

Human Activity Recognition based on Machine Learning Classification of Smartwatch Accelerometer Dataset

Dušan S. Radivojević

Department of Thermal engineering and Energy, „VINČA" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia

Nikola S. Mirkov

Department of Thermal engineering and Energy, „VINČA" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia

Slobodan Maletić

Department of Thermal engineering and Energy, „VINČA" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia

This paper presents two Machine Learning models that classify time series data given from smartwatch accelerometer of observed subjects. For the purpose of classification we use Deep Neural Network and Random Forest classifier algorithms. The comparison of both models shows that they have similar performance with regard to recognition of subject's activities that are used in the test group of the dataset. Training accuracy reaches approximately 95% and 100% for Deep Learning and Random Forest model respectively. Since the validation and recognition, reached about 81% and 75% respectively, a tendency for improving accuracy as a function of number of participants is considered. The influence of data sample precision to the accuracy of the models is examined since the input data could be given from various wearable devices.

Keywords: Human Activity Recognition, Machine Learning, Smartwatch.

1. INTRODUCTION

Monitoring human activities via smart devices, such as smartphones and smartwatches, and processing them by machine-learning-based algorithms, becomes an attractive field of research with various important applications. Human activity recognition could be an important tool in health monitoring of elderly patients [1], especially of patients suffering from various chronic diseases [2]. An important recent application appeared in sports for calorimetric calculations of physiological energy during training [3]. It is useful for detection of various cardiovascular diseases in combination with blood pressure logging wearable device [4] and even for prevention of stroke [5]. There are many other applications that could benefit from human movement recognition using motion data collected from smart devices like smart watches or phones. In most cases data acquired is often stored in the form of temporal streams of data, and the success of recognizing human activities depends on the algorithms applied, as well as on the availability of training data.

Sensors that are in use for monitoring human activities are accelerometer and gyroscope that are often manufactured in one integrated casing called Inertial Measurement Unit (IMU). These devices are able to detect linear and rotational motion relative to three perpendicular axes providing measurements for six degrees of freedom (6DOF) motion. The accelerometer sensor is able to sense gravitational acceleration as well. That fact significantly complicates data processing and use of the machine learning is for that purpose very useful. Measurement of gravitational acceleration in three dimensional system is often in use to display orientation in various electronic devices. Because of that

accelerometer is more often found in these devices than gyroscopes. Human activity recognition based only on accelerometer sensing data could therefore, be very important for practical application and widely used.

There is a growing tendency to use machine learning models such as artificial neural networks (ANN) to provide classification methodology for large amount of data generated by various monitoring devices e.g. health monitoring [6]. To answer this challenging issue, and within this context, this paper presents an application of two machine learning models that classify time series data given from smartwatch accelerometer of observed subjects. For the purpose of classification, Deep Neural Network and Random Forest classifier algorithms are used, and the comparison between models is conducted for the test group of the dataset. The improvement of accuracy as a function of the number of participants is considered, alongside with the estimation of the optimal number of participants. The influence of data sample precision to the accuracy of the models is examined, since the input data could be given from various wearable devices.

Deep learning neural networks (DNN) represent recent advancement in which computational neural network models are composed of multiple processing layers with representations of data in multiple levels of abstraction. These methods have significantly improved performance regarding speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics [7]. Random forests (RF) [8] is another well known machine learning algorithm, which shows relatively good results of classification even if it was trained on small amount of training data. That characteristic of RF model makes it suitable for this research. These two models are also appropriate for image classification [9].

The study in this paper is based on WISDM dataset that has smartwatch and smartphone, accelerometer and gyroscope records of 51 subjects for 18 different daily living activities [10]. The database was created to investigate the possibility of biometric identification of subjects [11]. In present research however, we use this

Received: August 2020, Accepted: November 2020

Correspondence to: Dušan S. Radivojević

„VINČA" Institute of Nuclear Sciences,
Mike Petrovića Alasa 12-14, 11000 Belgrade, Serbia

E-mail: dusanr@vin.bg.ac.rs

doi: 10.5937/fme2101225R

© Faculty of Mechanical Engineering, Belgrade. All rights reserved

FME Transactions (2021) 49, 225-232 225

dataset to investigate the possibility for recognition of unknown person's activity, based on model prediction over smartwatch accelerometer sensing dataset. Our research is based on use of time series of sensing data directly, without additional featurizing. That approach, with minimal computation for input data preprocessing, is more convenient for potential use in embedded or edge in IOT devices with limited hardware resources.

The paper is structured as follows. In the following section we describe the object data collection in greater detail, after that we describe machine learning models and software employed in the study. In third section we describe training, validation and test process and obtained results, followed by discussion on influential parameters. We conclude the article by giving recommendations on practical use of machine learning models for movement recognition.

