The Winning Solution to the AAIA'15 Data Mining Competition: Tagging Firefighter Activities at a Fire Scene

Jan Lasek
Interdisciplinary PhD Studies Program,
Institute of Computer Science,
Polish Academy of Sciences,
ul. Jana Kazimierza 5, 01-248 Warsaw, Poland
Email: janek.lasek@gmail.com

Marek Gagolewski
Systems Research Institute,
Polish Academy of Sciences,
ul. Newelska 6, 01-447 Warsaw, Poland
and
Faculty of Mathematics and Information Science,
Warsaw University of Technology,
ul. Koszykowa 75, 00-662 Warsaw, Poland
Email: gagolews@ibspan.waw.pl

***Abstract*—Multi-sensor based classification of professionals' activities plays a key role in ensuring the success of an his/her goals. In this paper we present the winning solution to the *AAIA'15 Tagging Firefighter Activities at a Fire Scene* data mining competition. The approach is based on a Random Forest classifier trained on an input data set with almost 5000 features describing the underlying time series of sensory data.**

I. Introduction

HUMAN activity recognition based on sensor inputs, cf., e.g., [1], [7], [14], is essential in many practical applications. In particular, a fire scene constitutes a dynamic environment in which valid, precise, and fast human decisions play a key role. Here, the aim is to achieve success in an emergency rescue mission, having in mind safety of the involved firemen [4] and his/her ability to save other peoples' lives and – in the second place – property, wealth, etc. It is worth noting that an automated decision support system may be used to increase the widely-conceived quality of an agents' behavior. One of its most fundamental components relies on a proper detection of an action a fireman is actually performing at a given moment. The topic of *AAIA 2015 Data Mining Competition: Tagging firefighters' activities at a fire scene* [8] aimed to deliver accurate model for recognising firefighters movements and activities based on multi-sensor data. In consecutive sections we explain the winning approach in very detail. The proposed solution was implemented in the `R` environment for statistical computing [11]. The solution is available on–line as a Git repository at https://github.com/janekl/AAIA15_Data_Mining_Contest.

The paper is organized as follows. In the section to follow, we describe the analyzed data set and define the evaluation metric used. In Section III we discuss main challenges that the data set brought. In Section IV we present the winning solution in detail and indicate its advantages, limitations, and possible extensions for future work. Finally, Section V concludes the paper.

II. Problem statement

The main purpose here is to design a model for a classification problem with two class attributes. The first class denotes the main activity of a firefighter. This class is referred to as *posture* class and it has 5 distinct labels (`crawling`, `crouching`, `moving`, `standing` and `stooping`). The second class, called *action* class denotes a particular activity of a firefighter and consists of 16 labels (4 labels associated with movement along ladder or stairs: `ladder_down`, `ladder_up`, `stairs_down`, `stairs_up`, 2 labels regarding forward movement: `walking` and `running`, labels describing firefighters' operational movements: `manipulating`, `nozzle_usage`, `signal_hose_pullback`, `signal_water_first`, `signal_water_main`, `signal_water_stop`, `striking` and `throwing_hose` and a `no_action` label).

The evaluation metric employed in the competition is the weighted average of *balanced accuracy* for the two classes. Below we recall the definition of this measure. For each label l_i within a class attribute we define *classification accuracy* as

$$\mathrm{acc}(l_i) = \frac{|\{j : l(x_j) = l_i \wedge p(x_j) = l_i\}|}{|\{j : l(x_j) = l_i\}|},$$

where $l(x_j) = l_i$ denotes the true label for instance x_j and $p(x_j)$ denotes the label assigned by a classifier. If a class attribute C assumes L possible labels, then the balanced accuracy score for that class is defined as

$$\mathrm{BAC}(C) = \frac{1}{L}\sum_{i=1}^{L} \mathrm{acc}(l_i).$$

Now, we may consider the weighted average of balanced accuracy scores for *posture* and *action* classes, which is given by

$$\mathrm{EvaluationMetric} = \frac{1}{3}\mathrm{BAC}(\texttt{posture}) + \frac{2}{3}\mathrm{BAC}(\texttt{action}).$$

