Activity Recognition and Quantifying Data Scarcity in Modeling Human Behavior for Healthcare Analysis

A Thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Tahera Hossain

Graduate School of Engineering

Kyushu Institute of Technology

Kitakyushu, Fukuoka

Thesis Committee

Professor Sozo Inoue

Professor Kenichi Asami

Professor Mitsunori Mizumachi

Professor Takeshi Ikenaga

Kyutech

Kyushu Institute of Technology

February, 2021

ABSTRACT

In this thesis, we proposed novel methods to analyze daily life activities of elderly people and quantifying data scarcity for modeling human behavior inside nursing care facilities in order to improve the overall care system. We addressed the existing challenges related to data collection, complex activity understanding and activity recognition by introducing different computational approaches that recover missing data and utilize spatial and temporal contexts of the data to achieve higher recognition accuracy in various modalities. As well as, in this thesis, we proposed a specific sample size determination method as a precursor to build more accurate models of human behavior. These proposed methods will be immensely helpful for the development of human activity recognition systems, mainly for elderly persons and nursing care facilities to improve human behavior understanding in real-life settings. The outline of the thesis is as follows:

In chapter 1 and 2, we briefly introduce the thesis work and literature review.

In chapter 3, we proposed a method which can improve activity recognition while having missing data without any data recovery. It is not explored or solved by others without any data imputation techniques. In our proposed approach, we explicitly induce different percentages of missing data randomly in the raw sensor data to train the model with missing data. Learning with missing data reinforces the model to regulate missing data during the classification of various activities that have missing data in the test module. This method can effectively improve the recognition accuracy from 80.8% to 97.5% in a developed synthetic dataset. Afterwards, we tested our approach with activities from two challenging benchmark datasets with a combination of 21 features. We achieved a better result of handling different levels of missing data in the dataset without any data imputation techniques.

In chapter 4, we introduce activity recognition using LoRaWAN that aim to sustain connections among IoT devices over a long distance. We explored low-power wide-area network (LPWAN) technology for sensing experiment to know the possibility of using LoRaWAN protocol and network for activity recognition. As well as, investigate the number of sensor nodes connected with a single gateway which have an impact on the performance of sensors ultimate data sending capability in terms of data loss. For human activity recognition, we have achieved recognition accuracy of 94.44% by Linear Discriminant Analysis (LDA), 84.72% by Random Forest (RnF) and 98.61% by K-Nearest Neighbor (KNN) from collected data. Later, in a real nursing care setting experiment, 42 LoRaWAN sensors environmental sensors data is collected to know the data loss ratio. We observe 5% data loss happened by the sensors with a single gateway. In a simulated environment, we checked the activity recognition performance with 5%, 30%, 50% and 80% data loss environment and have found recognition accuracy of 81.94% LDA, 80.55% RnF and 91.66% by KNN while 5% data are lost. Through our proposed framework it can open a new opportunity

to significantly increase the sensing range in nursing care center by LoRaWAN.

In chapter 5, we assess the daily life activity data obtained during the 4-months experiments at a nursing care center. Through this work, we tried to know whether the daily life nursing care activity data are dependent on subject (e.g., staff or target resident), or date, and whether the obtained data are meaningful and informative for activity understanding. We collected 38,076 activity labels, 46,803 record details, and 2834 hours of sensor data during this experiment. From this data, we revealed the varieties and dependency of activities and care details which can be a measure of any healthcare outcome. These findings can be essential elements for activity recognition, having many intra-class relationships among activities in nursing care center.

In chapter 6, we explore to recognize different types of head and mouth related activities of elderly people which can help to measure the health status. We propose an activity recognition framework to collect head and mouth related behavioral activities (e.g., head nodding, headshaking, eating and speaking) along with other regular activities. We have found that the accuracy is 93.34% by SVM, 91.92% by RnF, 91.64% by KNN, and 93.76% by CNN.

In chapter 7, we proposed a sample size determination method based on uncertainty quantification (UQ) for a specific Inverse Reinforcement Learning (IRL) model of human behavior. The main insight behind our method is that the probability of model parameters given training data can be updated from prior to posterior through Bayesian inference. For illustration, we provided an example with a specific hypothetical cost scenario and decision-making rule for MS (Multiple Sclerosis) behavioral modeling, under which our method indicated 935 samples is optimal.

Chapter 8 offers a brief discussion of the thesis. In Chapter 9, we conclude the thesis with conclusion and some future work issues.

This thesis contribution can help to create various tools to aid stakeholders, such as domain experts and end users, in exploring human behavior understanding. This work will become particularly important with the rise of behavior-aware user interfaces that automatically reason and act in response to people's behaviors in almost every aspect of their lives.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all those who have made this rewarding journey a meaningful experience. I wouldn't have been able to complete my doctorate without the support of so many people whom I want to thank now.

Firstly, I would like to thank my advisor Professor Sozo Inoue, for his unconditional support, mentorship and guidance to helped me stay on track, for directing my research to exciting challenges and applications. He has been an inspiring supervisor and has helped me grow as a researcher with his insightful suggestions, ideas and feedback. Your countless discussion, encouragement, motivation and trust always motivated me to take all new challenges throughout this doctoral program.

I would like to thank my thesis committee members Professor Kenichi Asami, Associate Professor Mitsunori Mizumachi and Professor Takeshi Ikenaga from Kyushu Institute of Technology for their all feedback to improve this work.

I would like to extend a special thanks to Professor Atiqur Rahman Ahad, with whom I first came to Japan for PhD research interview. I am indebted to him for spending countless hours for guiding me into this profession, settling my life in Japan and always being available with advice when I needed it most.

I would like to thank Assistant Professor Nikola Banovic, Electrical Engineering Computer Science and Director of Computational HCI Lab at the University of Michigan, for his support and opportunities to me to work directly in his lab for 3-months and continuous support and guidance for last one and half year. You gave me a new insight of my research. Also, I owe much to Assistant Professor Xun Huan at University of Michigan. Two of your guidance helped me to get the strength and courage to start work in a completely new direction into my research. Your valuable input about algorithm and dataset used in this work in Chapter 7. I would like to thank Wanggang Shen PhD candidate of University of Michigan for your all efforts to share workload and run the algorithm in a high computational resource. Also, like to thank Anindya Das Antar, Snehal Prabhudesai for being part of this research. Thanks to Nel Escher and Divya Ramesh for your friendship while I was at University of Michigan, Ann Arbor.

I like to thanks Paula Lago a postdoctoral researcher in my lab, for being like a sister to me in Japan. You guided me a lot for handling my professional and personal matters throughout this journey. Thanks to Paula Lago, Moe Matsuki, Nattaya Mairittha and Tittaya Marittha for being my travel partner for UbiComp and many international conferences where we joined together and enjoyed a lot. Your friendship brought me happiness during my time at Kyushu Institute of Technology.

I would like to thank all of my collaborators who have contributed to this research. I thank my lab mates over the years—Shingo Takeda, Shu Rin, Yusuke Doi, Hiroki Goto, Nour Al Bogha, Shafiqul Islam, Sayeda Shamma Alia, Farina Faiz, Shotaro Yoshinaga, Yusuke Nishimura, Haru Kaneko, Kohei Adachi, Brahim Benaissa, Defry Hamdhana, Muhammad Fikry and thanks to the members of the Sozo Inoue Laboratory with whom I had valuable discussions at seminars and monthly research presentation time.

I would like to thank the members of the nurse care facility of Sawayaka Club Inc., the nursing staffs, the subjects, and their families for their cooperation in collecting the data for this research in Chapter 4 and Chapter 5. I would like to thank the students of Sozolab who voluntary work for this experiment.

I am grateful to some Japanese friends who supported me a lot during my life in Japan. Specially, my Japanese local guardian Mr. Kenji Kurokawa and Mrs. Keiko Matsuoka for their all support and help.

To my elder sisters Shahera Hossain and Shakila Parveen, to teach me to survive a day by day.

To my special younger sister Tahia, you supported me a lot on the lonely days of this program. My brother Minhaz and other sisters - Tani and Tabia. You all will always be my little sisters and brother, how older you be.

Thanks to two little angles Rumaisa Fatima and Zubair Umar for giving me renewed energy to work at home during the last year of my doctoral program.

Last but not least, to my father and mother, for giving your life to educate me. I owe both to you for your unconditional love. You sacrificed a lot so that I could lead a better life.

Publications of this Thesis

Journal Articles (Peer Reviewed):

1. Tahera Hossain, M.A.R. Ahad, Sozo Inoue, "A Method for Sensor-Based Activity Recognition in Missing Data Scenario", Sensors, 20(14):3811, pp. 23 pages, 2020-07-08.

2. Sozo Inoue, Paula Lago, Tahera Hossain, Tittaya Mairittha, Nattaya Mairittha, "Integrating Activity Recognition and Nursing Care Records: The System, Deployment, and a Verification Study", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), Vol. 3, No. 86, pp. 24 pages, 2019-09-09.

3. MAR Ahad, TT Ngo, AD Antar, M Ahmed, Tahera Hossain, D Muramatsu, Y Makihara, Sozo Inoue, and Y Yagi, "Wearable Sensor-Based Gait Analysis for Age and Gender Estimation", Sensors, 20(8):2424, pp. 24 pages, 2020-04-24.

Book Chapters:

1. Md Shafiqul Islam, Tahera Hossain, Md Atiqur Rahman Ahad and Sozo Inoue, “Exploring Human Activities Using eSense Earable Device”, Activity and Behavior Computing, pp. 17 pages, Smart Innovation, Systems and Technologies, Vol. 204, Springer, 2021.

2. Haru Kaneko, Tahera Hossain, Sozo Inoue, “Estimation of Record Contents for Automatic Generation of Care Records”, Activity and Behavior Computing, pp. 19 pages, Smart Innovation, Systems and Technologies, Vol. 204, Springer, 2021.

Conference/Workshop Proceedings:

1. Haru Kaneko, Tahera Hossain and Sozo Inoue, “Implementation of Care Records Automatic Generation Function in a Care Record Application”, 11th Biennial Conference of the Asia-Pacific Association for Medical Informatics (APAMI), November 21, 2020, Japan.

2. Sayeda Shamma Alia, Paula Lago, Kohei Adachi, Tahera Hossain, Hiroki Goto, Tsuyoshi Okita, Sozo Inoue “Summary of the 2nd Nurse Care Activity Recognition Challenge Using Lab and Field Data”, 2nd Nurse Care Activity Recognition Challenge using Lab and Field, HASCA Workshop, (UbiComp’20), September 2020.

3. Md. Sadman Siraj, Md. Ahasan Atick Faisal, Omar Shahid, Farhan Fuad Abir, Tahera Hossain, Sozo Inoue, M.A R. Ahad “UPIC: User and Position Independent Classical Approach for Locomotion and Transportation Modes Recognition”, HASCA Workshop, Sussex-Huawei Locomotion (SHL) Challenge competition, (UbiComp’20), September 2020.

4. Tahera Hossain, Md Shafiqul Islam, Md Atiqur Rahman Ahad, Sozo Inoue, "Human Activity Recognition using Earable Device", ACM Int’l Conf. Pervasive and Ubiquitous Computing (Ubicomp’19), Poster, September 2019.

