Classical Machine Learning Approach to Classify Complex Nurse Activity

Md. Ahasan Atick Faisal
University of Dhaka
atickfaisal@gmail.com

Md. Sadman Siraj
University of Dhaka
sadman.siraj@ieee.org

Md. Tahmeed Abdullah
University of Dhaka
tahmeedabdullah@ieee.org

Omar Shahid
University of Dhaka
omarshahid232@gmail.com

Farhan Fuad Abir
University of Dhaka
farhan.fuad@ieee.org

M. A. R. Ahad
University of Dhaka
atiqahad@du.ac.bd

ABSTRACT

Nursing activity recognition add a new dimension to the healthcare automation system. But nursing activity recognition is very challenging than identifying simple human activity like walking, cycling, swimming, etc. due to intra-class variability between activities. Besides, lack of proper dataset does not allow researchers to develop a generalized method for nursing activity or comparing baseline methods on different datasets. Nurse Care Activity Recognition Challenge 2020 provides a dataset of twelve nursing activities. In this paper, we have described our (**Team Hex Code**) approach where we have emphasized on developing method which can cope up with real world data with noise and uncertainty. In our method, we have resampled our data to deal with variable sample frequency of dataset and we have also applied feature selection method on extracted feature to have the best combination of feature set for classification. We have used random forest classifier which is a classical machine learning algorithm. Applying our methodology, we have got 78% validation accuracy on dataset. We have trained our model on lab dataset and validate them on field dataset.

CCS CONCEPTS

• **Learning** → **Supervised**; • **Data Pre-processing** → *Resampling*; • **Feature** → Time Domain and Frequency Domain; • **Algorithm** → Random Forest.

KEYWORDS

activity recognition, accelerometer, inertial measurement unit, nurse care, smart hospital

ACM Reference Format:

Md. Ahasan Atick Faisal, Md. Sadman Siraj, Md. Tahmeed Abdullah, Omar Shahid, Farhan Fuad Abir, and M. A. R. Ahad. 2020. Classical Machine Learning Approach to Classify Complex Nurse Activity. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2020 International Symposium on Wearable Computers (Ubi- Comp/ISWC '20 Adjunct), September 12-16, 2020.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Human activity recognition deciphers to interpret human motion and it has been studied frequently over the last decade. Research and applications have been prioritized on illustrating and recognizing physical activity [4, 18]. Due to the abundance and easy availability of wearable technologies and a huge leap in data technology, tasks related to physical activity recognition using data collected from wearable or body-worn sensors have drawn significant attention. The technology is flexible to use in analyzing static as well as dynamic gestures and behaviors. In recent times, many commercial products like a wrist band or high configuration smartphone are available that can recognize walking, sitting, cycling, swimming, etc. human motion by an in-built application which may indicate an abstract idea of how active a human have been in a day [14]. However, activity recognition in the healthcare domain has received high importance in recent times, but this research sector is not mature enough. Research in this area is even more difficult as data related to healthcare is mostly unavailable publicly. Therefore, applications of activity recognition in the healthcare domain have not been explored broadly as it could have been.

The accelerometer sensor has been seen to be most frequently used for the physical activity recognition system in recent times. It is preferable for its usability that easy setup for recording data using any wearable devices like a band or watch and most importantly by a smartphone [1]. In general, accelerometer data draws out coordinates of the three-axis as well as rate of changing speed over time [11]. These types of data are very consummate in recording repetitive body postures like walking, running, cycling, swimming, climbing, etc. [7]. Bao and Intille [4] represented an overview of research work related to the physical activity recognition system through acceleration data. Different studies have shown that wrist, hip, ankle, knee, biceps are some ideal positions to attach an accelerometer as these positions are the key points for the motion of the human body [22]. Mantyjarvi et al. [19] recognized human posture and ambulation using acceleration data that was collected using a band attached to the hip. Antar et al. [3] provided a comprehensive study on some comparative approaches to classify smartphone accelerometer data.