Please cite this paper as: Lasek J., Gagolewski M., The winning solution to the AAIA'15 Data Mining Competition: Tagging firefighter activities at a fire scene, In: Ganzha M., Maciaszek L., Paprzycki M. (Eds.), Proc. FedCSIS'15, IEEE, 2015, pp. 375–380. doi:10.15439/2015F418

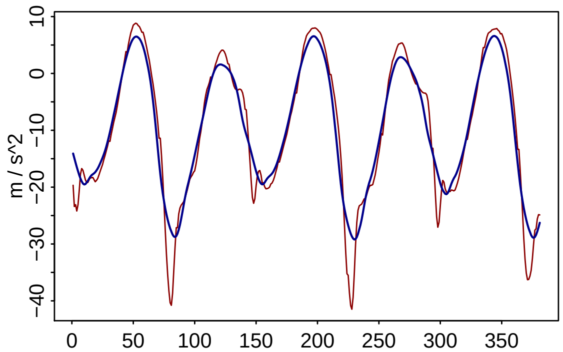

Fig. 1. Plot of the raw series (red) along x-axis of the accelerometer recordings at right hand for pair of labels (`moving`, `running`) and the smoothed series with 20-moving average filter (blue).

During the competition, the solutions were evaluated against approximately 10% of test data. An evaluation metric is the essence of a contest for both its organizers and participants. Through design of an evaluation metric, the organizers define their goal that they want to achieve. On the other hand, the participants need to tailor their models to optimize a given evaluation metric.

To train a statistical model, a training set consisting of 20.000 instances, each tagged with a pair of labels for *posture* and *action* class, was used. Each instance consists of basic statistics on the vital functions of a firefighter and a set of 42 time series which came from x/y/z–axis recordings from gyroscopes and accelerometers attached at 7 points on the body of a firefighter (left hand, right hand, left arm, right arm, left leg, right leg, torso). Each time series consists of 400 recordings (every 4-5 ms) over ca. 2 seconds. Test set consists of 20.000 instances as well. The goal was to develop a model for tagging instances in the test set with a pair of labels for the two class attributes. Both the training and test data set are of size approximately 2.4 GB (uncompressed *csv* files).

III. Main challenges

A. The same action, different results

Among one of the many challenges we find that the data set was inherently noisy. Moreover, the samples of activities in the training and test sets were due to activities of different firefighters. We observed that this had a significant impact on the classifier's score: our scores in terms of the evaluation metric were as high as 98% on a hold-out validation set. This is a considerably high score bearing in mind that the given classification problem is presumably not an easy task. However, in related studies as high accuracy scores were reported [10], [5]. The scores on the official leader-board were significantly lower – with the best scores being equal to ca. 85% during preliminary evaluation. The fact that an instance may come from different source is a great challenge in any application domain.

B. Imbalance of labels distribution

Another problem which required proper handling was related to the imbalance of labels proportion within each class attribute. Table I presents the pair of labels within *action* and *posture* for the test set.

Since the evaluation metric in the competition was the discussed balanced accuracy score, no label was distinguished and misclassification rate has equal weight for every label within each class. This metric treats each label within a class as being equally important (equal weights), regardless of its *a priori* distribution in the data. This distribution of labels varied significantly on the training set. For example, only about 0.5% of all instances constitute for the `signal_hose_pullback` label while about 32% for the `manipulating` label within the *action* class. This means that we are given over 60 times more instances having the former label. Such an uneven distribution of labels requires proper handling by a model. To overcome the problem of imbalanced label distribution, we trained individual classifiers in an ensemble (to be precise, using the below-discussed Random Forest method) based on a stratified subsamples of training set in which each label was represented in an equal amount. The proper balancing of the training set enabled to tailor a model for the evaluation metric employed in the competition.

IV. The winning solution

Let us describe the implemented approach towards feature extraction and model building for activity tagging problem. The model used was based on the Random Forest classifier which is an ensemble of decision trees. It is observed that in practical situations it often yields high accuracy scores [2], [6]. Another advantage of the Random Forest classifier is that it is a fast method: its training and prediction phase can be parallelized. Is is also relatively easy to handle (i.e., its parameter setup) as compared to other advanced ensemble methods. Both the described below feature extraction and the final model training procedures (included in the GitHub repository) can be performed on a single machine within a couple of hours. In our computations we used a 4–core 2.0 GHz CPU 16 GB RAM machine. The described parameter optimization steps were performed on a cluster of 10 8–core 3.40 GHz CPU 16 GB RAM machines to speed up the computations.

