[Type the company address]

2021

**Activity recognition with healthy older people using a batteryless wearable sensor Data Set**

REPORT

Absas-LE0379

[Type the company name]

Table of Contents

[Abstract 2](#_Toc88491706)

[Introduction 3](#_Toc88491707)

[PROBLEM DEFINITION 4](#_Toc88491708)

[Literature review 5](#_Toc88491709)

[Dataset description 6](#_Toc88491710)

[Results and Discussions: 7](#_Toc88491711)

[Exploratory investigations : 7](#_Toc88491712)

[Machine Learning work: 9](#_Toc88491713)

[Performance measure 9](#_Toc88491714)

[Summary or Conclusion: 11](#_Toc88491715)

[References 11](#_Toc88491716)

# Abstract

Human Activity Recognition (HAR) has been one of the booming scientific topics in the last few decades. It helps scientists to collect data through battery-less sensors or, we can say other multimodal activity recognition methods to perform different scientific experiments to determine different human aspects like postural transitions and other important health variables.

# Introduction

Human Activity monitoring has become a vital area of research in the health care domain. The rise in popularity of smart wearable devices like smartwatches with embedded sensors has facilitated collecting high-quality data both easily and effectively. This area of research is highly intriguing as it finds applications across a wide range of domains. Some of the interesting applications include monitoring the physical activity and health condition of the geriatric population, predicting the robot's motion using sensors, developing systems that help the older adults walk, etc. Human Activity Recognition is nowadays an active research field that aims to understand human behaviour by interpreting sensory information gathered from people and the environment they live in (Reyes-Ortiz JL., 2014). These days, one of the most active ways of recording human activities are handheld or wearable smart devices, which are equipped with multiple sensors to record the combination of different Human activities. Sensors have a big role in making smartphones or wearables more functional and aware of the environment. This makes it possible to collect vast amounts of information about users' daily lives and activities (E. Bulbul, 2018).

Moreover, one of the benefits of today's smartphones or other devices developments is that they combine inertial sensors such as accelerometers, gyroscopes, magnetometers, and so on that can be used to detect human activities (Md Atiqur Rahman Ahad, 2021). The underlying motivation behind the project is to contribute to healthcare, gaming, and exergaming such as full-body motion-based games (O. C. Ann, 2014). Also, HAR is used widely for monitoring the activities of elderly people staying in rehabilitation centres for chronic disease management and disease prevention (Yoshimitsu, 2014). Our solution approach is based on classification models.

The primary objective of this project is to come up with an innovative and robust system to monitor human activity and to classify the positioning of a user into one of the four classes, Sitting on the bed, Sitting on the chair, Lying on the bed, Ambulating, where ambulating includes standing, walking around the room., using a batteryless wearable sensor. The idea is to model this as a learning problem given the quality data of human activity belonging to the four classes mentioned above. The data has been collected to train the models and build inference systems for predicting unobserved data sets. The experiments are based on the battery-less wearable sensor data collected by four different users. The battery-less wearable users contributed a couple of hours of data for each activity. The battery-less wearable is embedded with highly precise sensors like Accelerometer, Gyroscope, sensors for measuring the orientation, recording gravity, step count, rotation motion etc. The signals captured by these sensors are well indicative of the hand motion and enable us to predict the user's activity. The experiments were conducted on machine learning models like naïve Bayes and random forest. We have found some interesting observations such as, an increasing number of sensors for tracking the activity and considering larger time windows to improve the model's accuracy.

# PROBLEM DEFINITION

Human activity recognition is the problem of human body gestures or motion via sensors and determining human activity or action. This project uses different machine learning models to discover multiple human activity patterns and specifically detect which activity a person is doing through signals received by sensors, by analyzing the patterns in the results through a predictive model, and design and generating individualized output results to further predict the accuracy of the results.

# Literature review

In this section, we review some of the research advances in the area of activity prediction. Tapia et al. have focused on activity recognition in a home setting by using simple sensors installed in home environments. The motivation behind this work is to provide a system for medical professionals to monitor the changes in the daily activities of elderly patients, which would be useful to identify and diagnose the tumour at an early stage. Most works in this domain monitor human activity by attaching sensors to any of the body parts. There are a class of activities like grooming, brushing, cooking etc., which may be more easily recognized not by watching for patterns in how people move but instead by watching for patterns in how people move things. So, placing sensors on the things that humans interact and using it for activity monitoring was the main focus. They extracted features from the existing environmental state-change sensors and recognized the activity by applying naive Bayes classifiers on top of the features. Though this is a new direction, the activities of our task walking, sitting etc. does not generalize well to this domain.

Liu et al. studied computational methods for estimating energy expenditure in human physical activities. They observed that accelerometer sensors are insufficient for computing the energy expenditure as the activities with similar accelerations may have different energy profiles. So, they found that the works that use physiological sensors such as heart rate sensors and skin temperature sensors have a better performance. They observed the following as important features to be extracted from the data. 1) Time-domain features from the sensors like the mean, standard deviation, median statistics, signal power, etc. 2) Frequency domain features like FFT to calculate the dominant frequency corresponding to the highest amplitude in the signal. 3) They also found that demographic features like age, location to be important for this task. They stated that linear and non-linear regression models like Support vector regressors perform at high accuracy for estimating the energy expenditures.