5. M Ahmed, AD Antar, Tahera Hossain, S Inoue, MAR Ahad,"POIDEN: Position and Orientation Independent Deep Ensemble Network for the Classification of Locomotion and Transportation Modes", HASCA Workshop, Sussex-Huawei Locomotion (SHL) Challenge competition, (UbiComp'19), September 2019.

6. SS Saha, S Rahman, ZRR Haque, Tahera Hossain, S Inoue, MAR Ahad, "Position Independent Activity Recognition Using Shallow Neural Architecture and Empirical Modeling", Sussex-Huawei Locomotion (SHL) Challenge competition, ACM Int'l Conf. Pervasive and Ubiquitous Computing UbiComp, (UbiComp'19), September 2019.

7. Thanh Trung Ngo, Md Atiqur Rahman Ahad, Anindya Das Antar, Masud Ahmed, Daigo Muramatsu, Yasushi Makihara, Yasushi Yagi, Sozo Inoue, Tahera Hossain, Yuichi Hattori, "OU-ISIR Wearable Sensor-based Gait Challenge: Age and Gender", 12th IAPR Intl. Conf. on Biometrics (ICB), June 2019, Greece.

8. Tahera Hossain, Sozo Inoue, "A Comparative Study on Missing Data Handling Using Machine Learning for Human Activity Recognition", Activity and Behavior Computing (ABC), May 30 2019, USA.

9. Sozo Inoue, Tittaya Mairittha, Nattaya Mairittha, and Tahera Hossain, " Integrating Activity Recognition and Nursing Care Records: the System, Experiment, and the Dataset", Activity and Behavior Computing (ABC), May 30 2019, USA.

10. Tahera Hossain, Sozo Inoue, "Sensor-based Daily Activity Understanding in Caregiving Center", Ph.D. Forum, 17th IEEE International Conference on Pervasive Computing and Communications (PerCom), Kyoto, Japan. March 2019.

11. Tahera Hossain, "Activity Recognition and Wireless Sensor Network Optimization", Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, Pages 1767-1770, Broadening Participation Workshop (BPW), Singapore, (UbiComp'18), October 2018.

12. Tahera Hossain, Tahia Tazin, Md Atiqur Rahman Ahad, Sozo Inoue, "Activity Recognition by Using LoRaWAN Sensor", ACM Int'l Conf. Pervasive and Ubiquitous Computing (Ubicomp'18), Poster, October 2018, Singapore.

13. Tahera Hossain, Yusuke Doi, Tahia Tazin, Md Atiqur Rahman Ahad, Sozo Inoue, "Study of LoRaWAN Technology for Activity Recognition", ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2018 International Symposium on Wearable Computers (UbiComp-ISWC'18), Workshop on Human Activity Sensing Corpus and Applications (HASCA), October 2018. Singapore.

14. Swapnil Sayan Saha, Shafizur Rahman, Tahera Hossain, Sozo Inoue, Md Atiqur Rahman Ahad, "Supervised and Neural Classifiers for Locomotion Analysis", ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2018 International Symposium on Wearable Computers (UbiComp-ISWC'18), (Sussex-Huawei Locomotion (SHL) Challenge competition), Singapore, October 2018.

15. Tahera Hossain, Hiroki Goto, Md Atiqur Rahman Ahad, Sozo Inoue, "A Study on Sensor-based Activity Recognition Having Missing Data", 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR), Kitakyushu. Japan, June 2018.

16. K Khabir, M Siraj, M Habib, Tahera Hossain, and M Ahad, “A Study on DSR Routing Protocol in Adhoc Network for Daily Activities of Elderly Living”, 7th Intl. Conf. on Informatics, Electronics Vision, Japan, June 2018.

17. Tahera Hossain, Hiroki Goto, Sozo Inoue, "Improving Activity Recognition for Missing Data", 9th EAI International Conference on Mobile Computing, Applications and Services (MobiCASE) Student Workshop, Osaka. Japan, March 2018.

18. H Goto, Tahera Hossain, and S Inoue, “Improving Accuracy for Missing Sensor Data in Activity Recognition with Smartphones”, DICOMO2018 Symposium, July 2018.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Automatic recognition of physical activities or human activity recognition (HAR) has emerged as a key research area in mobile and ubiquitous computing [38, 42]. The purpose of HAR is to understand people's everyday behaviors through the study of data collected from individuals and their neighboring living environments. HAR is important for assisted living [39, 95], healthcare monitoring system [36], fall detection, and elderly support system [77, 158]. The advent of smartphones and wearables has enabled a wide application spectrum for HAR including human-computer interaction, fitness tracking, and transportation mode recognition [22, 80, 136, 58] while Internet of Things (IoT) can be thought as of a large scale multi-modal sensing system consisting of devices at the edge (e.g. embedded and wearable computing devices, mobile computing) and also devices on the clouds. This ability to collect a large amounts of human behavior traces data collected from several different sensors on people' personal, mobile, and wearable devices, as well as from smart environments, offers a new direction of data to study human behavior at large scale. Vision-based approaches to action identification have been well established for more than a decade, but improvement is not enough [43, 79]. Plenty of methods [11, 2] are dealing different kinds of actions but little work has been done on elderly people's normal and critical activity detection for domain specific manner. Although HAR systems based on video camera are common for various security applications, in smart environments they pose numerous challenges related to privacy and space limitations. As vision-based sensors (cameras) is regarded as exposing privacy so users are more reluctant to their use inside home. That is why for this thesis we concentrated on non-vision-based sensing.

In this regard, wearables open up new possibilities for the identification of different forms of human behavior [4]. The wearable offers good scope for the assessment of various forms of human activity without having any privacy issues. There are various types of sensors available in wearable devices. In particular, human movement data can be obtained from an accelerometer and a gyroscope sensor. Such sensor are worn on different parts of the body and provide 3-axis acceleration and orientation, respectively. Afterwards, machine learning algorithms can be used to detect different types of human activities on these data [101]. It is easier to detect and identify different human motions by accelerometer and gyroscope sensors. In this regard, smartphones have a vital role to play in contributing to this area. Different sensors are already installed on a smartphone, such as a three-axis acceleration sensor that adjusts the data based on different movements, and an angular velocity sensor (also known as a gyro sensor) that measures an object's angle (posture), angular velocity, or angular acceleration. There is a considerable amount of sensor data due to continuous monitoring and data collection that carries significant information for study. Focusing these issues, researchers face a variety of problems related to proper data processing, data annotation, misclassification, etc.

Though, the ever growing elderly population through world-wide [107] increases the demand of care-giving center, assisted living center. The demand of working staff at the healthcare center/nursing care center are also increasing to help and support elderly person good health. Advancement of technology can help to monitor patient activity in healthcare center where acting on patient behaviors could improve their healthcare outcomes. In this regard, it is essential to improve domain specific human activity understanding [94]. Existing human activity recognition system mostly focus to recognize regular activity like walk, sit, run, stair up/down etc . and data are collected in controlled laboratory environment [21]. In care-giving center, it is aim to recognize daily activities such as taking bath, taking medicine, cooking, taking food and toilet using, as these activities will provide an indication of health status [71]. In conventional way, staff record patient information by using handwriting recording system. Mobile application in smartphone [104] play a vital role here to make the recording system automated. It helps to get large amount of annotated data from real world nursing care staffs. In nursing facilities, it is essential to have patient activity data and care records data. Care records data have details information of patient daily activities such as amount of food intake, medical assistance details, bathing method and different assistance demands which they need from care staff. Care staff for doctor need these care records information of patient to monitor their health condition or predict the need of future assistance for a particular patient.

In this regard, recent research focuses on improving the accuracy of activity recognition through feature selection [118, 12], improved segmentation of data or the use of state-of-the-art learning algorithms [119, 81] in domain specific manners. However, because of the stochastic nature of human actions, the accuracy of these approaches remains unsuitable for real life applications. The lack of actual test data and the complexity of moving information from one environment to another owing to the great heterogeneity between individuals. The need for large quantities of labeled data and the selection of features that are highly dependent on the sensors used make it difficult for the user to use these models in real-world settings.

In this regard, this thesis main contribution is to address real life activity recognition challenges and propose prospective solutions for those. We investigate that there is huge difference among real life activity recognition [148] from real nursing care data and activity recognition in controlled environment laboratory data [85, 94, 8]. Real-life activity recognition and assessment is not straightforward. Class imbalances, overlapping activity and wrong timestamps are involved while data collected in wild. Also, recognition of regular activity and complex activity from a real-life situation is very challenging [98]. Because the development of context recognition systems needed significant amount of labeled sensor data to train the models. Data collection involves missing sensor data, poor network coverage in nursing care facility. As well as the nature of data in application specific domain is unknown. Furthermore, it is also important to know the amount of data while modeling patient behavior. These problem findings motivate the need for strategies that can improve the activity recognition in nursing care domain. Thus, in this thesis, we address few challenges from real life nursing care activity recognition and give proposal to solve those.

1.2 Problem Definition

There are many challenges involved in real-life nursing care activity understanding. In this thesis, we have addressed the existing following challenges related to data collection, complex activity understanding and activity recognition by introducing different computational approaches in various modalities for better activity understanding in nursing care domain.

- **Problem 1.** ***Sensor data is missing***, while collected real field data for activity recognition. Also, missing pattern is important to know to propose a method for activity recognition improvement.

- **Problem 2.** ***No strong technological infrastructure and weak network coverage for data collection***, it is needed to significantly increase the sensing range in nursing care center for proper activity data collection.

- **Problem 3.** ***Important to know the nature of domain specific activity***, in the application of recognition of nursing activities, activity classes—the categories of activities—are specified in a domain-specific manner. We need to now the nature and dependencies of the nursing care activity for both caregivers/staffs and residents/elderly people.

- **Problem 4.** ***Head and mouth related complex personal scale behavioral activities are important to measure for the full records of patients***, understanding and recognize intaking activities are also very important for nurse care of elderly people.

- **Problem 5.** ***How much data to collect for modeling human behavior? – is an important question to know for modeling patient behavior,*** while data collection is a significant burden to patients/elderly. With having a promising computational model, it will be possible to predict current days symptoms and future days symptoms for any critical patient. But it is important to know the amount of samples needed (dataset size) for model designers.

1.3 Outline of the thesis

The rest of this thesis is organized in the following chapters:

- In Chapter 2, presents a detailed literature review and motivation of this thesis. This include the different area and challenges in nursing care center which are key issues to solve in real-life activity recognition field.
- In Chapter 3, we propose a novel method to improve activity recognition while having missing data without any data recovery. For the missing data pattern, we considered data to be missing in a random pattern, which is a realistic missing pattern for sensor data collection.
- In Chapter 4, we proposed an activity recognition framework by exploiting LoRaWAN (Long Range Wide Area Network) protocol for Nursing Care for improving the sensing capacity with proper activity recognition.
- In chapter 5, we discuss and integrate activity recognition and nursing care records to revealed the nature of nursing care activities from real-life data. We tried to know the staff activities as well as elderly peoples activities from 4-months real nursing care experiments.
- In chapter 6, we propose a framework to detect head and mouth related personal scale behavior activities which activities are difficult to detect with usual wearable sensors with other regular activities.
- In chapter 7, we proposed a Bayesian approach for quantifying data scarcity to modeling human Behavior via Inverse Reinforcement Learning. Our proposed sample size determination method can be a decision making support to model designers to estimate the number of samples (dataset size as part of data collection) for modeling human behavior.
- Finally, in chapter 8, we conclude a brief discussion and overall findings and contribution of this thesis with a conclusion remarks and suggestion for future work.

