincides with that in [18] and then in [19]. Thus, suggesting that these methods can exhibit a lower generalization error. We believe that the truncation of the model parameters could be producing a regularization effect.

The classification results of the MC-SVM and the MC-HF-SVM with $k = 8$ bits for the test data are depicted by means of a confusion matrix in Table 1, where estimates of the overall accuracy, recall and precision are also given. 789 test samples were evaluated with approximately equal number of samples per class. Both confusion matrices show similar outputs varying slightly in the classification accuracy of the activities *walking downstairs* and *walking upstairs*. They also expose some false predictions mostly in the dynamic activities. Static activ-

Method	MC-SVM							MC-HF-SVM $k = 8$ bits						
	Walking	Upstairs	Downstairs	Standing	Sitting	Laying	Recall %	Walking	Upstairs	Downstairs	Standing	Sitting	Laying	Recall %
Walking	109	0	5	0	0	0	95.6	109	2	3	0	0	0	95.6
Upstairs	1	95	40	0	0	0	69.8	1	98	37	0	0	0	72.1
Downstairs	15	9	119	0	0	0	83.2	15	14	114	0	0	0	79.7
Standing	0	5	0	132	5	0	93.0	0	5	0	131	6	0	92.2
Sitting	0	0	0	4	108	0	96.4	0	1	0	3	108	0	96.4
Laying	0	0	0	0	0	142	100	0	0	0	0	0	142	100
Precision %	87.2	87.2	72.6	97.1	95.6	100	89.3	87.2	81.7	74.0	97.8	94.7	100	89.0

Table 1. Confusion Matrix of the classification results on the test data using the traditional floating-point MC-SVM (Left) and the MC-HF-SVM with $k = 8$ bits (Right). Rows represent the actual class and columns the predicted class. The diagonal entries (in bold) show the number of test samples correctly classified.

ities instead perform better, particularly the *laying* activity which obtained an accuracy of 100%.

5 Conclusions

In this paper, we proposed a new method for building a multiclass SVM using integer parameters. The MC-HF-SVM is an appealing approach for use in AmI systems for healthcare applications such as activity monitoring on smartphones. This alternative that employs fixed-point calculations, can be used for AR because it requires less memory, processor time and power consumption. Moreover, it provides accuracy levels comparable to traditional approaches such as the MC-SVM that uses floating-point arithmetic.

The experimental results confirm that even with a reduction of bits equal to 6 for representing the learned MC-HF-SVM model parameter β , it is possible to substitute the standard MC-SVM. This outcome brings positive implications for smartphones because it could help to release system resources and reduce energy consumption. Future work will present a publicly available AR dataset to allow other researchers to test and compare different learning models.

Acknowledgments. This work was supported in part by the Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environments, which is funded by the EACEA Agency of the European Commission under EMJD ICE FPA n 2010-0012.

References

1. Davies, N., Siewiorek, D.P., Sukthankar, R.: Activity-based computing. *IEEE Pervasive Computing* **7**(2) (April 2008) 20–21

2. Ekholm, J., Fabre, S.: Forecast: Mobile data traffic and revenue, worldwide, 2010-2015. In: Gartner Mobile Communications Worldwide. (July 2011)
3. Cook, D.J., Das, S.K.: Pervasive computing at scale: Transforming the state of the art. *Pervasive and Mobile Computing* **8**(1) (February 2012) 22–35
4. Allen, F.R., Ambikairajah, E., Lovell, N.H., Celler, B.G.: Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models. *Physiological Measurement* **27**(10) (2006) 935
5. Rodríguez-Molinero, A., Pérez-Martínez, D., Samá, A., Sanz, P., Calopa, M., Gálvez, C., Pérez-López, C., Romagosa, J., Catalá, A.: Detection of gait parameters, bradykinesia and falls in patients with parkinson’s disease by using a unique triaxial accelerometer. *World Parkinson Congress, Glasgow* (2007)
6. Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **10**(2) (2010) 1154–1175
7. Ravi, N., D, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: *In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence (IAAI, AAAI Press* (2005) 1541–1546
8. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl.* **12**(2) (March 2011) 74–82
9. LeCun, Y., Jackel, L., Bottou, L., Brunot, A., Cortes, C., Denker, J., Drucker, H., Guyon, I., Mller, U., Sckinger, E., Simard, P., Vapnik, V.: Comparison of learning algorithms for handwritten digit recognition. In: *International Conference on Artificial Neural Networks.* (1995) 53–60
10. Ganapathiraju, A., Hamaker, J., Picone, J.: Applications of support vector machines to speech recognition. *Signal Processing, IEEE Transactions on* **52**(8) (aug. 2004) 2348 – 2355
11. Wawrzynek, J., Asanovic, K., Morgan, N., Member, S.: The design of a neuro-microprocessor. *VLSI for Neural Networks and Artificial Intelligence* **4** (1993) 103–7
12. Anguita, D., Gomes, B.A.: Mixing floating- and fixed-point formats for neural network learning on neuroprocessors. *Microprocess. Microprogram.* **41**(10) (May 1996) 757–769
13. Anguita, D., Ghio, A., Pischiutta, S., Ridella, S.: A hardware-friendly support vector machine for embedded automotive applications. In: *Neural Networks, 2007. IJCNN 2007. International Joint Conference on.* (aug. 2007) 1360 –1364
14. Rifkin, R., Klautau, A.: In defense of one-vs-all classification. *Journal of Machine Learning Research* **5** (2004) 101–141
15. Keerthi, S.S., Shevade, S.K., Bhattacharyya, C., Murthy, K.R.K.: Improvements to platt’s smo algorithm for svm classifier design. *Neural Comput.* **13**(3) (March 2001) 637–649
16. Vapnik, V.N.: *The nature of statistical learning theory.* Springer-Verlag New York, Inc., New York, NY, USA (1995)
17. Platt, J.C.: Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In: *Advances in Large Margin Classifiers,* MIT Press (1999) 61–74
18. Anguita, D., Sterpi, D.: Nature inspiration for support vector machines. In: *Proceedings of the 10th international conference on Knowledge-Based Intelligent Information and Engineering Systems - Volume Part II. KES’06, Berlin, Heidelberg, Springer-Verlag* (2006) 442–449
19. Neven, H., Denchev, V.S., Rose, G., Macready, W.G.: Training a binary classifier with the quantum adiabatic algorithm. *Arxiv preprint arXiv08110416* (x) (2008)