2. DESCRIPTION OF DATASET

The WISDM database contains a collection of data of 51 subjects [10]. For the training and testing phase of the study, we have divided the dataset into two groups of 45 and 6 subjects, respectively. Each subject performed 18 different daily activities that lasted for approximately 3 minutes. These activity segments are cut into pieces of 100 accelerometer records for all three sensor axes. Those array segments are concatenated for all three axes so that final input shape of the models is flat array of 300 recorded values. Graphical presentation

of input signals from one subject with four selected activities is presented on Figure 1.

Sensor data rate is 20 Hz, which means that each data segment duration is about 5 seconds. Validation dataset consist of last segment data series of the each activity for all subjects in the training group. Validation segments are used only for validation process in each iteration of the training activity. Prediction of the model on this input segment group could be considered as identification of the training group's activities because it is generated from the same subjects that were in the training group. Finally, records of the last 6 subjects are used in its entirety only for testing. Trained model has not used this data during training so it could be considered as process of recognition for daily activities of unknown persons.

From eighteen given daily living activities in the dataset [11], ten are chosen as important such as walking, jogging, climbing the stairs, etc., which showed best prediction results. Activities in the group of eating showed great dispersion of predicted results, probably because of great diversity of movement during lunch of the given subjects. There are also activities that have similar sensor readings like sitting and typing or walking on the flat and taking the stairs so that the model could easily mix them up. For each of ten chosen activities that are coded with letters from "A" to "S", indexes from 0 to 9 are assigned, to be output values for appropriate data segments, Table 1.

Some statistical properties of data input segments are given in Table 2. These statistical values are calculated based on randomly chosen segments, of all selected activities, from one subject so they can be compared.