Chapter 2

Background and Motivation

2.1 Introduction

Activity recognition has been researched for a long time now, research and implementations have concentrated largely on the recognition of regular physical activity while data are collected mostly in controlled laboratory environment. However, activity recognition of daily activity and complex activity from a real-life scenario is often very difficult. Many wearable computing devices can now identify, count steps and provide a general description of how involved we have been on a day when we walk, bike or run. However, research in behavior aware applications domains such as health care, nursing care, assisted living centers are not very mature [49]. One of the main reason is that it is more complicated and difficult to evaluate. In health care applications, complex activity recognition has concentrated on the recognition of patient activities or activities at home [46, 97, 144] and noticed nurses and caregivers. Recognition of nursing care activities (both nurses and residents) can have many applications. Particularly for daily activities understanding in nursing domain can be of great benefit for caregivers to checking compliance with care routines [74] for a given patient, identification of risk activities that require special care, creation of automated records to decrease documentation time [69], make decisions or diagnoses in emergency situations etc. Advancement of technologies can help track patient behaviour in healthcare centers where patient behavior can enhance their health outcomes. But human activities are complex and dynamic [110]. Activities may be carried out simultaneously at various levels of granularity [108]. Also it is complicated due to the lack of publicly accessible data [94, 9].

On the other hand, the difficulty of deploying learning-enabled HAR systems in the wild (real life) involves coping with data from multiple noisy sensors with varying sampling rates, misaligned timestamps and missing data. In addition, there is no good technical infrastructure in the nursing care center and it could be in retired areas with poor network coverage for data collection. Since the development of context recognition systems required a large amount of labeled sensor data to train the models so robust data collection system is important. Despite of this, sensor data is difficult to interpret and annotate after data collection, making it difficult and expensive to produce large training sets. It is also often necessary and crucial to know the appropriate amount of real field data to predict patient behavior precisely while collecting patient data is a significant burden to patients.

One of the domains that can benefit immensely from activity recognition is the nursing domain, but it has not been researched because of many challenges [83, 48]. Thus, our focus is on to addresses the real settings problem of nursing care centers and proposed prospective solution for those. In these contest, in the following sections we will elaborately describe few challenges of real-life activity recognition in nursing care centers.

2.2 Real field activity recognition challenges

2.2.1 Missing Sensor Data

Sensor data loss scenario is common issue in wireless sensor network while various sensor data traverse from one source to another [61, 163]. Human activity recognition will not possible properly if we will not able to evaluate all sensors data properly. In the real-world scenario, assisted living facility centers have big challenges to collect all sensor data properly. In certain cases, this missing data is often extreme, which brings a major challenge to sensor data applications. However, conventional methods of data estimation cannot be used directly in the wireless sensor network and existing algorithms of estimation do not provide satisfactory accuracy or have high complexity [150].

The missing value issue is prevalent in datasets [37, 120]. A significant amount of raw data will be lost if the missing data is directly discarded, which will decrease the quality and reliability of the results of the study and create a great loss of resources [47]. Data loss or the presence of incomplete data can be happen for low-battery power in sensors, longer distance between sensors and access points, failures of sensors to send data properly, hardware failures, synchronization problem, signal strength fading, packet collisions, weakness in Wi-Fi or network, environmental interference, etc. [57, 88, 16, 162]. It has a greater impact on the quality of the collected sensor data. Missing data are either removed, or missing values are replaced by the last non-missing values; however, it may contain important information, especially for activity recognition. Nevertheless, it is not a wise idea to just remove the missing content from the dataset. There are comparatively fewer works in the field of activity recognition domain that unambiguously addressed the problem of missing data [149, 156, 134]. These missing values create significant difficulties for data processing approaches such as classification, prediction and other machine learning methods [67]. It is often not possible to deal with missing values, particularly when the number of missing data is high.

In different areas of research, data interpolation techniques for missing data were proposed in [78]. Saeed et al. [133] proposed an approach for handling missing sensory features and realistic samples by adversarial autoencoder. They designed a fully-connected classification network for missing modalities. On the other hand, Pedro et al. [51] instituted four algorithms for handling missing data. In their work, they simulated various amounts of missing data. Note that the researchers mostly focused on data imputation techniques to recover missing data. Data imputation techniques are time-consuming for real settings.

During the identification of real settings sensor data loss pattern, it was observed that sensor data missing pattern follows a random missing pattern [161, 88, 51]. The nature of missing data is extremely unpredictable. Missing patterns can be distributed at any locations in the data. The distribution of missing regions is purely random; therefore, we did not compare with any uniform random model or clumped random model. Uniform random model follows some specific time and specific missing amount. Here, 'uniform' implies that the missing occurrence is evenly spaced, whereas, 'random' indicates random spacing, and 'clumped' decodes that the missing data are distributed in clusters. These are not realistic random models for sensor network's missing pattern.

Based on our study, there is almost no prior work that considers a realistic missing pattern in a dataset, as well as examining any missing values without imputation. Earlier, the effect of missing data situations, when missing values correspond to specific feature values, were studied. The impact of any possible good combinations of features for handling missing data for activity recognition were explored [62]. The randomness was varied at different levels to study the recognition results. In this thesis, we proposed a method to improve activity recognition performance having missing data in the dataset [62]. To demonstrate the robustness of our approach, initially, the method was evaluated on simulated data. After obtaining good results, we explored the method on two difficult benchmark datasets. Usually, most of the datasets were created using few subjects, , and not well adapted for real world applications. A project called "HASC Challenge" gathered a large-scale corpus of human activity data to resolve the situation. More than 6,700 accelerometer data with 540 subjects were gathered through this project through the cooperation of 20 teams. Hence, this dataset is indeed a difficult one to challenge [85]. On the other hand, another dataset we considered is a combination with normal and complex activity classes with varying class imbalance issue [30]. The nature of class-imbalance and complex activities are also aligned with real-settings application. Hence, we considered these two benchmark datasets to evaluate our approach.

2.2.2 Lack of strong technological infrastructure in nursing care

Human activity recognition with low strength sensor is another daunting task due to the lack of proper data collection in real life. There is no strong technological infrastructure in nursing care and they might be in retired places. One key limitation of current wireless sensing technologies is the small sensing range. Although the communication range of WiFi can be 20–50 meters indoors, its sensing range is only 3–6 meters due to the weak target-reflected signal used to sense target movement. Furthermore, existing wireless sensing solutions require at least one of the transceivers to be close to the target, which may not be feasible in certain scenarios for nursing care. LoRaWAN (Long Range Wide Area Network) is one of the major low-power wide-area network (LPWAN) technologies that aim to sustain connections among IoT devices over a long distance. Due to weak network coverage in nursing care. LoRaWAN can open a new opportunities to significantly increase the sensing range in nursing care center.

Machine-to-machine communication has progressed rapidly in healthcare settings. IoT-themed applications in care-giving center need to improve for providing smart care facilities in near future. However, IoT health care is not without its obstacles. IoT applications have particular criteria such as long-range, low data rate, low energy consumption and cost-effectiveness. Generally used short-range radio technologies (e.g., ZigBee, Bluetooth) are not suited to situations involving long-range transmission. Therefore, IoT applications requirements have driven the emergence of a new wireless communication technology: low-power wide-area network (LPWAN). LPWAN is increasingly gaining popularity in industrial and research communities because of its low power, long range, and low-cost communication characteristics [165]. It provides long-range communication up to 10–40 km in rural zones and 1–5 km in urban zones [146, 111]. Therefore, these promising aspects of LPWAN have prompted recent experimental studies on the performance of LPWAN in outdoor and indoor environments [14, 55]. In summary, LPWAN is highly suitable for IoT applications in healthcare monitoring service. This technology open up exciting new areas for data mining research and data mining applications.

However, the researches of LoRaWAN application is still limited. Recently, the research community started to publish their work on LoRa technology. Many studies have done about the indoor coverage of LoRaWAN [116, 139]. Regarding LoRaWAN coverage the author [122] analyze the coverage of LoRaWAN in a sub-urban area. A LoRaWAN scalability study is present in [18, 17]. Stable wireless sensor network is necessary for better healthcare service. It is estimated that there will be 50 billion connected devices by 2020. Therefore, the LPWAN technology can play a vital role for healthcare monitoring service in future. LoRaWAN has some special criteria that is LoRaWAN gateway is able to support up to 20,000 IoT devices. It can operate successfully at ranges exceeding 15 km in sub-urban settings and more than 2 km in dense urban environments.

Researchers have started proposing this new technology for human activity recognition [61]. But the scalability of the LoRaWAN network with multiple end nodes is a challenge to keep all nodes active at the same time. Packet loss and packet receive ratio is a crucial factor in healthcare services. In this thesis we used LoRaWAN network to send various sensor data for human activity recognition in nursing care domain. We explore LoRaWAN-based sensors to verify the activity recognition. In single device data collection, we observe no packet loss happens. Packet loss happens when multiple devices are connected at the same time. We study the relation between packet loss and accuracy in simulation environment. The objectives of our work is to verify the activity recognition by using LoRaWAN based network; to explore the relation between data loss and accuracy; and to estimate the data loss in realistic nursing care indoors experiment with many sensors. The result provides promising prospect for LoRaWAN sensor for improving healthcare monitoring service. We proposed this framework for healthcare monitoring system for nursing care data collection and activity recognition.

2.2.3 Understanding real life complex activity with proper data annotation

There are limited research especially cases used at hospitals and nursing care facilities [115, 15, 74]. There are many challenges in wearable sensors to explore it in some real fields [5, 99]. For machine learning algorithms, data with a training label is required while collecting label data is costly. Sensor data is also difficult to interpret. Sensor data collections are typically performed with several participants, and the data obtained from different participants for the same task set may not be of a similar nature. Therefore, the intra-class activity variations are prominent. In addition, the data relating to two separate activities can be identical in nature, resulting in inter-class variability. Class imbalance can occur if an activity is done for a longer time. These are some important nature of activities which we need to know for domain specific activity understanding. In controlled laboratory environment it is comparatively easy to control these types of crucial factors but it is most difficult in wild. On the other hand, it is difficult to choose the proper sensing modality and location [135] of the sensing that is used to monitor complex activities.

Through commercial developed products it is possible now to identify general summary of daily activities. It will give an overview of how involved we are on a day. But domain-specific applications like healthcare/nursing activities research are not mature yet. One of the main reason is domain specific human activities are complex to understand with proper data annotation. Furthermore, the nature of the domain specific activities is not clear. In the application of recognition of nursing activities, activity classes—the categories of activities—are specified in a domain-specific manner. Moreover, such activities have imbalance varieties, for example, the number of occurrences between classes, start times per day and length. Also, activities are performed based on the demand basis.