Although researchers of this era have developed robust and sturdy learning algorithm for predicting location, user identity

UbiComp/ISWC '20 Adjunct, September 12-16, 2020,

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

but estimating user activities which has proven much more challenging. In hospitals, there is no available automated system that monitors their healthcare stuff's activity due to the lack of data and challenges. Inoue et al. [13] provided a first-ever nursing activity dataset using mobile sensors. Nurse care activity recognition is challenging than other daily activities as experimental setup or settings in which nurses perform an activity is very dynamic and robust, moreover, nursing activities vary due to different patients as nurses usually perform activities to a patient. For example, if a nurse gives an injection to a patient, then the activity of giving injection must vary according to the environment and position of the patient. These types of incidents introduce challenges that are not studied in other daily activity recognition applications. In other words, in nurse care activity, we see intra-class variability which depends not only on the nurses/subjects, as in other domains but also on the receiving patient. On the other hand, nurse care activities are quite similar in nature, therefore, its so difficult to classify activities. For example, collecting blood samples or pushing an injection to a patients body contains similar types of features.

Nurse Care Activity Recognition Challenge, a part of the HACSA workshop provided a dataset of 12 different nursing activities for classification purposes. The aim of the challenge is to create a bridge for the automation process of nurse care documentation, care routines, checking compliance, etc. [17]. This challenge opens up a great opportunity for the researchers to mitigate the gap and compare different methods of activity recognition on nurse care activity. Kadir et al. [16] proposed a feature extraction based method on motion capture data using K-Nearest Neighbors (KNN) classifier to predict nursing activities. In another study, Haque et al. [10] provided a Gated Recurrent Unit (GRU) model with an attention mechanism to recognize the nursing activity, where it used location-based as well as motion-based features. Besides, a classical approach using only acceleration data and random forest classifier performs better on recognizing nursing activity [21].

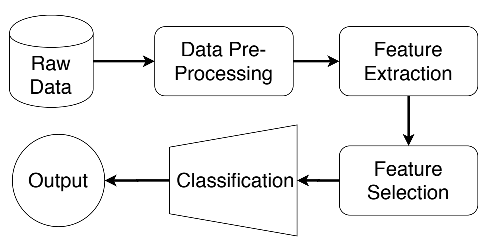

Figure 1: Block diagram overview of our method

In this paper, we have developed a method for Nurse Care Activity Recognition Challenge 2020 dataset to recognize the nursing activity. We have used a very simple method composed of feature extraction, feature selection, and supervised learning classifier. Our core concept is to build a generalized framework of predicting complex activities from real world data. Therefore, we have applied a several steps of data pre-processing and select the final extracted feature vector by applying a set of feature selection method which ensures to have smart features for classification. Figure 1 has described the abstract concept of our method. After giving an overview in section 1, we have organized the paper as follows: Section 2 has described the dataset with the core information. Section 3 has represented our methodology of all necessary steps with the required description. In Section 4, we have discussed the challenges that we have faced as well as the motivation behind our method and results with their evaluations. Finally, we have drawn the conclusion of the paper in Section 5.

2 DATASET DESCRIPTION

The dataset [12] contains nurse activity data collected from the 3-axis accelerometer sensor of a mobile phone attached to the subject's right hand. The data were collected both in an experimental setup in the lab and in field. The lab data was collected in the Smart Life Care Unit of the Kyushu Institute of Technology, Japan. In the lab experiment, 2 professional nurses participated. The real field data was collected in a Care Facility in Japan and data of 6 nurses for training and 3 nurses for testing are available from the real-life situation.