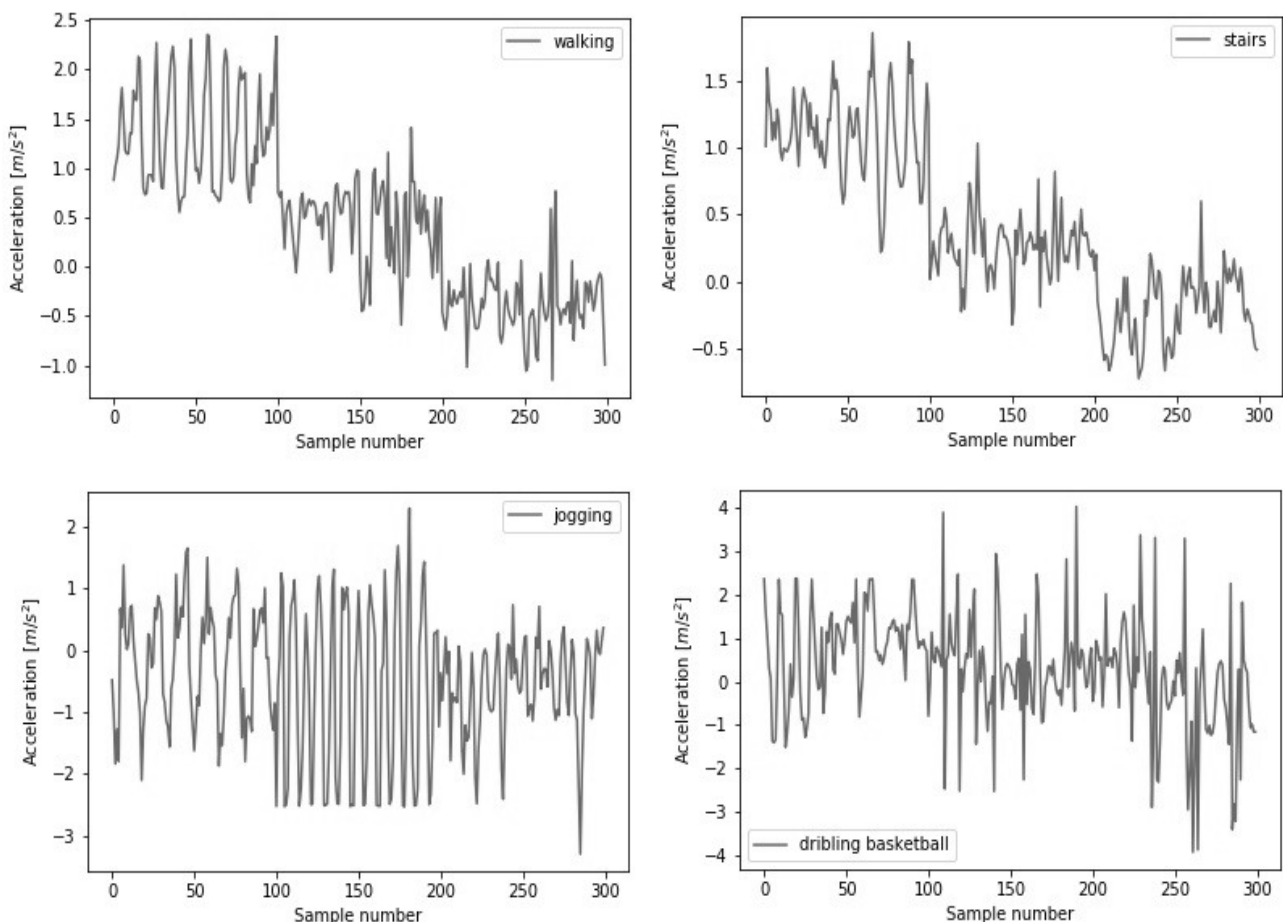


Figure 1. Input signals used in classification representing four different activities of a single subject

Table 1. Table of chosen human daily activities with its original code and given index

Activity	Walking	Jogging	Stairs	Sitting	Standing	Brushing Teeth	Playing Catch	Dribbling Basketball	Writing	Folding Clothes
Code	A	B	C	D	E	G	O	P	Q	S
Index	0	1	2	3	4	5	6	7	8	9

Table 2. Statistical properties of used time series

Activities	Mean	Variance	Skewness	Kurtosis
walking	0.470	0.664	0.111	-0.863
jogging	-0.475	1.211	-0.293	-0.594
stairs	0.387	0.370	0.348	-0.849
sitting	0.628	0.325	0.612	-1.498
standing	0.459	0.238	-0.294	-1.499
brushing teeth	-0.269	0.837	-0.263	-1.331
playing catch	0.328	1.159	-0.395	2.338
dribbling basketball	0.323	1.570	-0.334	1.085
writing	-0.530	0.113	-0.004	-0.831
folding clothes	-0.364	0.462	0.300	-0.038

Table 3. Influence of different data setup percentage on DNN model accuracy

Training/Validation/Testing data ratio	Training accuracy [%]	Validation accuracy [%]	Test accuracy [%]
87:3:10	96.51	81.62	75.59
78:2:20	95.79	77.94	71.48
69:1:30	94.94	82.12	55.84

Accelerometer time series were converted from textual format to 64 bit floating point numbers. Models are trained and tested with this dataset. For the next testing, dataset was cast into 32 bit precision floating point numbers and models are trained again, and the last dataset was cast into 16 bit precision, and the process was repeated. The purpose of this data manipulation was the examination of the influence of data precision on models' classification accuracy. For all cases the accelerometer data was appropriately normalized.

Both DNN and RF models are provided with the same data input so they could be qualitatively analyzed in terms of ability to perform classification of human activities.

Data collection prepared in described way consists of 17238 samples divided in the following way: 87% of data is used for training, 3% for validation, and 10% for testing, Table 3. Cross validation of time-series is performed using K-fold Cross Validation technique.

Besides DNN and RF models, to be able to compare the performance of the models, we have introduced K-nearest neighbors (KN) as a control model.

3. MODELLING AND EVALUATION

3.1 Machine Learning Models

For classifications we use Deep Neural Network and Random Forest classifier algorithms, both implemented in open-source *Keras/TensorFlow* [12] and *scikit-learn* Python programming language libraries, respectively. The additional control model, the K-Nearest Neighbors Classifier is accessed through *scikit-learn* library as well.

The Deep Neural Network Model is sequential with two hidden dense layers and one dense output layer. First hidden layer has 300 neurons, input shape of 300 input data array and ReLu activation function. The second hidden layer has 500 neurons and *tanh* activation function. The output layer has 10 neurons and each of

the indices corresponds to one of the selected human activities. The output layer activation function is *softmax* that gives probability for each output neuron according to the model prediction for the given data input.

The second considered model is Random Forest classifier from *scikit-learn* library. While trained the Random Forest model has been iteratively improved, so that the number of trees has been increasing, until specified convergence criteria was satisfied.

The third model, which serves as a control model, is K-Neighbors Classifier (KN) from *scikit-learn* Python library. This model is suitable to illustrate the difficulty of the considered problem, and is used here for statistical testing of algorithms' performance. In KN model, classification is computed from a simple voting of the nearest neighbors of each point [13].

3.2 Training process and results

The chosen learning model is Supervised Learning. In supervised learning classification, search criteria allow decision on whether a sample belongs to a certain class of patterns. The identification of decision functions is based on examples where it is known a priori to which class they belong [14]. According to known result of classification, the model weights are adjusted to minimize the error for given input data. This process is known as supervised training of machine learning model.