A. Feature extraction

Our approach was particularly focused on the phase dealing with features' extraction. The extracted features were based on literature [5], [9], [10] as well as the authors' experimental ideas. The processed training/test dataset is of size about 1.4 GB. For each activity we derived over 4700 features describing a particular activity. First of all, we filtered the data with a moving average window of size 20, see Figure 1. Since the sensor recordings were gathered at a 4.5 ms. resolution, this roughly corresponds to averaging the arriving over a window of 0.1 second. This step was not crucial for the model performance, however, it allowed to filter out the noise slightly.

Please cite this paper as: Lasek J., Gagolewski M., The winning solution to the AAIA'15 Data Mining Competition: Tagging firefighter activities at a fire scene, In: Ganzha M., Maciaszek L., Paprzycki M. (Eds.), Proc. FedCSIS'15, IEEE, 2015, pp. 375–380. doi:10.15439/2015F418

TABLE I
COUNTS FOR PAIRS OF LABELS FOR THE TWO CLASSES – TRAINING SET.

	crawling	crouching	moving	standing	stooping
ladder_down	0	0	465	0	0
ladder_up	0	0	476	0	0
manipulating	0	1764	331	2356	1898
no_action	0	87	0	491	0
nozzle_usage	0	492	0	443	0
running	0	0	4324	0	0
searching	459	0	0	0	0
signal_hose_pullback	0	0	0	98	0
signal_water_first	0	0	41	496	0
signal_water_main	0	46	0	405	0
signal_water_stop	0	0	0	277	0
stairs_down	0	0	644	0	0
stairs_up	0	0	1157	0	0
striking	0	0	0	1022	0
throwing_hose	0	0	0	234	930
walking	0	0	1064	0	0

Next, for each of the time series, we derived basic summary statistics: quantiles (denoted with qx in Tables III and IV for $\mathtt{x} \in \{0.01, 0.05, 0.1, 0.2, \ldots, 0.9, 0.95, 0.99\}$), standard deviation (sd), skewness, kurtosis, amplitude (defined as the difference between 0.99-quantile and 0.01-quantile of the series), the signal energy (ener; defined as the sum of squares of consecutive recordings), the ratio between its maximal absolute value and the median and minimal and maximal of the first differences of a series (deriv1min and deriv1max). We extracted a set of quantiles and standard deviations on the time series processed by the Fast Fourier Transform, to its real, imaginary and modulus (ModFFT) parts independently. Additionally, we recorded first 5 Fourier coefficients of the real and imaginary part of the transformed series. We also extracted quantiles and standard deviation of the periodogram (Period) of each time series. Further, for each pair of time series we computed the linear correlation coefficients (cor).

We also extracted several experimental features for counting the number of peaks in the series based on their sub-chunks in which they exceeded the mean by one or two standard deviations. We imposed a constraint that the minimal length of a sub-chunk is 5 (for the filtered series). Finally, we counted the number of times a given series crosses 0 and its mean.

Another property of Random Forest model is that it has an inherent method of evaluation of feature relevance. Tables III and IV present the 50 most important features for two individual classification tasks for the two class attributes. The criterion of our choice according to which features are evaluated is the mean decrease in Gini Impurity Index for classification (column M.D.Gini). As far as the vital functions are concerned, median respiratory rate reading med.rr is present in the top 50 list for *action* classification problem.

Let us note that the number of features derived is large and some of them do not posses a clear interpretation. However, due to Random Forest model described in the next section – which includes an inherent method of selecting relevant attributes – we were able to handle and select relevant content from this rich set of features.

B. Classification model

For the purpose of tagging the activities we used the balanced Random Forest [2]. By the balanced Random Forest classifier we mean an ensemble of trees that are trained on subsamples of training set in which every label within a given class is represented in an equal amount. This model was used in a stepwise approach. In the first step, we trained the model which aimed to recognize the *posture* of a firefighter. In the second one, we trained the model to recognise the main *action* of a fireman, given *posture* class attribute. This approach is analogous to the classifier chaining method in multilabel classification tasks [12]. In the tagging phase for new data we plug the predicted *posture* labels by the first model as an input for the model for the *action* class. The combined predictions complete the tagging phase for test set.