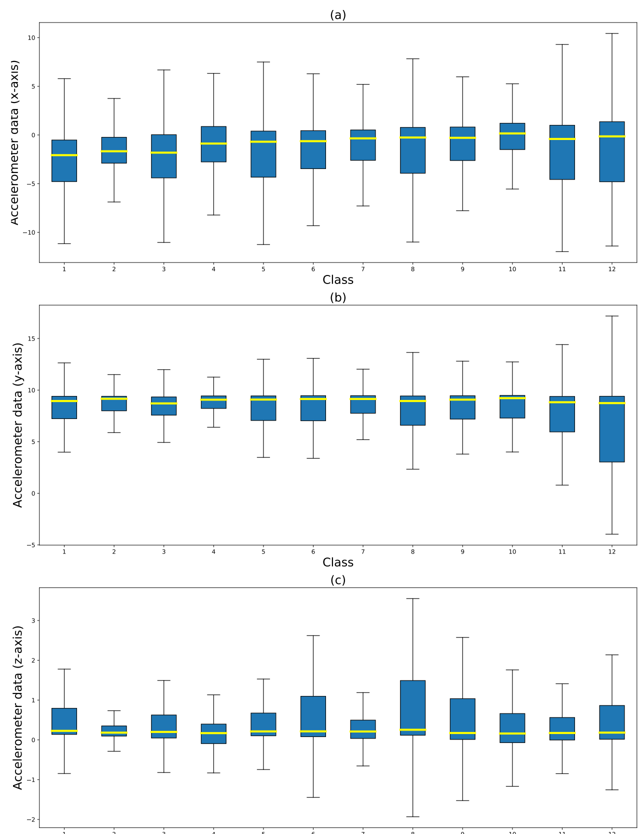

Figure 2: Visualizing the data distribution of accelerometer data in channels (a) x-axis, (b) y-axis and (c) z-axis.

The nurses perform 12 different activities and the activity classes are, C_1: Guide (from front), C_2: Partial assistance, C_3: Walker, C_4: Wheelchair, C_5: All assistance, C_6:Partial assistance (from the front), C_7: Partial assistance (from the side), C_8: Partial assistance (from the back), C_9: Position change to right lying position, C_{10}: Position change to left lying position, C_{11}: Lower body lifting and C_{12}: Horizontal movement. The classes are grouped under three superclass,

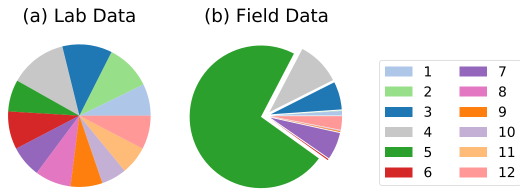

Figure 3: Distribution of classes in cleaned (a) Lab data and (b) Field data. The colors show the 12 activities provided in the dataset.

(A) Help in mobility ($C_1 - C_4$), (B) Assistance in Transfer ($C_5 - C_8$), (C) Position change ($C_9 - C_{12}$). The distribution of sensor data in x, y, and z axes for each of the 12 classes are shown in Fig. 2. The distribution of classes is highly imbalanced for field data, which shown in Fig. 3.

3 METHODOLOGY

Our method has focused on developing a simple learning technique that is computationally low lost, therefore, we have applied a classical machine learning algorithm on smartly extracted features. We have emphasized more on data pre-processing and feature extraction so that the learning algorithm can classify activity using the best feature combination.

3.1 Data cleaning and Pre-processing

The provided raw dataset contained label files and sensor data files. The label files contain the label for a range of data in time and the data file containing the sensor data of the accelerometer in 3 axes with the timestamps. The cleaning of data was done in four steps:

- **Unshuffling** The sensor data was provided in a shuffled way which was unshuffled by sorting the data samples according to date and time.
- **Duplicates** There were a lots of duplicate rows in the provided file. These duplicate data samples were dropped.
- **Label-matching** For many data in the data file, labels are not provided which were dropped.
  After these two steps only 1.84% of the raw data remained which was still discontinuous or irregularly sampled.
- **Resampling** The sensor values were sampled in an inconsistent frequency over the dataset. The data were resampled to a consistent 20Hz frequency throughout the dataset for extracting features in the frequency domain. Figure 4 shows, resampling the data smoothed the data and fill up missing samples in the time domain. As a result of resampling, the number of data samples increased by 19.73%.

3.2 Feature Extraction

The dataset was provided with only 3 sensor channels; accelerometer data for each of the three axes. We assume the data for the x-axis, y-axis and z-axis channel for i^{th} sample be denoted by $Acc_{x(i)}$, $Acc_{y(i)}$ and $Acc_{z(i)}$ respectively. Prior to the feature extraction process, an additional 5 channels were created from these 3 channels.