To resolve the above-mentioned challenges, it is important to evaluate or analyze real field data that can be used by researchers to propose and compare the performance of different methods [74, 71]. In this thesis, we suggest a framework that combines activity records and activity label records that are routinely used by workers in the nursing sector. While this can cause inaccuracy due to self-labelling, but it is possible to increase the number of label collections by easy recording. We analyze intra-class relationship of real life nursing care activity data which has important information to revealed the nature of real field data. Also, in this thesis, we introduce a new platform for eSense earables to recognize head and mouth related activities alongside with some regular activities.

2.2.4 Data Scarcity for Modeling Human Behavior

After having a promising computational model, quality and amount of training data is often the single most dominant factor that determines the performance of a machine learning model. Large amount of data play an important role for a machine learning model to be computationally efficient. Once we have enough training data then it can perform good. But exactly how much training data do you need? It depends on many factors. Mostly, it depends on the task we are trying to perform and the performance we want to achieve. Also, model complexity, input features we have, the noise in the data are important factors. Optimum size of the training data set is necessary to achieve high classification accuracy.

To date most approaches to scaling up machine learning to large data sets have attempted to modify existing algorithms to deal with large datasets in a more computationally efficient and effective manner. However, despite of being able to collect massive amount of data about people, there are data samples were data will be scarce. For example, medical domain where condition allows for collecting data from only a few patients. There is still desire to use machine learning to model even behaviors of those stakeholders. But it is not clear how much data we need to be confident that the model is capturing the behaviors correctly.

Data scarcity means few data points when it is difficult to get data or the data is small as compared to the amount needed. Data scarcity is the most obvious limitation for a machine learning and deep learning model. There are no precise rules for knowing how much data we need to support for building of a good model. If we feed a model poorly, then it will only give us poor results. Theoretically, big amount of training samples is always good for a model. There is no method to estimate exact amount of training samples which will be good for a model. In real world scenario, it is not always possible to have large amount of data. Specially in health care intervention where we can able to collect data only from few patients. Advancement of technology can help to monitor patient activity in healthcare center. This technology can be used to recognition of patient activity record, nursing care activities, nursing care work along with resident's care [104]. Analysis of these care records data can improve the support of the nursing facility center [70] and able to provide better life to patient. It helps to evaluate elderly peoples day to day health status. For this purpose, either sensor/video based system can help us to collect these type of patients regular life data and clinicians will able to analyze those data for behavior prediction. In this regard, it is always challenging to capture appropriate amount of real field data to predict patient behavior.

Although large amounts of good quality training data will result in a good model, collecting behavior instances is challenging. In this situation, if data is collected without deliberation, it might result in two possible scenarios - under collection (enough data is not collected) and over collection (more than necessary data is collected). In the first case, the model will not learn well if the data is scarce. Techniques like data augmentation are used in such scenarios, however, they cannot replace actual data. Data Augmentation approaches [123] is done under the assumption that more information can be extracted from the original dataset through augmentations. These augmentations artificially inflate the training dataset size by either data warping or oversampling. Using data augmentation it is possible to artificially expand the size of a training dataset by creating modified versions of training data in the dataset. Having more training data can result in more skillful models, and the augmentation techniques can create variations of the data that can improve the ability of the models. The performance of deep learning neural networks often improves with the amount of data available. Functional solutions such as dropout regularization [143], batch normalization [75], transfer learning [157] [140], and pretraining [45] have been developed for deep learning based analysis on smaller datasets. Conventional data augmentation methods are mainly developed for image and video analysis tasks. A key challenge for data augmentation is to generate new data that maintains the correct label, which typically requires domain knowledge. However, it is not obvious how to carry out label-preserving augmentation in some domains, e.g., wearable sensor data. Moreover, using data augmentation we can enlarge the training dataset size but there are no precise rules for knowing how much data we need to support for building of a good model. If we collect much less data, in such a case, one might have to go back into the field for more data collection. This poses several challenges - the same representative patients might not be found, the field might be far away, etc. On the other hand, collecting too much data would not result in a poor model, but rather waste precious resources like time. It might also be cumbersome for patients if wearable sensors are used to collect data.

In this thesis, we address this data scarcity problem for modeling human behavior. We propose a sample size determination method. Our approach can be use as a tool to help decision-making for model designers to selecting optimum amount of data when modeling patient behavior.

2.3 Contributions of the Thesis

One of the domains that can benefit immensely from activity recognition is the nursing domain. Domain specific activity recognition is till now limited because of many challenges. Although activity recognition in laboratory settings and real life settings has huge difference. To reduce this gap our focus is on to addresses the real settings problem of nursing care centers and proposed prospective solution for those. Therefore, we contribute to addressing a few real-life activity recognition challenges in this thesis.

The main contributions of this thesis are:

1. ***A Method for improving Sensor-Based Activity Recognition in Missing Data Scenario***

   We propose a novel method to improve activity recognition while having missing data without any data recovery. For the missing data pattern, we considered data to be missing in a random pattern, which is a realistic missing pattern for real field sensor data collection. In our proposed approach, we explicitly induce different percentages of missing data randomly in the raw sensor data to train the model with missing data. Learning with missing data reinforces the model to regulate missing data during the classification of various activities that have missing data in the test module [62]. This approach demonstrates the plausibility of the machine learning model, as it can learn and predict from an identical domain. We examined the method for different missing percentages, varied window sizes, and diverse window sliding widths [63]. We examined real settings nursing care dataset for evaluating missing pattern and missing sample ratio. As such, our method can improve activity recognition in nursing care data collected in wild in the presence of missing data. We explain this in *Chapter 3*.

2. ***Proposed an activity recognition framework by exploiting LoRaWAN (Long Range Wide Area Network) protocol for Nursing Care***

   We propose a framework for activity recognition by using LoRaWAN protocol and network to send various sensor data. This work explores exciting new opportunities to significantly increase the sensing range with the introduction of LoRaWAN for Nursing Care Center. After having a activity recognition framework in laboratory settings experiment. We examined the performance of the LoRaWAN in both laboratory environment and real nursing care environment. In LoRaWAN technology, the amount of sensor nodes connected with a single gateway have an impact on the performance of sensors ultimate data sending capability in terms of data loss. We investigated this issue from real nursing care data to check the feasibility of using LoRaWAN sensor for future healthcare monitoring center [64, 65]. We explain the proposed framework and results and findings in *Chapter 4*.

3. ***Integrating Activity Recognition and Nursing Care Records***

   In this thesis, we suggest a framework that combines activity records and activity label records that are routinely used by workers in the nursing sector. We have collected 4-months nursing care data from real nursing care center in Japan. It has been shown that by using our system it can possible to increase the number of label collections by easy recording. We analyze intra-class relationship of real life nursing care activity data which has important information to revealed the nature of real life data. We investigated the dependency of activities to staff users, target residents, and days. Using the obtained data, we revealed the nature of the data, including dependency of activities to several factors, and the nature of timestamps of self-labeling.

We investigated "staff activity" for support service requirements as well as "elderly people activity" from this data. This work also help to examined elderly people health status from the data and types of activities recorded by staff. In nursing care facility, activity recognition is challenging because the nature of activities is not clear. Hence, this thesis revealed the nature of real settings activity data for nursing care center [71, 69]. We explain this in Chapter 5.

4. ***Exploring Human Activities Using eSense Earable*** We can measure nurse activities, but patient's are also very important to make complete records, therefore in this thesis we propose a framework using eSense earable to detect head and mouth behavioral activities with good accuracy. We develop a smartphone application for data collection from the eSense. Through eSense earables, we performed experiments for seven activities. Four activities are related to head and mouth (namely, *eating*, *speaking*, *head shaking*, and *head nodding*). Three regular activities (i.e. *walk*, *stay*, and *speaking while walking*). Detecting these activities accurately using only accelerometer and gyroscope sensors data quite challenging. By using our propose framework [66], it will be possible to detect complex head-mouth related behavioral activity along with regular activity in nursing care center. Our proposed framework and results are explained in *Chapter 6*.

5. ***A Bayesian approach for quantifying data scarcity when modeling human behavior via Inverse Reinforcement Learning (IRL)***

   After having a promising computation model, it is always important to know the amount of data needed to train the model. It is necessary to have a well-informed idea of how much data to collect for both resource conservation and to obtain an accurate model parameter estimate. In the thesis, we proposed a sample size determination approach based on uncertainty quantification (UQ) for a specific Inverse Reinforcement Learning (IRL) model of human behavior. Our approach can be use as a tool to help decision-making for model designers to selecting optimum amount of data when modeling patient behavior. We elaborately explained our approach in *Chapter 7*.

2.4 Summary

In order to solve problems in the field of real nursing care center, it is important to address the challenges of real field nursing care center. There are several challenges associated with data collection, noisy data, daily life complex activity understanding from real field data as well as having a well-informed idea of how much data to collect is imperative for both- conservation of resources as well as obtaining a well-trained model. In this context, if a technology that can detect and understand the regular and complex activity of the nursing care center in advance by combining unevenly distributed sensors and analyze and developed algorithm using advance machine learning techniques, the impact will be extremely large.

In this chapter, we have summarized the related work on the issues studied in this thesis and the main contribution. In the next chapters we will explain elaborately about our proposal and findings.

Chapter 3

A Method for Sensor-Based Activity Recognition in Missing Data Scenario

3.1 Introduction

Human activity recognition (HAR), using sensor-based systems, is one of the most prominent fields of research[1]. One of the primary goals of HAR is to understand the daily behaviors of people through the interpretation of information from different sensors collected from people and their surrounding living environments. Sensor-based human activity recognition has a significant impact on healthcare monitoring, assisted-living, surveillance, entertainment, etc. [42]. One of the main reasons for the exploration of HAR-related research is the increase in elderly population all over the world [72]. The monitoring of elderly people's daily activity becomes crucial to improve their quality of life and to prevent any sudden accidents, such as falls [148, 76]. Through advanced technology, it is possible to detect human body movement. There are lots of works on video-based action or activity recognition [74, 3, 1, 6]; however, video-based methods have privacy concerns that deter many users from video-based systems, and on the contrary, encourage users to adopt sensor-based activity analysis. Apart from the privacy issue, video-based activity recognition has issues with occlusions, segmenting multiple people in the scenes, etc. [1, 3]. Many researchers explored sensor-based activity recognition techniques [160, 56]. In this scheme, wearable sensor data were collected from each individual. Afterward, through near-by access points, data were transmitted to servers. Usually, various features are extracted from these time-series continuous data. Finally, machine learning approaches were explored for recognition of different activities [6].