The Deep Neural Network model is trained by TensorFlow's backpropagation algorithm. Model fitting was conducted through 30 epochs including the shuffle of the training data. Random shuffle option is included because it can reduce total training time and increase test prediction accuracy [15]. Training and validation precision was logged during the learning process of DNN model, Figure 2. Accuracy of the model prediction on training data reached 95.3%, and on validation data 80.6%.

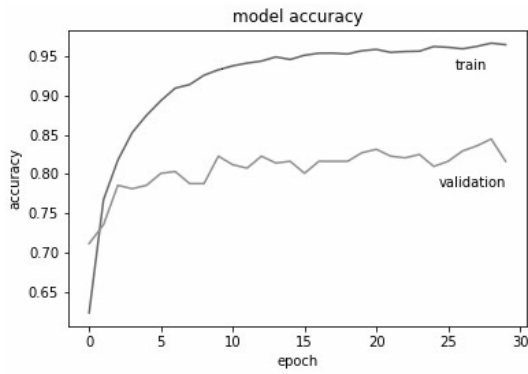


Figure 2. Prediction accuracy of Deep Learning model

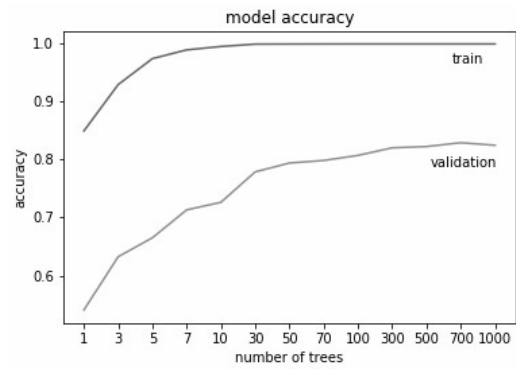


Figure 3. Prediction accuracy of Random Forest model

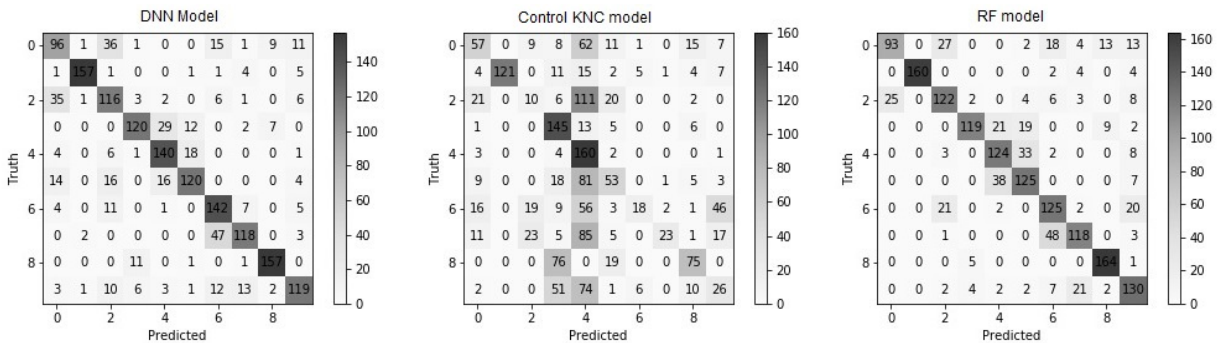


Figure 4. Confusion matrices for Deep Neural Network (DNN), K-nearest neighbors (control model) and Random Forest (RF) models

The Random Forest model is trained, verified and tested with the same data samples as DNN model. The model performance is tested according to the number of classifiers, Figure 3. It was found that with the increase of the number of trees to more than 700, there is no more significant increase of precision for validation group of samples. Training precision reaches 100% through iteration, which is higher accuracy than for the DNN model. However, achieved validation precision is little better in comparison to the last epoch of DNN prediction and reached 81.8%.

The control model - K-nearest neighbors classifier has achieved 60.2 % of training accuracy and 50.8 % of validation accuracy.

3.3 Human motion recognition test results and statistical analysis

Test predictions are based on human activities recognition for subjects that were not in the training group. This is important for the present study because of the fact that every subject, as a human being, has its own way of physical movement [11]. Better prediction result could be achieved if the whole dataset was split in a way that all subjects are represented in training and testing dataset groups. We are considering here a more objective approach where test group must be made from entirely different subjects than training group.

Prediction performance of the presented models is conducted with test data that consists of 1700 samples. Testing was done with 1000 randomly selected samples from mentioned test dataset in 50 repetitions. Predicted array of labels was compared with original labels from dataset, and binary (True/False) array was created for each test and each model. Achieved mean value of prediction accuracy was 75.9 %, 75.2 %, and 40.3 %

with standard deviation of 1.3 %, 1.1 %, and 1.7 % for DNN, RF, and KNC models respectively.