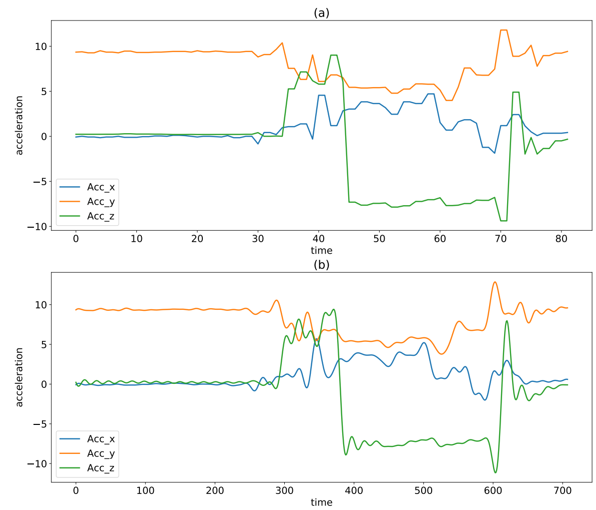

Figure 4: The plot of 3-axis accelerometer data of user 1 performing activity 8 in the lab showing the differences in (a) Unsampled data (b) Resampled data at 20Hz

These additional channels include the magnitude and the jerk of the original channels.

- **Accelerometer Magnitude** Using the magnitude along the three axes, the magnitude of acceleration for any given sample i is calculated as the Norm of the values of three axes.

$$Acc_{mag(i)} = \sqrt{{Acc_{x(i)}}^2 + {Acc_{y(i)}}^2 + {Acc_{z(i)}}^2} \tag{1}$$

- **Jerk** The rate of change of acceleration with respect to time is called Jerk. Assuming the time difference between two samples as $\Delta\tau$, the first-order derivative of the accelerometer data of each channel is calculated using the central difference technique to get a jerk channels for each of the accelerometer channels. The jerk of channel c of the i^{th} sample is calculated using Equation 2.

$$Jerk_{c(i)} = \frac{Acc_{c(i+1)} - Acc_{c(i-1)}}{2\Delta\tau} \qquad c = x, y, z \tag{2}$$

- **Jerk of Magnitude** This channel is calculated by taking the first order derivative of the Accelerometer Magnitude channel.

$$Jerk_{mag(i)} = \frac{Acc_{mag(i+1)} - Acc_{mag(i-1)}}{2\tau} \tag{3}$$

Window selection and feature vector computation: Each channel in the re-sampled data was segmented into sliding windows of an empirically chosen size of l samples with an overlapping of r. For M data samples in the training set, we get N windows.