Researchers are also conscious of real-time applications for hospitals and nursing homes [104, 7]; however, data collection in a controlled-environment or in the wild has challenges [69]. A real field study was introduced in [69] to facilitate nursing

[1]This Section is reproduced from the article Tahera Hossain, M.A.R. Ahad, Sozo Inoue, *"A Method for Sensor-Based Activity Recognition in Missing Data Scenario"*, Sensors, 20(14):3811, 2020-07-08. with permission of the editor

home activity logs and activity recognition. This experiment in nursing care facilities introduce real life activity recognition challenges from field data. Data were collected for four months from a nursing care center. The data included nursing care records at nursing care facilities, and data from activity labels and sensors obtained through smartphones. Evaluation of the data showed a lower recognition accuracy in the real-life situation [69]; however, a controlled experiment using motion capture and smartphone sensors have demonstrated that it is possible to recognize nursing activities by means of motion sensing with higher accuracy [94]. One limitation for the recognition of complex activities in the real field is the use of a single inertial sensor in the phone. Moreover, recognition of regular activity and complex activity from a real-life situation is challenging. Along with these issues, data loss or missing data are a crucial concern while collecting data in the wild. Usually, human walking, jogging, running, sitting, standing, stair-up, stair-down, cooking activities, sports activities, etc. are recognized from standard sensors without any corrupted or missing data [74, 109]. There are more than 150 datasets for the sensor-based activity arena [109]. Usually, these datasets obtain continuous data from the accelerometer or other sensors, and as such, there is no issue of missing data [109].

In a real-life scenario, however, missing data are prevalent in any wireless sensor networks or body-area wireless sensor networks [156]. Data are lost or the presence of incomplete data can be due to various reasons. A few prominent reasons for data missing are: distance within sensors and nearest access point, sensor malfunction, lack of sufficient power in the battery, collision of data packets, transmission problems, dropped connections, troubles in sensor synchronization, corrupted readings, weak Wi-Fi signal strength, limited network coverage, and so on [10, 50, 113, 156, 134, 132, 159]. Mobile devices have limited memory capacity and computational ability that may cause data to be missed [124].

There are three categories or patterns of missing data. Usually, data can be missed at any random order at any period (i.e., missing in random order (MAR)), or data can be missed in a periodic manner (i.e., data not missing in a random order), or the third pattern is the case where data are completely missed or lost, in any random manner [156]. The latter pattern occurs mainly when the battery is off, or when the system has technical problems. It can be predicted as data missing entirely.

The random missing pattern is more realistic because, in real nursing care or a healthcare center, staff use wearable sensors (e.g., accelerometer and gyroscope) for each individual. These sensors' data are then transferred to near-by access points and important data are sent to a server using a wireless sensor network. In some cases, sensing, computing, and communications can be performed on a single chip, which reduce the cost and facilitate the deployment in even larger numbers. Each node can sense its local environment, perform necessary computations and processing, and send the sensory data through multi-hop nodes. Due to these computational complexities, sensors data can be continuously received, and sometimes be missed without following any patterns. In the real field, time-series sensor data are not continuously missed, or not missed by following a specific pattern, or not missed having a rhythmic period.

During the missing pattern identification, distinguishing a complete missing pattern is relatively easier; however, in real-life data collection, the most common missing pattern is missing at random. The random missing pattern is difficult to decipher. Researchers have been engaged on various data imputation techniques, which are not explored in our work [10, 88]. The main contributions of this research are highlighted below: In this work, missing data were not recovered or reconstructed. Instead, human activity recognition performance having missing data was explored based on the most realistic scenario in wireless sensor network. Without exploring any specific data imputation technique, through this approach, any overhead related to data imputation processing time and storage can be reduced. We propose an approach where activity recognition accuracy can be enhanced while having missing data in the dataset. Based on our study, there is no work considering the missing at random pattern without data imputation for sensor-based activity recognition in the manner that we consider in this work. Two challenging benchmark datasets were explored to justify the approach.

For this study, we generated various random missing percentages (2%, 3%, 5%, 7%, 8%, and 10%) on two benchmark datasets. We unearthed different statistical features from the time-series data for classification. To engulf the entire observation of the dataset, an overlapping windowing technique was followed. In the original dataset, there was no missing information. We refer to these data as 'clean data'. Afterward, we created different percentages of randomly-missing datasets for evaluation. In the recognition process, we classified the action instances in three scenarios:

- When both training dataset and testing dataset were clean;
- When training dataset was clean but the testing dataset had missing data; and
- When both training and testing dataset had missing data.

In this paper, we explored two challenging benchmark datasets, the human activity sensing consortium (HASC) dataset [85] and the Single Chest dataset [30] for the experimental proofs.

3.2 Proposed Activity Recognition Method

Time series sensor data are required for sensor-based human activity recognition. In usual cases, the missing data issue is not considered or generated while collecting data from sensors. In this study, we addressed the missing data problem to evaluate the sensor-based activity recognition performance in a clean data environment and different percentages of missing environment. Our approach was not to recover missing sensor data, but instead, we propose a method to increase recognition performance while having different percentages of missing data by using a machine

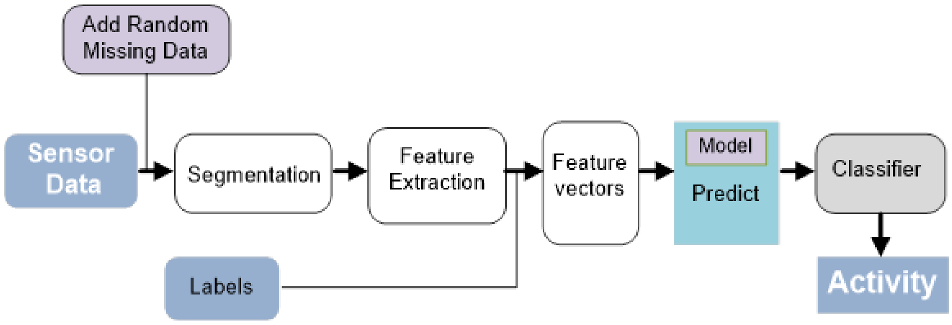


FIGURE 3.1: Illustration of activity recognition work flow considering missing data pattern.

learning approach. The basic recognition process is shown in Figure 3.1. In this block diagram, data are collected from different sensors for different actions. Later, we extracted different statistical features from either uninterrupted or missing data at random pattern. Afterward, for activity classification, we exploited different machine learning classifiers to classify activities. In the initial dataset, there were no missing data; however, in our system, we amended the raw sensor data with different percentages of random missing data.

3.2.1 Induce Missing at Random Pattern

In this section, we explore the recognition performance in multi-sensory networks. A more realistic assumption is that missing patterns will follow the missing at random (MAR) pattern, as many sensors try to send data in real-time. Another terminology for MAR can be stated as missing conditionally at random. This missing pattern depends on some random conditions. The missing data mechanism is an important issue during the missing pattern selection. In a sensor network, missing completely at random is a straightforward situation, as it can only happen when some sensor nodes become faulty. Some replacements for the missing values can be 'NA', 'None', 'NaN', etc. In this work, the pattern we adopted was 'NA', which stands for data that are 'not available' . Let D be the data from sensor, D_o be the observed or available data, and D_m be the data that are missing. Hence, the probability that data are missing $P(r)$ at random can be formalized as: $P(rD_o, D_m) = P(rD_o)$.

To induce the missing at random pattern in the dataset, we generated a random permutation of the sensor data with varying percentages and replaced those time-series tuples by 'NA' value. This was purely a missing at random pattern. This was a normal distribution, and there was no specific pattern or location or clustering of the missing values. Missing patterns can be distributed at any locations in the data. The random patterns were located differently, and no specific routines were followed; therefore, the distribution was purely missing at random. For creating different random missing percentages, we induced 2%, 3%, 5%, 7%, 8%, and 10% random missing data in an actual dataset. We separated each percentage of missing dataset for evaluation.

3.2.2 Methodology

A traditional machine learning model is trained and tested on the data, which has the same input feature space and the same data distribution. If the data distribution is different between train and test data, the performance of the model degrades [73]. In our approach, considering the realistic scenario, we induced missing values in both training and testing modules. In that case, the model learns with the specific percentage of missing input examples. The model initially learns the missing data behavior from the training data and it often aids for generalization. In turn, it improves the robustness of the model. So, while the model learns through having missing examples, and even tests with also missing examples, this method results in a model that will perform better when having missing data in both training and testing modules.

Therefore, we studied the effect of missing environment in the feature values [63]. Consequently, in this work, we created a missing environment in raw sensor data. We considered that missing data may occur during the period of raw sensor data collection in a practical environment. Here, we first created a simulated dataset to evaluate our approach of handling missing data. In the simulation dataset, we evaluated the performances on a clean environment. Afterward, we observed the performances in 10%, 20%, 30%, 40%, 50%, 60%, 70%, and 80% missing data on the simulated dataset.

We tested our proposed approach on two challenging benchmark datasets. These are covered in the following subsections. On the two benchmark datasets, we considered up to 10% missing data. If we chose more than 10% missing data, the recognition performance degrades drastically. After inducing the missing data pattern, we considered that the missing data were existent only in the test data module, while the training data had no missing cases. We built up a model with good quality data, and tested with the random missing data. Afterward, in our approach, in another setup, we added missing data in both training and testing sets. We studied the recognition performance – having missing data in the test module, as well as having missing data in both modules. During feature computation, an overlapping windowing approach was considered for data segmentation. The overlapping windows assist in distinguishing the smooth changeover of the signal. It also plays a significant role when the signal interval is not independent of the signals of the other intervals. It helps to cover all data information at the margin of the windows. Overlapping windows can handle the transitions more precisely. Finally, we exploited classifiers to recognize the classes smartly.

One of the most important parts of classification is to extract useful information from the sensor signal. We not explored any feature extraction method, though there are various feature extraction strategies for activity recognition. In this paper, we specifically concentrated on statistical features on time-series data. We unearthed several features. Table 3.1 depicts the 21 statistical features that are extracted from the rudimentary signals to generate smart features. Our study suggests that this

TABLE 3.1: Extracted features on the HASC and Single Chest datasets.

1. Mean-X	2. Mean-Y	3. Mean-Z
4. Variance-X	5. Variance-Y	6. Variance-Z
7. Skewness-X	8. Skewness-Y	9. Skewness-Z
10. Kurtosis-X	11. Kurtosis-Y	12. Kurtosis-Z
13. Max-X	14. Max-Y	15. Max-Z
16. Min-X	17. Min-Y	18. Min-Z
19. Median-X	20. Median-Y	21. Median-Z

set can profoundly assist in distinguishing distinctive activities. For the selection of important features, we considered [62], where we found that a combination of different features is useful in the extraction of useful information that has missing data. The extracted statistical features were: mean, variance, skewness, kurtosis, maximum value, minimum value, and median absolute deviation (MAD) for the x-axis, y-axis, and z-axis of the accelerometer data. The following paragraphs summarize the features.