To statistically assess performance of the algorithms, series of statistical tests are conducted, following the examples from [16,17]. We have performed a multiple pairwise comparison of algorithms, between DNN and RF model, as well as comparison of both models with the KN model, and the results are summarized below in Table 4 for t-test. Wilcoxon signed rank test for the DNN/RF pair, shown in Table 5, showed that the result of slightly better average performance of DNN model has statistical significance at $p < 0.05$.

To assess model's performance further we have conducted a series of tests on obtained classification confusion matrices. the confusion matrix is a specific table that allows visualization of the prediction performance of ML models. The confusion matrices produced in testing of DNN, KN and RF models are represented in Figure 4. As can be seen, most of the predictions are located on the diagonal in the given matrix of DNN model and also very similar representation is for RF model. In ideal case, with 100% of accuracy, none of the predictions should be out of matrix diagonals.

The confusion matrix of the control KN model shows poor classification performance relative to previous two. It can be noticed in results of DNN and RF model that some activities lead to misclassification more than others. These examples are labels 0 and 2 that correspond to walking and climbing stairs and also labels 4 and 5 that correspond to standing and brushing teeth respectively. There is also confusion with label 7 that correspond to dribbling basketball activity that is often predicted as label 6 that correspond to playing catch with tennis ball. Activities that have best predictions are jogging and writing, with labels 1 and 8 respectively. The reasons for these results are probably in the fact that there was lots of

diversity in subject's movements and that there were relatively low number of participants, which influences efficiency of the model prediction.

Table 4. t-test statistics for multiple pairwise comparison, the critical t-value (alpha=0.05, DOF=98) t-critical = 1.99

Pairwise test	Mean	t-value	p-value
DNN/RF	758.7/752.2	2.70	8.28(-03)
DNN/KN	758.2/402.68	120.70	1.29(-102)
RF/KN	752.2/402.68	122.75	4.84(-99)

Table 5. Wilcoxon signed rank test for DNN/RF pair

Parameter	Value
W -value	251
Mean Difference	9.7
Sum of pos. ranks	1024
Sum of neg. ranks	251
Z -value	-3.371
Mean (W)	637.5
Standard Deviation (W)	103.59
Sample Size (N)	50
p -value	0.0002

To evaluate studied classification procedures statistically, we will also undertake multiple pairwise comparisons of confusion matrices produced by classifications in test runs. The study follows comparison methods advocated in [18], where various different test are suggested which act as decision rules in the sense of statistical hypothesis tests. In these tests by concordant elements of confusion matrices we denote the elements for which the classification was correct for that particular category, graphically they are represented on the main diagonal of the confusion matrix. A single binomial contrast test, compares global proportion of concordant elements in two confusion matrices. The global proportion of concordant elements for two compared matrices π_A , and π_B , respectively is the sum of the number of elements on main diagonal divided by total number of classified elements. The total number of classified elements is 1700 in all cases. The null hypothesis in this test is that classification procedures are equal in statistical sense by virtue of equality of concordant element proportions in confusion matrices. The null hypothesis is rejected if absolute value of difference is above specified threshold. The results of these tests are shown in Table 6.

Table 6. Pairwise testing of confusion matrices using single binomial contrast test

Pairwise test	n/m	π_A	π_B	$abs(\pi_A - \pi_B)$
DNN/RF	1700	0.756	0.753	0.003
DNN/KN	1700	0.756	0.405	0.351
RF/KN	1700	0.753	0.405	0.348

With an overall chi-square test, the squared z-vector is assumed to follow chi-squared distribution with k degrees of freedom, if all column/row concordance proportion estimators are equal (i.e. if null hypothesis holds), so a single test for rejection of null hypothesis of whole equality between two confusion matrices is used, Table 7. In Kolmogorov-Smirnov test we assume that once the confusion matrices are turned into vectors, in consistent manner, the sample values are drawn from the same distribution. We then test discrepancy measure

between them to draw conclusion about rejecting the null hypothesis. Test statistics and p-values for this test are given in Table 8.

Table 7. Pairwise overall Chi-squared tests for confusion matrices

Pairwise test	Chi-squared statistic	p-value
DNN/RF	23.799	0.004631
DNN/KN	625.549	6.859e-129
RF/KN	N/A	N/A

Table 8. Pairwise Kolmogorov-Smirnov tests for confusion matrices

Pairwise test	Kolmogorov-Smirnov statistic	p-value
DNN/RF	0.120	0.470
DNN/KN	0.130	0.368
RF/KN	0.180	0.078

3.4 Influence of data precision on the accuracy of the models

Many applications that use sensor readings often need some kind of data processing like noise reduction, averaging or upscaling [19]. Also there are different manufacturers and models of accelerometer sensors and they have different precision of the output data sensing results. There are also studies that investigate the possibility of using low precision data formats that could positively effect on model training and prediction time, energy and memory consumption [20]. For this purpose, the influence of sensor data precision on model prediction accuracy is examined. Data given from database was converted from string to floating point format with precision of 64 bits. First examination was with this data precision. For the second examination data was cut-down to 32 bits and finally to 16 bits for third. Models are trained and tested with given data precisions. Results are presented in the Figure 5.