$$N = \frac{M - l}{l \times (1 - r)} + 1 \tag{4}$$

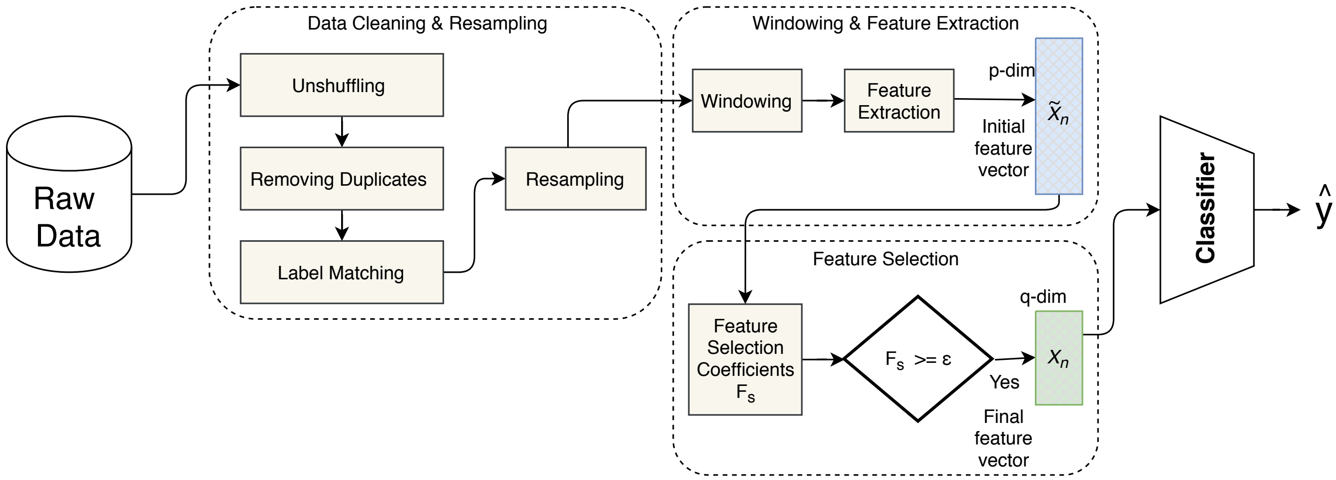

Figure 5: Block diagram of our proposed method

For the n^{th} window,a feature vector $\tilde{X_n}$ of length $p = 319$ (128 time domain,184 frequency domain and 7 other) was calculated both extracting features in both time and frequency domain from the generated channels. From the time domain, statistical features like *Mean*, *Standard deviation*, *RMS*, etc., signal processing features like *Mean Crossing rate*, *Maximum rate of change*, *Auto-Correlation* and some physics-based features like *Linear velocity* and *Kinetic energy* were calculated. After converting the channels of a window to frequency domain, the features like *Entropy*, *Energy*, *Skewness*, *Fast Fourier Transform coefficients*, etc. were extracted. Additionally, *Correlation*, *Co-variance* and *Signal Magnitude Area (SMA)* were also calculated for the acceleration data channels. The complete list of features is illustrated in the Table 1.

Table 1: The name of the features and the number of each feature extracted from a window.

Time Domain Features		Frequency Domain Features	
Name of feature	Number	Name of feature	Number
Minimum	8	Max Power	8
Maximum	8	Entropy	8
Standard Deviation	8	Center Frequency	8
Average	8	Peak Frequency	8
Variance	8	Energy	8
Peak to Peak Range	8	Skewness	8
Max Rate of Change	8	Kurtosis	8
Average Rate of Change	8	Number of Peaks	8
Standard Deviation of Rate of Change	8	First 15 FFT coefficients	120
Mean Absolute Deviation (MAD)	8		
Inter-Quartile Range	8		
Auto-correlation	8		
Mean Crossing Rate	8		
Linear Velocity	8		
Kinetic Energy	8		
Signal Magnitude Area	8		

3.3 Feature Selection

All of the calculated features might not be useful for the purpose of classification. Some of the features had a negative effect on performance. In other words, including those features in the classification process decreased the classification accuracy. Moreover, there were some redundant features that had no effect on performance. A feature selection framework is required to effectively remove these negative and redundant features. To achieve this we have used 6 different feature scoring techniques for each feature.

- **Mutual Information [6]:** The measurement of dependency between two features which is calculated based on the entropy estimation from k nearest neighbors distances. This metric will provide if any two feature is independent of each other. A higher value means higher dependency.
- **Chi-square(χ^2) Test [15]:** Measures the χ^2 statistics between two non-negative features. Irrelevant features that are most likely to be independent of the class get a lower score in this process.
- **Tree-based Selection [9]:** This approach uses Decision Tree-based classifiers to calculate the importance of each feature.
- **Pearson Correlation Coefficient [5]:** Highly correlated features are most likely to be irrelevant. This approach calculates the correlation coefficient and p-value to rule out the redundant features.
- **Spearman Correlation Coefficient [20]:** This is a non-parametric measure of monotonicity between two features. Spearman correlation does not consider the feature vector to be normally distributed.
- **ANOVA F-value [8]:** An ANOVA test is performed on the dataset to calculate the f-value and p-value for the features and importance values are calculated based on that.

After calculating those six scores for each feature, we have taken the average values for all individual features across N windows and picked the $q \leq p$ best features based upon an experimentally calculated threshold value ϵ. Basically, these feature scores define how much significance that feature adds to our classification. In this case, we have found that $q = 180$ features are equal or above of our selected threshold score ϵ. We finally get our final set of 180-dimensional feature vectors $\{X_n\}_{n=1}^{N}$ which is fed into the classifier for training.

3.4 Classification

Given the feature vector X_n' of a test window, our goal is to assign a label $\hat{y}$ by learning a classifier on the training set $T = \{(X_n, y_c) : n \in [1, N]$and $c \in [1, 12]\}$ In this experiment, we have exploited random forest for classification. Random forest is a popular and important machine learning algorithm. It is well-known for providing a good general predictive performance with low over-fitting, and easy interpretability which ensures the fact that it is straightforward to determine the significance of each feature on the tree decision. In other words, it is easy to compute how much each feature is contributing to the decision.