Mean and variance: Mean value summarizes the data attributes for the three axes of accelerometer data. Variance is used to identify any sharp details of time series data. Consider that $X = x_1, x_2, x_3, \ldots, x_n$ are the time series of accelerometer data. So, the mean value is:

$$\bar{X} = \frac{\sum x_i}{n} \tag{3.1}$$

where, n is the number of training instances. The variance of any random variable X is the expected value of the squared deviation from the mean of $X : \mu = E[X]$ and variance is:

$$Var(X) = E[(X - \mu)^2]. \tag{3.2}$$

Skewness: By skewness, we can measure the lack of the symmetry of the graph. It computes the skewing rate or irregularity of the probability distribution of a random variable about its mean value. For a balanced dataset, the skewness will be equal to 0. Hence, for a normal distribution, we can have a skewness of 0. It can be computed as, $\gamma = E(\frac{X-\bar{X}}{\sigma^2})^2$, where $\bar{X}$ is the mean and σ^2 is the standard deviation.

$$Skewness = \frac{n}{(n-1)(n-2)} \sum \frac{\left(X_i - \overline{X}\right)^3}{S^3} \tag{3.3}$$

Kurtosis: Kurtosis denotes the peak of a frequency distribution curve. The distribution of flatness is measured by Kurtosis. It can be defined as an average value of the variation of the time series data. It measures whether the data are peaked or flat relative to a normal distribution. It also estimates the tailedness of the probability distribution of a real-valued random variable and the degree of asymmetry of the sensor signal distribution. The peakedness or kurtosis is given by, $\mu_4 = \sum(x - \bar{x})^k.f(x)$, where, all $\bar{x}$ are summation of values and $f(x)$ is the probability distribution function of x.

$$Kurtosis = \left[\frac{n\,(n+1)}{(n-1)\,(n-2)\,(n-3)}\sum\frac{\left(X_i-\overline{X}\right)^4}{S^4}\right]-\frac{3\,(n-1)^2}{(n-2)\,(n-3)} \tag{3.4}$$

Other features: Few other features have been extracted from the accelerometer data. For a particular signal, maxima (Max) and minima (Min) stipulate the maximum and minimum values of that signal. Moreover, the median absolute deviation (MAD) is computed as a feature. The median absolute deviation can be computed by,

$$MAD = \sum\frac{x_i - M}{n}, \tag{3.5}$$

where n = number of samples, $x_i = i^{th}$ sample value and M = median.

There are different classifiers and among them, we explored the support vector machine (SVM) [138] and random forest (RnF) [27] methods. As sensor-based data are not computationally-expensive, similar to the image or video data, we concluded that SVM or RnF classifier can perform without much processing time. For SVM, we exploited the radial basis function (RBF) as the kernel function. SVM is very well-known for its strength in classifying many binary or multiclass scenarios in computer vision, imaging, and other arenas. SVM has been employed in several thousand good research so far.

Apart from the SVM, the random forest classifier is another important classifier that is also widely explored by the researchers. The RnF's strength is its ability in classifying multiclass scenarios with a regression algorithm [41]. The random forest classifier has a set of basic classifiers of decision trees, which are produced in a random manner according to the sampled data in training dataset [41]. These decision trees are trained autonomously. In the test module, each class label is developed based on the multiple classifiers' prediction levels [126]. According to the study in [126], the most inconsistency-tolerant classifier is the random forests; therefore, if the missing rate is higher, it is supposed to perform reasonably better than a basic decision tree, Bayesian network, logistic regression, K-means, and K-nearest neighbors (KNN) algorithm [126]. For Random forest, we implemented R's approach based on the Breiman's random forest algorithm. In the implementation, there were 500 trees. This number was set to a relatively high number so that every input row can be predicted at least a few times.

3.3 Experimental Results

In this research work, we explored two benchmark datasets to evaluate the method: these were the human activity sensing consortium (HASC) dataset [85], and the single chest-mounted accelerometer (SCMA) dataset [30]. The HASC dataset comprises more than 6700 data, taken from the accelerometer. These were captured from 540 subjects. This is one of the few and largest datasets in terms of the number of subjects that were engaged to collect accelerometer-based dataset [85]. Usually, most of the datasets were created using 10 or 20 subjects, , and not well adapted for real world applications. A project called "HASC Challenge" gathered a large-scale corpus of human activity data to resolve the situation. More than 6,700 accelerometer data with 540 subjects were gathered through this project through the cooperation of 20 teams. Hence, this dataset is indeed a difficult one to challenge.

There are six activity classes – stay in a normal position, walking on the floor, jogging, skipping action, walking up the stairs, and walking down the stairs. The data were collected through two different modes:

- Segmented dataset: segmented data per activity per person; and
- Sequential dataset: several activities in a sequence by a person.

The latter mode (i.e., continuous and multiple activities) makes this dataset even more challenging. Apart from the multiple activities in a sequence, the activities were taken at a random order as well. Usually, each activity was taken for 10 s or above for the sequential dataset. The segmented data were measured for 20 s and in *.CSV format. Conventionally, the segmented dataset was considered as a training module, whereas the sequential dataset was taken as a testing module. The sampling frequency was 10 to 100 Hz; the age range of all young subjects was 21 to 32 years. The leave-one-person-out cross-validation pattern was considered in our research.

On the other hand, the single chest-mounted accelerometer [30] dataset (single chest dataset) was designed from 15 subjects who put the wearable accelerometers on their chests. There are seven activity classes in this dataset – standing; walking normally; working at a computer; standing-up, walking and going up or down stairs (a combined and complex activity); walking up or down a staircase; talking while standing; and walking and talking with someone. Though there is are relatively few subjects in this dataset, the activities are complex and challenging when compared to many other existing datasets. The sampling rate of each wearable accelerometer sensor was 52 Hz. Similar to the HASC dataset, the data were stored in a *.CSV format. Before explaining our results, we present the missing data concept in relation to our simulated dataset.

3.3.1 Results for the Simulated Data

In this simulated dataset, we introduced three patterns of signals or data – sine wave, sawtooth wave, and square wave. There were 1200 data entries in the simulated dataset for training and testing. Each data pattern contained 400 entries separately. Initially, these are continuous time-series data, and later, we added some random missing patterns to evaluate our method. This is done to justify or confirm the classification accuracy. At the beginning, we took continuous time-series data for the

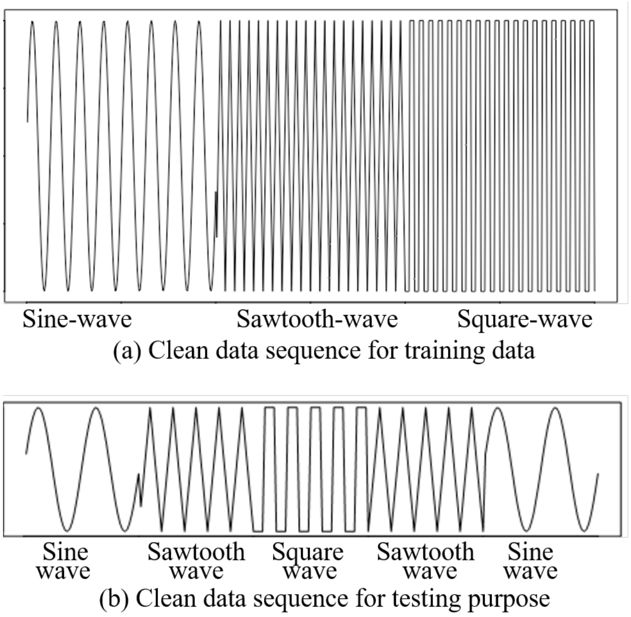


FIGURE 3.2: Time-series data having three different signals: sine wave, sawtooth wave, and square wave; (**a**) clean data sequence for training data; (**b**) clean data sequence for testing purposes.

three signals (Figure 3.2). The training data had three sequences (top image in Figure 3.2). The bottom image of Figure 3.2 has different five signals that were considered as the testing data. Then, we added some missing signals randomly in the sequence – both for training data and testing data. The added missing data were considered from 10% to 80%. However, up to 60% missing data in this short sequence was found to be reasonable. After that, the missing data rate was extremely high and hence, the recognition result fell drastically. The recognition results were computed for clean data as well as for different levels of missing data patterns (from 10% to 80%). These were accomplished when the training data were clean but the testing data had missing data in a random order.

As mentioned above, we computed the recognition rate when the missing data were present in both training and testing data. These results are demonstrated in Figure 3.3. It is evident from these results that by using our proposed approach when both training data and testing data have missing data patterns, the recognition results are always enhancing. According to our study, we have found that when there is no missing data, the recognition rate is 100% in this simulated data [62]. However, after adding missing data randomly in the sequences for testing data only, as well as, for both training and testing data – we can have a better recognition rate when both training and testing data have missing patterns randomly.

3.3.2 Results for the Benchmark Datasets

In this paper, missing datasets were augmented as per the following three conditions on the two benchmark datasets (HASC dataset and the single chest dataset):

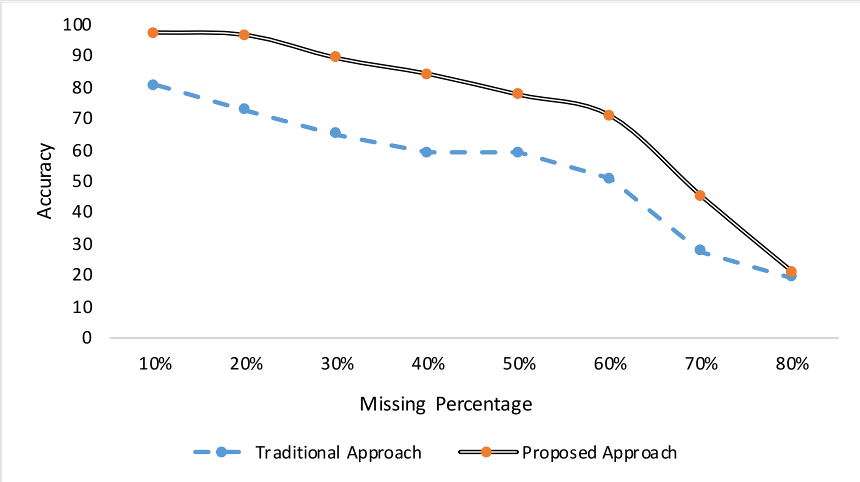


FIGURE 3.3: Recognition results for the simulated data with different data loss rates (from 10% to 80%). Traditional approach represents the results when the training data were clean while the testing data had missing patterns. Proposed approach demonstrates the results when both data had missing patterns.

- Activity classification without any missing data;
- Activity classification with missing data in the testing part; and
- Activity classification with missing data in both training and testing parts.

To demonstrate the variations and data loss patterns, we depicted two different actions and realistic missing patterns in Figure 3.4. In Figure 3.4, data sequences of two activity classes from the two datasets were demonstrated. Among the two datasets, the single chest dataset is an imbalance dataset in comparison to the relatively fairly-balanced HASC dataset. This shows how missing data were distributed among each activity class for different percentages of missing data. For missing data classification, we considered the missing at random pattern, which is a more realistic situation during real field data collection. We generated a random permutation of the sensor data with varied percentages. In this scheme, there was no specific patterns or locations or clustering. It is the sensible scenario while assembling sensor data from multiple sensors; missing data may occur at random patterns, which can be distributed at any location in the data. We embraced two activity classes: 'jogging' and 'going up down stairs' from two benchmark datasets to illustrate the distributions of data missing in various percentages of missing data. Missing rates were from 2% to 10%. Accelerometer data for x-, y-, and z-axes are shown for each case. In the real dataset, it was not considerable to have missing data for more than 8% to 10% in the context of activity recognition.