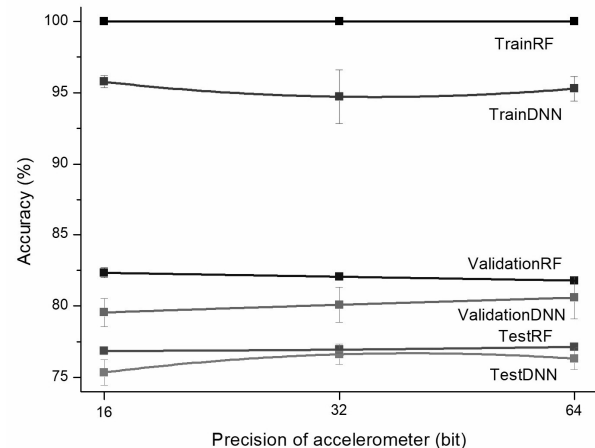


Figure 5. Precision of accelerometer data versus accuracy of models

As can be seen from Figure 5, the Random Forest model is very resistant to change in input data precision. Test prediction accuracy decreases in given precision range for only 0.3%. It is interesting that prediction accuracy of validation data unexpectedly increases with lower precision of data. However DNN model shows negative influence of the lower data precision on model prediction

accuracy. Negative trend is recorded for validation and test accuracy for approximately 1% in the given range.

3.5 Influence of the number of participants on test accuracy of the models

According to the presented results where validation and test prediction accuracy are significantly less than training model accuracy, it is obvious that models are relatively under-fitted. The reason for this could be that training data set is insufficiently large regarding the number of participants.

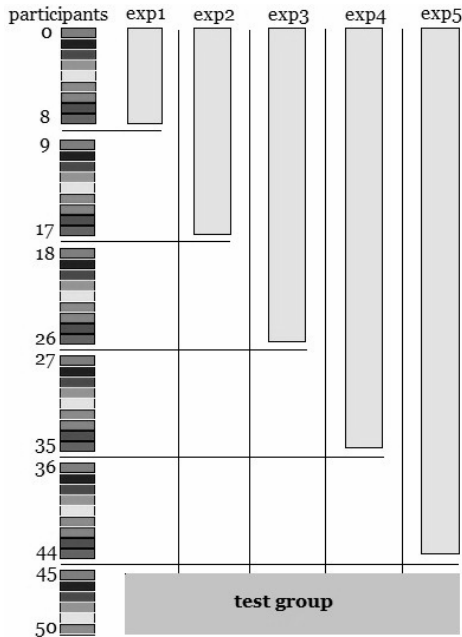


Figure 6. Participants schematic for the experiment

The following sequence of tests should determine if it is possible to increase prediction accuracy of a model by increasing the number of participants. It is known, however, that in the case of under-fitted models with small datasets, this is hard to achieve [21]. For this experiment the same group of first 45 subjects was used for training like in previous sets and they are separated into 5 groups with 9 subjects as shown Figure 6.

The test group is the same as previous one and consists of last 6 subjects. Experiment was done on DNN and RF models with 64 bit data precision. In the first experiment models were trained on first group of subjects and tested with last 6 subjects. In the second experiment models were trained with first two groups etc. In all 5 experiments test group was the same. Dependency of model's accuracy relative to number of participants is presented on Figure 7.

We can conclude that by increasing the number of participants the test accuracy has tendency of increasing. It is confirmed that for better testing accuracy the number of training participants should be significantly greater.

This result corroborates findings of other researchers. Similar research based on 17 subjects but with 10 seconds of sample time duration, and using input with features, achieved 70.3% of testing prediction for impersonal human activity recognition [22]. Also research with similar conditions and 51 subjects reached human activity recognition of 87.8% [11]. These results show

the tendency similar to the one presented here, and they all confirm the hypothesis, that better result of prediction could be established with more subjects involved. This tendency is also depicted in Figure 8, where dependence of loss function on involved number of participants is shown. This diagram is produced for DNN model with 64 bits data precision.