4 OVERALL RESULTS AND DISCUSSION

This dataset is very much challenging due to missing labels, shuffled data, variable window size and unstructured data pattern. In other words, its a real world data where there are lots of uncertainty and noise in dataset. We have cleaned up data and resampled them before feature extraction. Using Random Forest Classifier(number of trees= 70, maximum depth = 100, number of features = 200) our method got 78% validation accuracy and 75% weighted average of F1 score. Figure 6 shows the confusion matrix of our classification model on validation set. There we can see mostly data have been predicted as class 5, the main reason behind this prediction is our validation set. As we have validated our model on field data and from Figure 3 we can see that field data is mostly composed of class 5 activity. Therefore, our validation result predicts most of the activity as class 5.

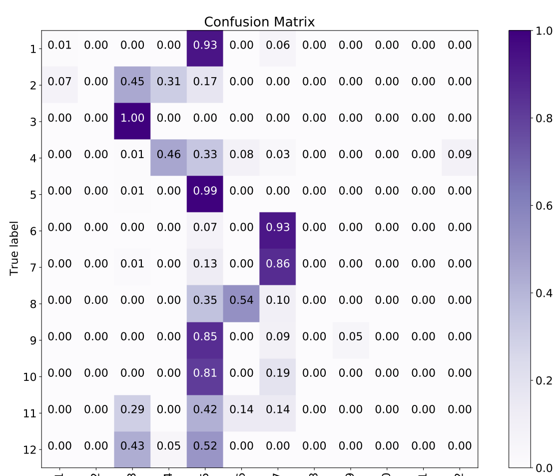

Figure 6: Confusion matrix

Our method has been trained on mostly lab dataset, some data of two activities have been taken from field dataset as those two activities were missing in lab dataset. Then, we have validated our dataset on the field dataset, which increases the possibility of good performance on the test set. Moreover, we have used a 95% overlapped sliding window of size 1000 on our resampled data, therefore, we have more data samples for training purpose. It has mitigated the lack of training samples as we have done a pseudo-data augmentation method here by taking a large overlap ratio. Furthermore, we have eliminated redundant features which ensures to have the best feature set for classification.

From Figure 3, we see the distribution of the activities in the dataset. Here, all the instances of individual activities in lab data are not the same, some activities have very less amount of data, Therefore, our model has not been trained uniformly, it may be biased towards some particular activity.

5 CONCLUSION

A very straight forward machine learning approach has been proposed in this paper to solve the problem of recognizing complex nurse activities. The method only uses accelerometer data collected from a mobile phone's sensor which is attached to the wrist of right arm of a nurse and classifies tasks into 12 activities. Our experimental result shows, this approach can provide promising results. However this method can be further improved using careful selection of window size and features and the overall performance can be more generalized by using Deep learning techniques. The recognition result for the testing dataset will be presented in the summary paper of the challenge [2].

REFERENCES

[1] Masud Ahmed, Anindya Das Antar, and Md Atiqur Rahman Ahad. 2019. An Approach to Classify Human Activities in Real-time from Smartphone Sensor Data. In *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 140–145.

[2] Sayeda Shamma Alia, Paula Lago, Kohei Adachi, Tahera Hossain, Hiroki Goto, Tsuyoshi Okita, and Sozo Inoue. 2020. "Summary of the 2nd Nurse Care Activity Recognition Challenge Using Lab and Field Data". *Proc. HASCA* (2020).

[3] Anindya Das Antar, Masud Ahmed, Mohammad Shadman Ishrak, and Md Atiqur Rahman Ahad. 2018. A comparative approach to classification of locomotion and transportation modes using smartphone sensor data. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 1497–1502.

[4] Ling Bao and Stephen S Intille. 2004. Activity recognition from user-annotated acceleration data. In *International conference on pervasive computing*. Springer, 1–17.

[5] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise reduction in speech processing*. Springer, 1–4.

[6] Alexander Bulinski and Denis Dimitrov. [n.d.]. Statistical estimation of the Shannon entropy. ([n. d.]). arXiv:math.ST/http://arxiv.org/abs/1801.02050v1

[7] Liming Chen, Jesse Hoey, Chris D Nugent, Diane J Cook, and Zhiwen Yu. 2012. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 6 (2012), 790–808.