The idea behind such a chaining method was driven by the fact that some combinations of activities and posture labels are mutually exclusive. For example, the posture cannot be equal to standing when the main activity is equal to ladder_up. When individual classifiers were trained, such inconsistencies were very common. We managed to reduce them by employing the mentioned stepwise approach. However, we did not succeed in eliminating them at all: our final submission still contained some fraction of prediction labels that were mutually exclusive. By mutually exclusive pairs of labels we mean such a combination of pairs of labels that were not observed in the training set (see Table II; these pairs are given in bold). Another way of reducing conflicts was to aggregate different submissions by, e.g., majority voting. We observed that aggregating individual submissions often produced a new submission with a higher preliminary evaluation score than each of the individual ones. This serves as a method for providing more stable and accurate predictions since, e.g., they are based on a larger number of trees.

We also experimented with one-vs-all and single class (i.e., a single class was obtained by mapping each pair of labels within (*posture*, *action*) classes to an individual class) versions of the model. However, the best results were achieved by

TABLE II
COUNTS FOR PAIRS OF PREDICTED LABELS FOR THE TWO CLASSES – TEST SET.

	crawling	crouching	moving	standing	stooping
ladder_down	0	**1**	459	**209**	0
ladder_up	0	**2**	452	**118**	0
manipulating	0	1576	12	1639	2438
no_action	0	71	0	467	**31**
nozzle_usage	0	454	0	1060	0
running	0	**13**	3974	0	**2**
searching	513	**42**	0	0	0
signal_hose_pullback	0	0	0	96	0
signal_water_first	0	**3**	10	580	0
signal_water_main	0	55	0	174	0
stairs_down	0	0	533	0	0
stairs_up	0	0	1442	0	0
striking	0	**13**	**7**	1026	**49**
throwing_hose	0	0	0	196	982
walking	0	**2**	1251	**46**	**2**

the described chaining method.

C. Parameter tuning

Due to the discussed issue of performing activities by different people, the model's parameter tuning process poses a real challenge. We were primarily interested in the parameters responsible for balancing the classifier (parameter `sampsize` in `R`'s `randomForest` package), the minimal number of instances in the leafs (parameter `nodesize`) and the number of sampled features to perform a test split (parameter `mtry`). In our methodology we experimentally set parameters by monitoring out–of–bag error accuracy estimates for each of the classes. Additionally, we run 3–fold cross validation for different pairs of parameters (`mtry`, `nodesize`). Our conclusions was that the parameter `nodesize` should be set to 1, i.e., the trees should be grown to maximal depth. Moreover, given `nodesize = 1`, setting parameter `mtry` to a couple of hundreds already provided stable and high accuracy scores. Finally, to balance training sets, in our initial trials we sampled more instances of most represented labels, i.e., `moving` within *posture* class and `manipulating` and `running` within *action* class as indicated by lower out–of–bag error estimates for those labels. However, significantly better preliminary scores were obtained just by sampling each of the labels in an equal amount. Although choosing parameter values based on leader–board score is quite a dangerous way of tuning them, we took this risk for those parameters as the test set instances differ significantly from the training set ones. In any case, sampling each of the labels equally appears to be a reasonable setup. The described sampling procedure was a crucial step of achieving high evaluation scores. Another advantage of this methodology is that the models are training using fewer instances from the training set. This may in turn prove useful to a reduce overfitting of the model to the training data. The number of trees in the forest was set to 700, i.e., a relatively large number accounting for the computation time of the model.

D. Final submission

During the competition, we submitted over 100 proposals, which were based on different ideas and changes in model parameters and enrichment of the training data with new features. Most of them relied on experiments with different setup of Random Forest model, but we also tried the Gradient Boosting Machine (GBM) model (tree-based) [3], [13]. However, the performance of a less complex Random Forest model was satisfactory and we devoted more time for optimising this model. Moreover, we primarily focused on the feature extraction step.