Rosenberger et al. focused on estimating activity and sedentary behaviour and compared the effectiveness of placing the accelerometer at the wrist to placing it at hip for this task. Healthy adults wore triaxial accelerometers on the hip and dominant wrist along with a portable metabolic unit to measure energy expenditure during various activities. They used area under the ROC to differentiate the activity from the sedentary behaviour. Their experiments concluded that placing an accelerometer at hip is better to estimate energy expenditure than to place it on a wrist. But the problem with this approach is that usual smartwatch won't be suitable for this approach. We need to use special sensors that give us the ability to perform this task. So far, we have described different directions of works in this domain. Now, we will discuss some of the important works that used significant features from raw data for classification tasks.

# Dataset description

The data is Activity recognition with healthy older people, they are using a battery-less wearable sensor which gives 75128 observation. There are nine features,

1: Time in seconds

2: Acceleration reading in G for frontal axis

3: Acceleration reading in G for vertical axis

4: Acceleration reading in G for lateral axis

5: Id of Sensor reading sensor

6: Received signal strength indicator (RSSI)

7: Phase

8: Frequency The class variable is,

9: Activity, 1: sit on bed, 2: sit on the chair, 3: lying, 4: ambulating

Only one person's activities will be analyzed, a 60-year-old female.

**Workflow:**

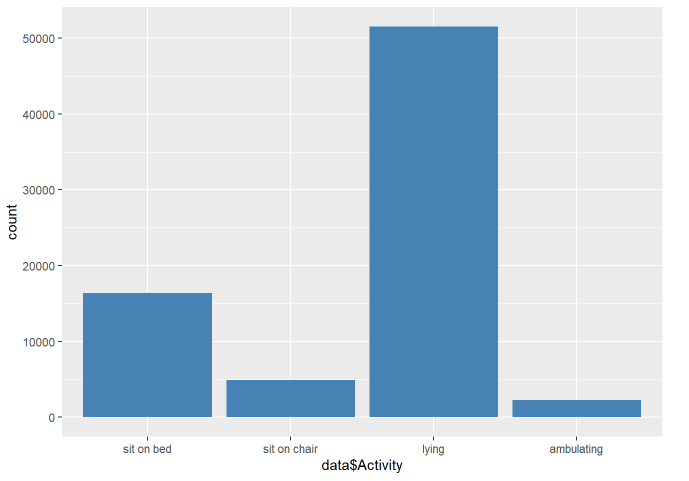
1. Import required libraries
2. Import dataset
3. Preprocess dataset
4. Exploratory data analysis
5. Split dataset into train and test
6. Build classification model
7. Evaluate model

# Results and Discussions:

## Exploratory investigations :

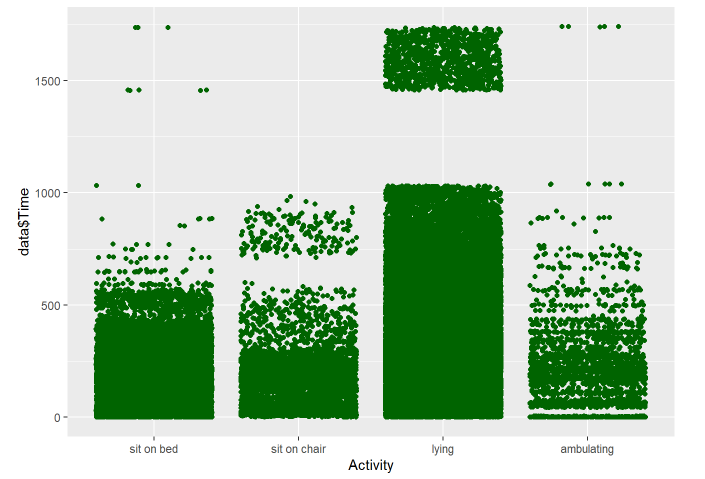
In EDA, we found that there is no missing values or duplicate rows in the dataset.

**Distribution of Activity:**



In this plot, we can see most of healthy older people doing lying activities.

1. **Distribution of Activity by Time:**



In this scatter plot, we can see the Activity distribution by Time. Most of older people spend Time on lying activity.

## Machine Learning work:

We have built two machine learning model to classify the activity:

**Random forest Model:**

Random Forest is an ensemble technique in which every tree in the ensemble is built from a sample drawn with replacement from the training set. During the decision tree construction, the splitting of nodes is based among the random subset of features and not chosen from all features.