[8] Nadir Omer Fadl Elssied, Othman Ibrahim, and Ahmed Hamza Osman. 2014. A Novel Feature Selection Based on One-Way ANOVA F-Test for E-Mail Spam Classification. *Research Journal of Applied Sciences, Engineering and Technology* 7, 3 (jan 2014), 625–638. https://doi.org/10.19026/rjaset.7.299

[9] K. Grabczewski and N. Jankowski. 2005. Feature selection with decision tree criterion. In *Fifth International Conference on Hybrid Intelligent Systems (HIS'05)*. IEEE. https://doi.org/10.1109/ichis.2005.43

[10] Md Nazmul Haque, Mahir Mahbub, Md Hasan Tarek, Lutfun Nahar Lota, and Amin Ahsan Ali. 2019. Nurse Care Activity Recognition: A GRU-based approach with attention mechanism. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 719–723.

[11] Zhenyu He and Lianwen Jin. 2009. Activity recognition from acceleration data based on discrete consine transform and SVM. In *2009 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 5041–5044.

[12] Sozo Inoue, Sayeda Shamma Alia, and Paula Lago. 2020. Nurse Care Activities Datasets: In laboratory and in real field. https://doi.org/10.21227/JEM3-AP07

[13] Sozo Inoue, Naonori Ueda, Yasunobu Nohara, and Naoki Nakashima. 2015. Mobile activity recognition for a whole day: recognizing real nursing activities with big dataset. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1269–1280.

[14] Wenchao Jiang and Zhaozheng Yin. 2015. Human activity recognition using wearable sensors by deep convolutional neural networks. In *Proceedings of the 23rd ACM international conference on Multimedia*. 1307–1310.

[15] Xin Jin, Anbang Xu, Rongfang Bie, and Ping Guo. 2006. Machine Learning Techniques and Chi-Square Feature Selection for Cancer Classification Using SAGE Gene Expression Profiles. In *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 106–115. https://doi.org/10.1007/11691730_11

[16] Md Eusha Kadir, Pritom Saha Akash, Sadia Sharmin, Amin Ahsan Ali, and Mohammad Shoyaib. 2019. Can a simple approach identify complex nurse care activity?. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 736–740.

[17] Paula Lago, Sayeda Shamma Alia, Shingo Takeda, Tittaya Mairittha, Nattaya Mairittha, Farina Faiz, Yusuke Nishimura, Kohei Adachi, Tsuyoshi Okita, François Charpillet, et al. 2019. Nurse care activity recognition challenge: summary and results. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 746–751.

[18] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. 2006. A practical approach to recognizing physical activities. In *International conference on pervasive computing*. Springer, 1–16.

[19] Jani Mantyjarvi, Johan Himberg, and Tapio Seppanen. 2001. Recognizing human motion with multiple acceleration sensors. In *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, Vol. 2. IEEE, 747–752.

[20] Leann Myers and Maria J Sirois. 2004. S pearman Correlation Coefficients, Differences between. *Encyclopedia of statistical sciences* (2004).

[21] Nazmus Sakib Patwary, Protap Kumar Saha, and Ifthakhar Ahmed. 2019. Nurse care activity recognition challenge using a supervised methodology. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. 788–792.

[22] Lin Sun, Daqing Zhang, Bin Li, Bin Guo, and Shijian Li. 2010. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations. In *International conference on ubiquitous intelligence and computing*. Springer, 548–562.

A APPENDIX

Sensor Modalities Accelerometer sensor channels in x,y and z axis.

Features Used Time domain: Min, Max, Standard, Deviation, Average, Variance, Peak to Peak Range, Max Rate of Change, Average Rate of Change, Standard Deviation of Rate of Change, Mean Absolute Deviation (MAD), InterQuartile Range, Autocorrelation, Mean Crossing Rate, Linear Velocity, Kinetic Energy, Signal Magnitude Area (SMA)

Frequency domain: Max Power, Entropy, Center Frequency, Peak Frequency, Energy, Skewness, Kurtosis, Number of Peaks, First 15 FFT coefficients

Programming Language and Libraries Programming language: Python

Libraries: Numpy, Pandas, Matplotlib, Scikit-learn, Sci-py

Window size and Post processing Window size: 1000

Post processing: N/A

Training and testing time Training: 150 seconds

Testing: 1.5 second

Machine Specification RAM: 13 GB, CPU: Dual-core 2.3 GHz, GPU: N/A