The best performing model consists of 700 trees, it has the number of attributes for performing test split equal to 300, stratified over class attribute with sample size of 400 for each *posture* and 90 for *action* label. The final submission was derived by majority voting of three classifiers (in fact, two–stage classifiers) with weights 1.5 (to avoid ties), 1, 1 respectively:

1) Random forest model with the minimum size a leaf in a single tree equal to 1 (attribute `nodesize = 1`)
2) Random forest model with `nodesize = 3` and
3) Random forest model with `nodesize = 1` trained on the dataset with exclusion of features associated with left arm of sensory data and a subset of quantiles $(0.01, 0.05, 0.2, 0.4, 0.6, 0.8, 0.95, 0.99)$.

The preliminary evaluation scores were 0.858, 0.8573, and 0.8567 for consecutive models. The averaging step was aimed to reduce the variance of a single model as well as to resolve the mentioned conflicts due to contradictory labels. We used this method as we concluded by our previous experimentation with averaging that it yields higher evaluation scores. However, in case of our final submission, it yielded a not significantly lower preliminary evaluation score than the best one used for aggregating. In any case, we believed that it would produce more stable and accurate predictions on the whole test set. The third model was trained on a subset of attributes since we observed that some recordings in the test set have constant values of these which we interpret as missing values. We also excluded a part of quantiles with the aim to reduce overfitting of the model to the training sample (however, this appeared not to be of help in this case). The final submission yielded score of 0.8577 during preliminary evaluation and 0.8391 on the whole test set – it was ranked the first among 79 submitted proposals.

E. Unsolved puzzles a.k.a. future work

By the end of the challenge, we were still left with some unsolved problems that became evident after the solutions were submitted.

First of all, our final submission still contained some mutually exclusive pairs of labels e.g. `ladder_down` and `standing` or `walking` and `standing`. This problem was limited to some extent by the two–stage classification as well as submissions averaging.

The other problem with our submissions was that our model never predicted the activity `signal_water_stop`. Perhaps feeding the classifier with more instances with this particular label could resolve this issue. This could also possibly apply to `signal_hose_pullback` label within *action* class as there where merely 98 instances tagged with this activity. Finally, as we already mentioned, our preliminary evaluation scores based on out–of–bag predictions from the Random Forest model were overly optimistic: the scores on training data were about 98% of the evaluation metric while the evaluation of our solution on the whole training data yielded much lower score of about 84%. This issue could be addressed by, e.g., performing evaluation and optimisation of a model via cross-validation, where the validation folds would contain activities performed by different firefighters. However, this could not be performed as the information on, e.g., firefighters identifiers performing a given action was not made available to the participants. Another possibility would be to derive more robust features with better generalisation properties for different people performing the same activities.

V. Conclusions

The *AAIA'15 Data Mining Competition: Tagging Firefighter Activities at a Fire Scene* contest was a very interesting and absorbing event. Taking part in such competitions requires some persistence as only few tested ideas prove to give an improvement for the classification score. It is enough to mention that our winning solution was submitted within 24 hours of the competition's deadline.

In our approach, we employed Random Forest classifier and spend much more time on pre-processing data and engineering new features. We believe that the key to success were good data. Using a more sophisticated model may constitute for improvement in classification, however, as in this competition raw time series data needed to be processed, we regarded feature engineering as a more important step. Moreover, proper balancing of training sample provided major gains in the evaluation metric employed in this competition.

The code for our submission is available at GitHub. Further enhancements of the proposed solution are possible. We hope that it will serve as a benchmark for even better performing models for the task of tagging activities at a fire scene.

Acknowledgments

Jan Lasek would like to acknowledge the support by the European Union from resources of the European Social Fund, Project PO KL "Information technologies: Research and their interdisciplinary applications", agreement UDA-POKL.04.01.01-00-051/10-00 via the Interdisciplinary PhD Studies Program.

References

[1] Atallah, L., Lo, B., Ali, R., King, R., Yang, G.-Z. 2009. Real-time activity classification using ambient and wearable sensors, IEEE Transactions on Information Technology in Biomedicine 13:1031–1039, http://dx.doi.org/10.1109/TITB.2009.2028575.