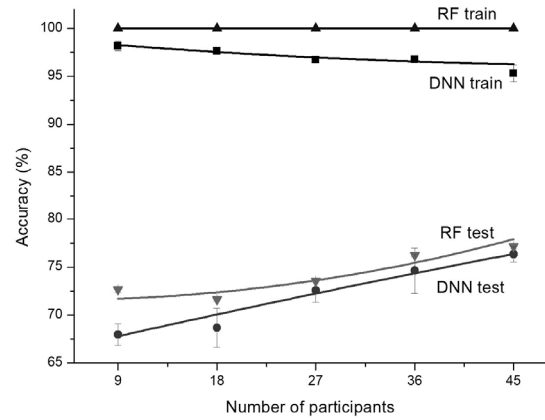


Figure 7. Number of participants versus accuracy of the models

The loss function is quantity that a model should minimize during training. In the present study, the loss function for DNN model is sparse categorical cross-entropy. It is obvious that loss function has tendency to decrease its value with increasing numbers of participants. In order to assess optimal number of participants, optimal exponential curve, fitted to results, is added.

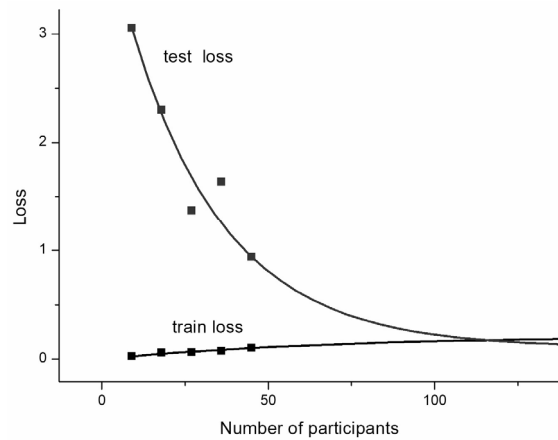


Figure 8. Loss function according to number of participants

Since train loss and test loss are approaching with increasing the number of participants, a value where two fitted curves intersect could be chosen for the optimum number of participants. In such a way we estimated optimal number of participants to 115. To be precise this is the number of subjects that should be in the training group of the dataset. Because this optimal number is far from given one, this approach should be repeated after increasing the number of subjects to this optimal value in the future study.

4. CONCLUSION

In this study we have shown that it is possible to use Machine Learning models to recognize human activity using smartwatch accelerometer dataset, with a varying

level of accuracy, depending on the employed ML model. The study considered the use of smartwatch accelerometer time series data directly, since it is favorable for various applications. Achieved precision of the models, on average, showed small advantage of Deep Neural Network model over Random Forest model. This was confirmed under statistical analysis of algorithms' performance. The influence of data precision on prediction accuracy of the models is not significant so it allows greater freedom in development and eventual savings in timing, memory or energy consumption. When influence of data precision and numbers of participants is examined, Random Forest model showed greater robustness over DNN model. It is also demonstrated that increasing the accuracy of human activity recognition is possible by using a larger number of involved participants.

SUPPLEMENTARY MATERIAL

For reproducibility purposes, interested readers can access code developed for this study directly through public repository: (<https://github.com/DusanRandD/Human-Activity-Recognition>).