Figure 3.5 and Figure 3.6 exhibit the activity counts of HASC and single chest dataset. HASC dataset has regular activity classes, whereas, the single chest dataset has two overlapping activities for each class. For example, the going up and down

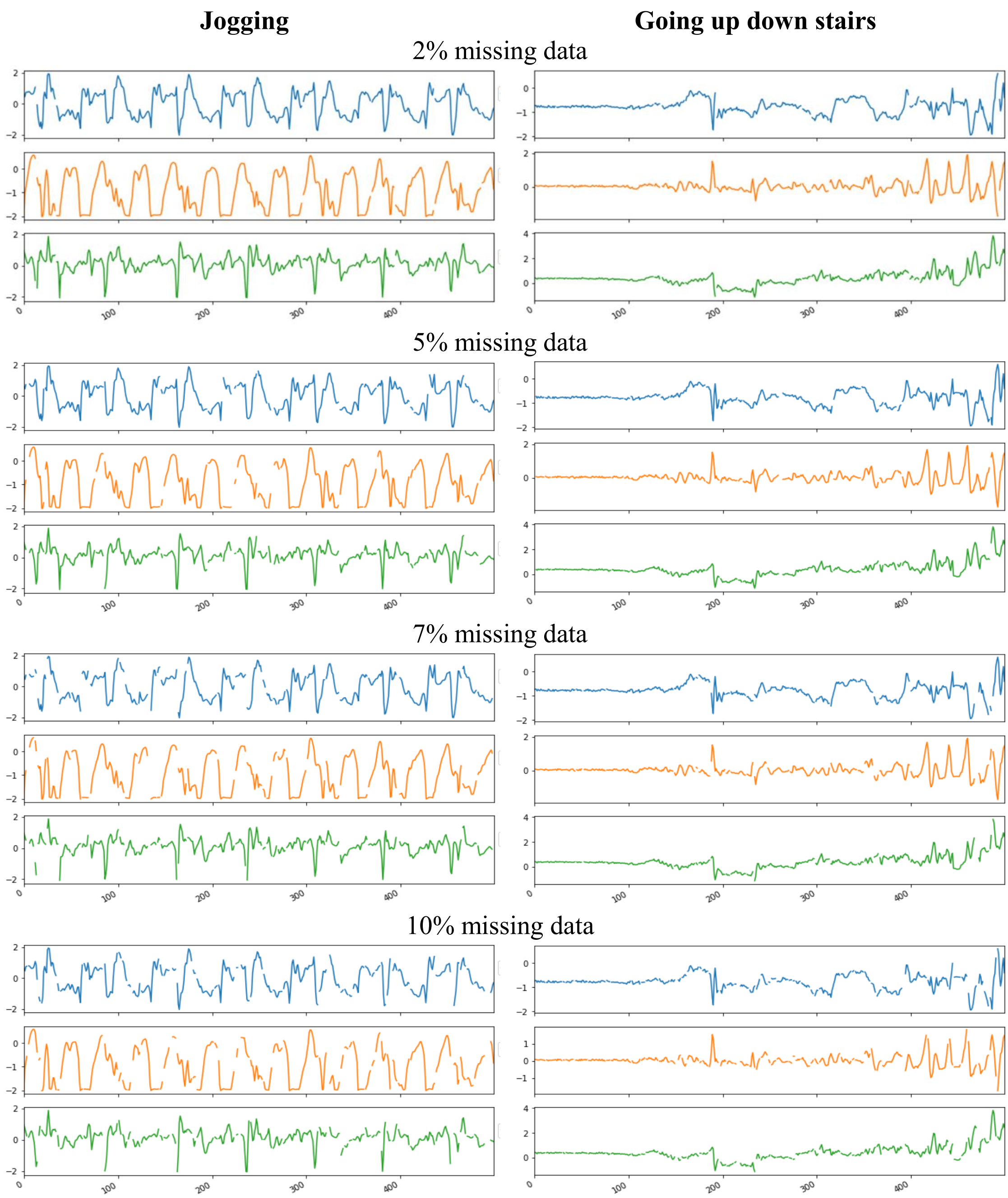


FIGURE 3.4: Data sequence of 'jogging' and 'going up down stairs' from two benchmark datasets that have different percentages of missing data. Missing rates are from 2% to 10%. Accelerometer data for x-, y-, and z-axes are shown for each case.

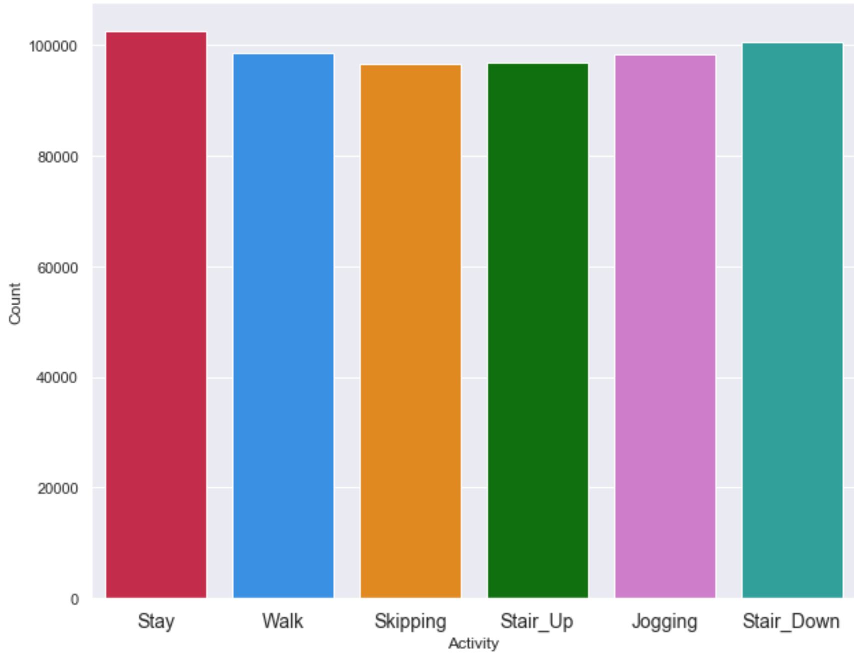


FIGURE 3.5: Activity record count in HASC clean dataset.

stairs together comprise as one class; whereas, the standing up-walking and going-up-down stairs combine as another class. We considered two datasets so that we can evaluate activity recognition performance having missing data in straightforward activity classes, as well as on several overlapping activity classes.

We have orchestrated different missing percentages of data randomly in these datasets. Then, we have observed the performances of recognition results in the following three environments:

- (Case-i) Clean data for both train and test module (Figure 3.7);
- (Case-ii) Clean train data and missing data in test module (Figure 3.8 and Figure 3.9); and
- (Case-iii) Missing data in both test and train modules (Figure 3.8 and Figure 3.9).

In Figure 3.7, we have demonstrated the recognition results for clean datasets (Case-i) in two classifiers. Random Forest classifier has performed better recognition results for both cases. In Figure 3.8 and Figure 3.9, we demonstrated the recognition results for both datasets on different levels of missing data in two environments: clean train data and missing data in test module (Case-ii); and missing data in both test and train modules (Case-iii). We accomplished these to demonstrate whether our simulated analyses have correlation with real datasets.

In Figure 3.8, the results are showing for HASC datasets, for both classifiers, in two environments: having missing data rate from 2% to 10%. More missing data produces lower recognition results, hence these are unacceptable. Hence, we have

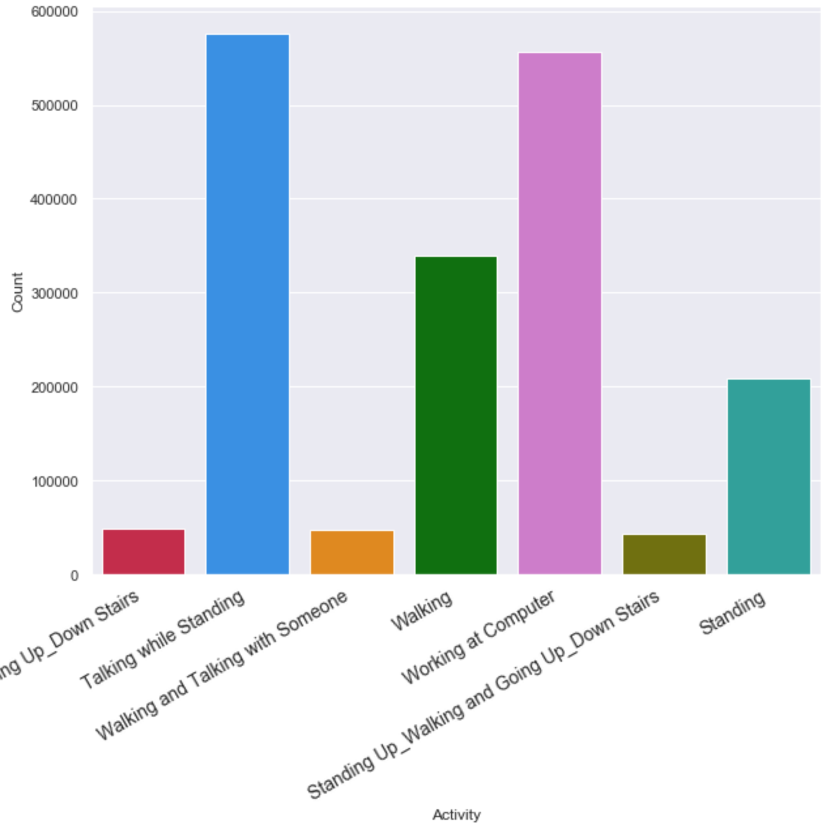


FIGURE 3.6: Activity record count in Single Chest clean dataset.

demonstrated the results up to 10% missing data rate. For the both classifiers, we have found better recognition rates by using our proposed approach when the both train and test datasets have missing information. In the similar fashion, Figure 3.9 demonstrates the approval of the above conclusions that our approach can produce higher recognition results when the both train and test data have missing patterns.

3.3.3 Effect of Different Window Sizes for Different Missing Rates

Window size and window sliding widths have impacts on handling missing data in the dataset. Figure 3.10a is the result of having missing data in the test module, while train data was clean in HASC dataset. We have evaluated the results for the window sizes 1 s, 1.5 s, and 2 s, having missing percentages of 2%, 5%, 7%, and 10%. Results demonstrate that the window size of 2 s has less impact of missing information rather than the window size of 1s. It is because of the fact that a larger window size contains more information of a particular activity class. When any specific activity class is affected more by the presence of missing data, it can handle missing information with a larger window size. On the contrary, a smaller window size may lead to a decrease in the accuracy of not having enough activity class information in that particular window period. Figure 3.10b shows the result of our proposed approach of having missing data in both train and test modules in window size 1 s, 1.5 s, and 2 s for the HASC dataset. It has shown that using the proposed approach

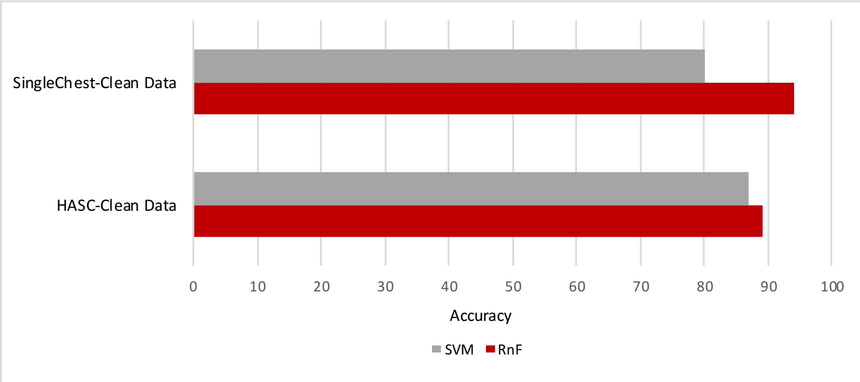


FIGURE 3.7: Recognition results for clean data for both datasets (in %).