[2] Breiman, L. 2001. Random forests, Machine Learning 45:5–32, http://dx.doi.org/10.1023/A:1010933404324.

[3] Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine, Annals of Statistics 29:1189–1232, http://dx.doi.org/10.1214/aos/1013203451.

[4] Krasuski A. 2014. A framework for Dynamic Analytical Risk Management at the emergency scene. From tribal to top down in the risk management maturity model. Proc. Federated Conference on Computer Science and Information Systems (FedCSIS'14), IEEE, pp. 323–330, http://dx.dox.org/10.15439/2014F371.

[5] Leutheuser H., Schuldhaus D., Eskofier B.M. 2013. Hierarchical, multi-sensor based classification of daily life activities: Comparison with state-of-the-art algorithms using a benchmark dataset, PLoS ONE 8:e75196, http://dx.doi.org/10.1371/journal.pone.0075196.

[6] Liaw, A., Wiener M. 2002. Classification and Regression by randomForest. R News 2:18–22, urlhttp://cran.r-project.org/doc/Rnews/.

[7] Maurer, U., Smailagic, A., Siewiorek, D.P., Deisher, M. 2006. Activity recognition and monitoring using multiple sensors on different body positions, Proc. International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06), IEEE, pp. 113–116, http://dx.doi.org/10.1109/BSN.2006.6.

[8] Meina, M., Janusz, A., Rykaczewski, K., Ślęzak D. Celmer, B., Krasuski, A. 2015. Tagging Firefighter Activities at the Emergency Scene: Summary of AAIA'15 Data Mining Competition at Knowledge Pit, Proc. Federated Conference on Computer Science and Information Systems (FedCSIS'15), IEEE, pp. 379–385, http://dx.doi.org/10.15439/2015F426.

[9] Mörchen, F. 2003. Time series feature extraction for data mining using Discrete Wavelet Transform and Discrete Fourier Transform, Technical Report No. 33, Philipps-University Marburg, Germany.

[10] Preece, S.J., Goulermas, J.Y., Kenney, L.P.J., Howard, D. 2009. A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data, IEEE Transactions on Biomedical Engineering 56:871–879, http://dx.doi.org/10.1109/TBME.2008.2006190.

[11] R Core Team. 2015. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.

[12] Read, J., Pfahringer, B., Holmes, G.J., Frank, E. 2009. Classifier Chains for Multi-label Classification, Machine Learning and Knowledge Discovery in Databases 5782:254–269, http://dx.doi.org/10.1007/978-3-642-04174-7_17.

[13] Ridgeway, G. et al. 2015. `gbm`: Generalized Boosted Regression Models, R package version 2.1.1, http://CRAN.R-project.org/package=gbm.

[14] Safonov, I., Gartseev, I., Pikhletsky, M., Tishutin, O., Bailey M.J.A. 2015, An approach for model assessment for activity recognition, Pattern Recognition and Image Analysis 25:263–269, http://dx.doi.org/10.1134/S1054661815020224.

TABLE III
EVALUATION OF FEATURE IMPORTANCE ACCORDING TO MEAN DECREASE IN GINI IMPURITY INDEX FOR THE **POSTURE** CLASS.