ACKNOWLEDGMENTS

The research was funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

REFERENCES

- [1] Wang, Z., Yang, Z., Dong, T.: A Review of Wearable Technologies for Elderly Care that Can Accurately Track Indoor Position, Recognize Physical Activities and Monitor Vital Signs in Real Time, *Sensors* (Basel, Switzerland), Vol. 17, No. 2, 341, 2017.
- [2] Axisa, C.: The role of human activity recognition in healthcare: a study focusing on patients suffering from chronic illnesses, Bachelor thesis, Faculty of Economics, Management and Accountancy, University of Malta, 2017.
- [3] Ndahimana D, Kim E-K.: Measurement methods for physical activity and energy expenditure: a review, *Clinical Nutrition Research*, Vol. 6, No. 2, pp. 68-80, 2017.
- [4] Arakawa, T.: Recent research and developing trends of wearable sensors for detecting blood pressure, *Sensors*, Vol. 18, No. 9, p. 2772, 2018.
- [5] Villar, J.R., González, S., Sedano, J. Chira, C. and Trejo-Gabriel-Galan, J.M.: Improving human activity recognition and its application in early stroke diagnosis, *International Journal of Neural Systems*, Vol. 25, No. 4, 1450036, 2015.
- [6] Punithavathy, K., Poobal, S., Ramya, M.M.: Performance Evaluation of Machine Learning Techniques in Lung Cancer Classification from PET/CT Images, *FME Transactions*, Vol. 47, No. 3, pp. 418-423, 2019.
- [7] LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, *Nature*, Vol. 521, pp. 436-444, 2015
- [8] Breiman, L.: Random Forests, *Machine Learning*, Vol. 45, No. 1, pp. 5-32, 2001.
- [9] Patel, D. R., Vakharia, V., Kiran M. B.: Texture Classification of Machined Surfaces Using Image Processing and Machine Learning Techniques, *FME Transactions*, Vol. 47, No. 2, pp. 865-872, 2019.
- [10] Weiss, G.M.: WISDM Smartphone and Smartwatch Activity and Biometrics Dataset, Department of Computer and Information Science Fordham University, Bronx NY, 2019.
- [11] Weiss, G.M., Yoneda, K., Hayajneh, T.: Smartphone and Smartwatch-Based Biometrics Using Activities of Daily Living, *IEEE Access*, Vol. 7, pp. 133190-133202, 2019.
- [12] Abadi, M., Barham, P., Chen, J. et al.: TensorFlow: A system for large-scale machine learning, in: *Proceedings OSDI16*, 265–283, 2016.
- [13] Goldberger, J., Roweis, S., Hinton, G., Salakhutdinov, R.: Neighbourhood Components Analysis, *Advances in Neural Information Processing Systems*, Vol. 17, pp. 513-520, May 2005.
- [14] Langer, H., Falsaperla, S., Hammer, C.: *Advantages and Pitfalls of Pattern Recognition*, Elsevier, Amsterdam, 2020.
- [15] Ke, Z-L, Cheng H-Y, and Yang, C-L: LIRS: Enabling Efficient Machine Learning on NVM-Based Storage via a Lightweight Implementation of Random Shuffling, arXiv:1810.04509, 2018.
- [16] Vuković, N., Petrović, M., Miljković, Z., A comprehensive Experimental Evaluation of Orthogonal Polynomial Expanded Random Vector Functional Link Neural Networks for Regression, *Applied Soft Computing*, Vol. 70, pp. 1083-1096, 2018.
- [17] Vuković, N., Miljković, Z., A growing and pruning sequential learning algorithm of hyper basis function neural network for function approximation, *Neural Networks*, Vol. 46, pp.210–226, 2013.
- [18] Rodríguez-Avi, J., Ariza-López, F. J., Alba-Fernández, V., Methods for comparing two observed confusion matrices, *Proceedings of the 21st AGILE International Conference on Geographic Information Science*, Lund, 2018.
- [19] Radivojević, D., Milošević, N.: Influence of selected software upscale on statistical characteristic of A/D conversion used in a DC Power Supply developed in the Institute Vinca (in Serbian), in: *Proceedings of the Congress of Metrologists*, 12-15.10.2015, Zlatibor .
- [20] Carmichael, Z., Langroudi, H., Khazanov, C., Lillie, J., Gustafson, J., Kudithipudi D.: Performance-Efficiency Trade-off of Low-Precision Numerical Formats in Deep Neural Networks, in: *CoNGA'19: Proceedings of the Conference for Next Generation Arithmetic 2019*, 25.03.2019, Singapore.
- [21] Schutten, M., Wiering, M.: An Analysis on Better Testing than Training Performances on the Iris Dataset, in: *Proceedings of Belgian Dutch Artificial Intelligence Conference*, 10–11 November 2016, Amsterdam, The Netherlands.
- [22] Weiss, G.M., Timko, J.L., Gallagher, C.M., Yoneda, K. and Schreiber, A.J.: Smartwatch-based activity recognition: a machine learning approach,

NOMENCLATURE

DNN	Deep Learning Neural Network
RF	Random Forest
IMU	Inertial Measurement Unit
AI	Artificial Intelligence
CSV	Comma Separated Value
WISDM	Wireless Sensor Data Mining
IOT	Internet Of Things

РАСПОЗНАВАЊЕ ЛЈУДСКИХ АКТИВНОСТИ КЛАСИФИКАЦИЈОМ АКЦЕЛЕРО- МЕТАРСКИХ ПОДАТАКА СА ПАМЕТНИХ САТОВА ПОМОЋУ МОДЕЛА МАШИНСКОГ УЧЕЊА

Душан С. Радивојевић, Никола С. Мирков,
Слободан Малетић

У приказаном истраживању се разматрају карактеристике два модела из домена машинског учења помоћу којих се врши класификација података добијених у виду записа акцелерометара паметних сатова посматраних испитаника. За потребе истраживања коришћени су алгоритми класификације "Deep Neural Network" и "Random Forest". Поређење оба алгоритма показује њихов релативно сличан успех по питању препознавања активности испитаника које су сврстане у тест групу података, док оба модела показују већу прецизност класификације у односу на контролни модел. Прецизност током тренирања модела достиже 95% и 100% у случају DNN и RF алгоритама појединачно. С обзиром да је успех валидације и препознавања износио 81% и 75% појединачно, разматрана је тенденција за унапређење прецизности у функцији повећања броја учесника испитивања. Утицај прецизности записа података на прецизност класификације помоћу посматраних модела такође је разматрана из разлога разноврсности порекла података због разноликости носивих уређаја.