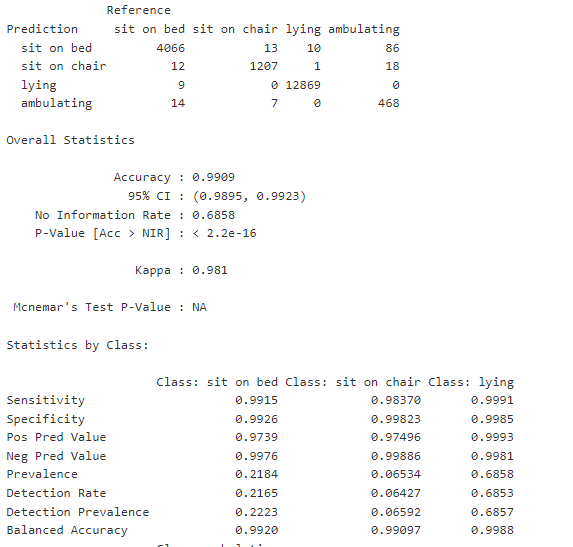
**Naïve bayes model:**

Naive means the algorithm used to classify objects is 'naive' or uniformed, making assumptions that may or may not be correct. We have used to classify the activity, and we get the best accuracy.

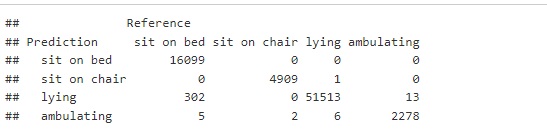
## Performance measure

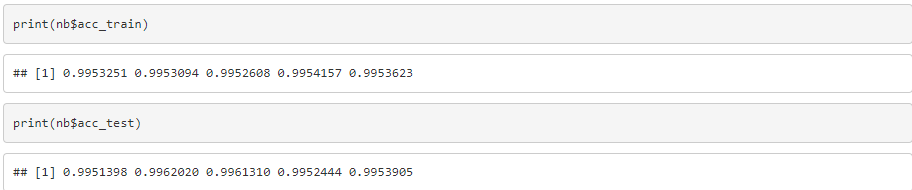
We use accuracy to quantify the performance of our models after considering the following reasons: as per the problem statement, we are only interested in predicting each class equally and accurately without preferring one above the others, hence discrediting the purpose of recall and precision metrics. This problem is a multiclass classification, where accuracy is more commonly used and more interpretable than ROC-AUC metrics. Model evaluation is the integral component of a project, and it targets to estimate the generalization accuracy of a model on future data. Model evaluation metrics are required to quantify the model performance. In this project, we focus on supervised classification learning models: Random Forest, Naïve Bayes. For evaluation, we have used classification accuracy, confusion matrix and precision call. Accuracy for classification models is as given below.

Random forest model evaluation result:



Naïve Bayes model evaluation result:





Random Forest & Naive bayes are wins as the best model. Random Forest is able to recognize human activities based on their behaviour with an outstanding 99% accuracy.

# Summary or Conclusion:

In a previous study (Tapiaet et al, 2004), they extracted features from the existing environmental state-change sensors and recognized the activity by applying naive Bayes classifiers on top of the features. We also have applied naïve Bayes, its gives 99.5% accuracy to classify the activity. The human activity recognition database is built by healthy older people using a batteryless wearable for the activities of daily living. The values are captured by a waist-mounted using a battery-less wearable with embedded inertial sensors. These results not only allow these models to be efficient but also give enhanced results in terms of Human Activity recognition, Human Computer interaction, and other platforms where Machine learning is necessary. However, there is always room for improvement, and we believe our best result can be made better as we progress in technology and data. We believe that by adding more features like Time when the activity is performed, we can draw more valuable insights and predict the activities.

# References

Reyes-Ortiz, J. L., Oneto, L., Ghio, A., Samá, A., Anguita, D., & Parra, X. (2014, September). Human activity recognition on smartphones with awareness of basic activities and postural transitions. In *International conference on artificial neural networks* (pp. 177-184). Springer, Cham.

Bulbul, E., Cetin, A., & Dogru, I. A. (2018, October). Human activity recognition using smartphones. In *2018 2nd international symposium on multidisciplinary studies and innovative technologies (ismsit)* (pp. 1-6). IEEE.

Md Atiqur Rahman Ahad, Anindya Das Antar and Masud Ahmed Series: Intelligent Systems Reference Library, Year: 2021, Volume 173, Page 7.

O. C. Ann and L. B. Theng, "Human activity recognition: A review," 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), Batu Ferringhi, 2014, pp. 389-393, doi: 10.1109/ICCSCE.2014.7072750.

Yoshimitsu, K., Muragaki, Y., Maruyama, T., Yamato, M., & Iseki, H. (2014). Development and Initial Clinical Testing of "OPECT." Operative Neurosurgery, 10(1), 46–50. doi:10.1227/neu.0000000000000214

Tapia, E. M., Intille, S. S., & Larson, K. (2004, April). Activity recognition in the home using simple and

ubiquitous sensors. In ​International conference on pervasive computing​ (pp. 158-175). Springer, Berlin, Heidelberg.

Liu, S., Gao, R., & Freedson, P. (2012). Computational methods for estimating energy expenditure in human physical activities. ​Medicine and science in sports and exercise​, ​44​(11), 2138.

Rosenberger, M. E., Haskell, W. L., Albinali, F., Mota, S., Nawyn, J., & Intille, S. (2013). Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. ​Medicine and science in sports and exercise​, ​45​(5), 964.