Rank	Feature name	M.D.Gini
1	q40.acc_right_leg_x	38.16
2	q20.acc_right_leg_x	36.25
3	q30.acc_right_leg_x	34.59
4	q01.acc_torso_x	34.23
5	q10.acc_torso_x	32.50
6	q70.acc_left_leg_z	30.99
7	q05.acc_torso_x	29.25
8	q20.acc_torso_x	29.12
9	q30.acc_torso_x	27.50
10	q50.acc_right_leg_x	27.04
11	q80.acc_left_leg_z	25.88
12	ener.acc_right_leg_x	25.79
13	q90.acc_left_leg_z	24.90
14	q60.acc_left_leg_z	23.37
15	q10.acc_right_leg_x	22.35
16	q40.acc_torso_x	21.73
17	q50.acc_torso_x	20.85
18	q40.acc_left_leg_x	19.65
19	q30.acc_left_leg_x	18.86
20	q60.acc_right_leg_x	18.42
21	q95.acc_left_leg_z	17.89
22	q60.acc_torso_x	16.08
23	q20.acc_left_leg_x	14.87
24	q70.acc_right_leg_x	14.65
25	q50.acc_left_leg_x	14.09
26	q60.gyr_left_leg_y	13.17
27	med.rr	12.71
28	q70.gyr_left_leg_y	12.48
29	q80.gyr_right_leg_y	12.24
30	q50.acc_right_hand_x	12.13
31	cor.acc_torso_x.acc_torso_z	11.33
32	q70.acc_torso_x	10.53
33	q99.acc_left_leg_z	10.24
34	q40.acc_right_hand_x	9.84
35	q60.acc_right_hand_x	9.57
36	q80.gyr_left_leg_y	9.46
37	q50.acc_left_leg_z	9.14
38	q40.acc_left_leg_z	8.95
39	q90.acc_torso_x	8.71
40	q95.acc_torso_x	8.62
41	ener.acc_left_leg_x	8.44
42	q60.acc_left_leg_x	8.31
43	Period.sd.acc_right_leg_x	8.09
44	q30.acc_left_leg_z	7.74
45	ener.acc_right_hand_x	7.74
46	q80.acc_torso_x	7.63
47	sd.acc_right_leg_x	7.55
48	ModFFT.sd.acc_right_leg_x	7.22
49	q80.acc_right_leg_x	6.73
50	q90.gyr_right_leg_y	6.65

TABLE IV
EVALUATION OF FEATURE IMPORTANCE ACCORDING TO MEAN DECREASE IN GINI IMPURITY INDEX FOR THE **ACTION** CLASS.

Rank	Feature name	M.D.Gini
1	cor.acc_left_leg_x.gyr_left_leg_y	12.03
2	cor.acc_right_leg_x.gyr_right_leg_y	10.56
3	cor.gyr_left_leg_y.gyr_right_leg_y	10.50
4	q50.acc_left_hand_x	8.71
5	q60.acc_left_hand_x	8.40
6	q30.acc_left_hand_x	8.31
7	q40.acc_left_hand_x	7.95
8	q70.acc_left_hand_y	7.87
9	q20.acc_right_hand_y	7.85
10	q30.acc_right_leg_x	7.54
11	deriv1max.acc_left_arm_z	7.17
12	ModFFT.sd.gyr_left_leg_y	7.15
13	q20.acc_right_arm_z	7.15
14	Period.sd.acc_left_leg_z	7.12
15	ener.gyr_left_leg_y	7.12
16	Period.sd.gyr_left_leg_y	6.97
17	q20.acc_torso_x	6.92
18	q20.acc_left_hand_x	6.90
19	q70.gyr_left_leg_y	6.74
20	q10.gyr_left_leg_y	6.70
21	q70.acc_left_hand_x	6.57
22	q20.acc_right_leg_x	6.56
23	q40.acc_left_leg_x	6.27
24	sd.gyr_left_leg_y	6.21
25	q10.acc_right_arm_z	6.20
26	q30.acc_left_leg_x	6.06
27	q50.acc_left_leg_x	5.91
28	q10.acc_left_hand_x	5.83
29	q40.acc_right_arm_z	5.76
30	ModFFT.sd.acc_left_leg_z	5.64
31	q30.acc_right_arm_z	5.59
32	q95.acc_left_hand_y	5.48
33	q80.gyr_right_hand_y	5.40
34	q80.acc_left_hand_x	5.37
35	ener.acc_right_arm_z	5.37
36	q20.acc_left_leg_x	5.31
37	q90.acc_right_leg_z	5.04
38	q80.gyr_left_leg_y	4.98
39	ener.gyr_right_hand_y	4.85
40	Period.sd.gyr_right_arm_x	4.81
41	q70.acc_torso_x	4.78
42	ener.acc_left_leg_x	4.68
43	deriv1min.acc_left_arm_z	4.68
44	ener.acc_left_hand_x	4.48
45	q60.acc_left_leg_x	4.43
46	q40.acc_left_arm_x	4.43
47	q95.acc_torso_x	4.41
48	q05.gyr_right_hand_y	4.38
49	sd.acc_left_leg_z	4.33
50	Period.sd.acc_right_leg_z	4.28