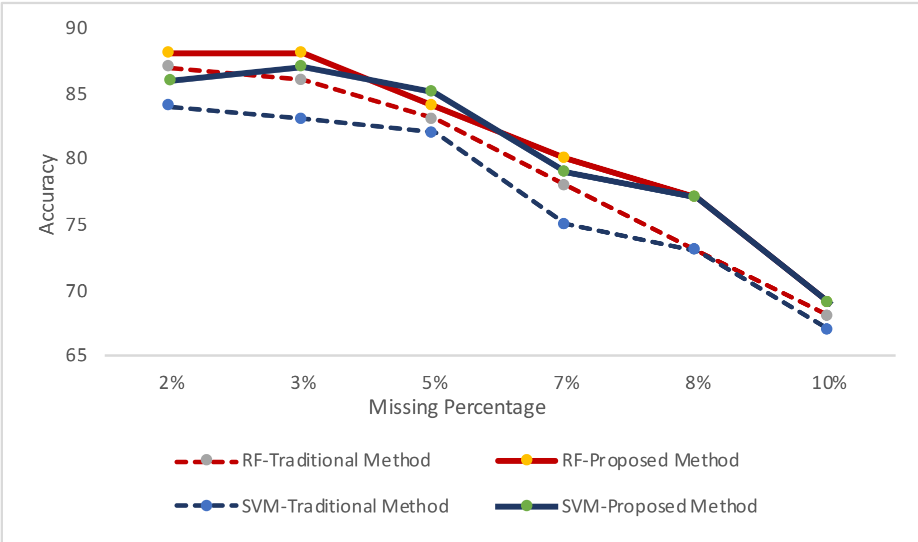


FIGURE 3.8: Missing data analysis for activities of human activity sensing consortium (HASC) dataset.

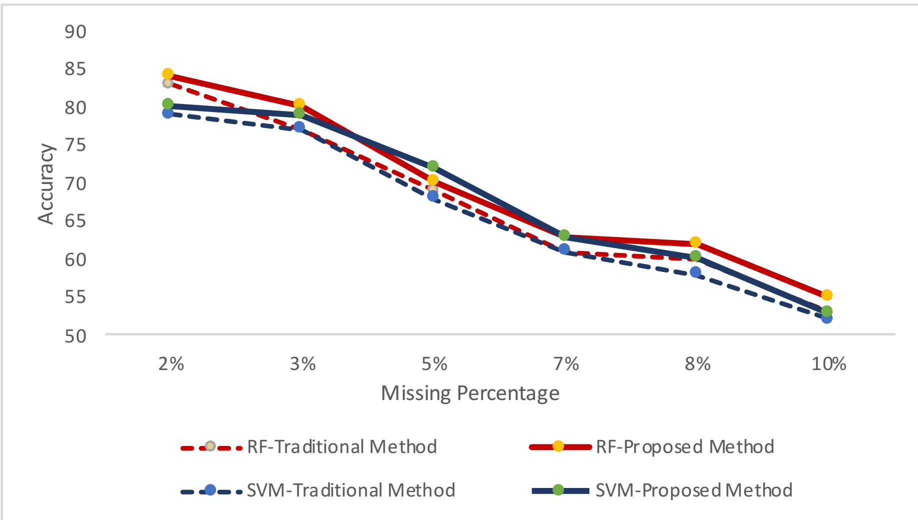


FIGURE 3.9: Missing data analysis for some activities of single chest dataset.

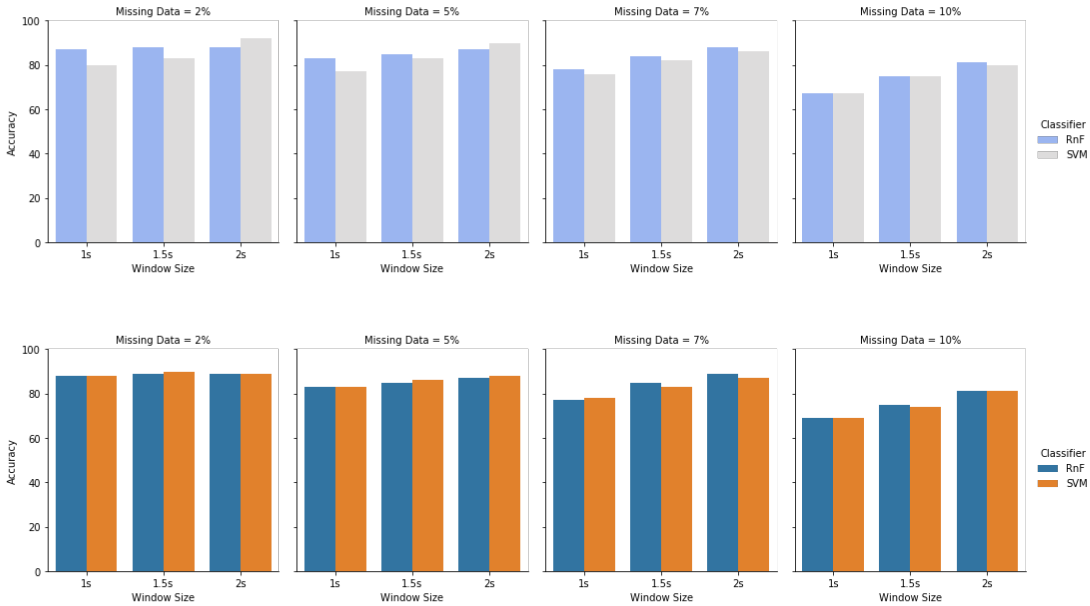


FIGURE 3.10: Evaluation of the HASC dataset for window size 1 s, 1.5 s, and 2 s. Missing percentage 2%, 5%, 7%, and 10% in two scenarios: (**a**) HASC dataset for the traditional approach; (**b**) HASC dataset for the proposed approach.

of handling missing data in both train and test modules can increase recognition performance. Figure 3.11 demonstrates the same evaluation for the Single Chest dataset.

3.4 Analysis and Discussion

3.4.1 Analysis for Synthetic and Experimental Dataset

The simulated data demonstrated some evaluations based on mean and variance. With simulated data, it shows recognition performance as 100%, when it has no missing data in the dataset. Afterward, we evaluated the performance of our approach by adding different percentages of random missing data (from 10% to 80%) in the simulated dataset. Regarding the missing data % variations in simulated data and on benchmark datasets, we would like to mention that the simulated data is a trivial dataset to test the model and to get feedback – based on which, we can work further; therefore, we created a higher % of missing data. The patterns are simple in the simulated dataset, therefore, we need to add more missing data. On the other hand, for real-life situation in wireless sensor network, the missing data are not of a higher rate. In reality, if more than 10% data are missing, for a sequence to decode various activities, there is a huge chance that one or more activities can be completely ignored or missed. Usually, in WSN, missing data at random occur, which is why we explored this strategy. Missing data for a large chunk of data means either the sensor network has experienced a major technical flaw, or there was an error in data transmission. Those issues can be handled by network administrators or

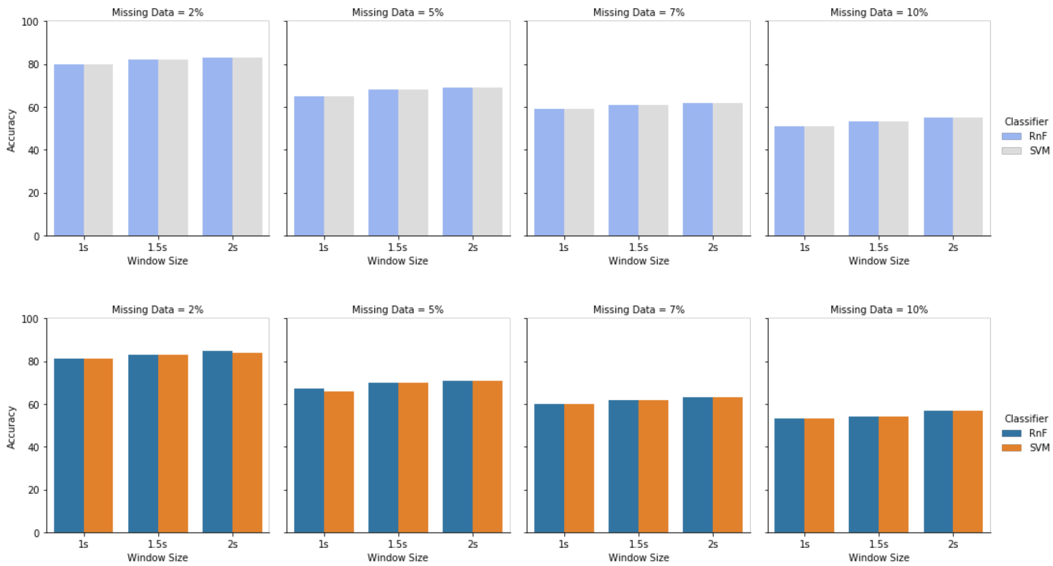


FIGURE 3.11: Evaluation of the single chest dataset for window size 1 s, 1.5 s, and 2 s. Missing percentage 2%, 5%, 7%, and 10% in two scenarios: (**a**) single Chest dataset for the traditional approach; (**b**) single chest dataset for the proposed approach.

network-based data loss assessment mechanism. We ascertain that the current work has successfully dealt with the missing data scenario in a realistic manner.

In our approach, we observed that when both modules have missing data, it always demonstrated better performance compared to the cases of having missing data in only during the testing time. With both HASC and single chest datasets, the random forest classifier provided better recognition results than the support vector machine classifier for both datasets when data are fully clean. With the HASC dataset, it is 89% and 87% with random forest and support vector machine classifiers, respectively. When we have evaluated the same evaluation with the single chest dataset, it achieved 94% and 80% by random forest and support vector machine classifiers, respectively. Evaluation with clean datasets depicted better performances by the random forest classifier. Sensor-based data are not computationally-expensive like image or video data; therefore, we considered that both classifiers can perform without much processing time.

In the evaluation time with HASC dataset using the random forest classifier and a combination of 21 features, we achieved a better result of handling different levels of missing data in the dataset. When a dataset has 3% missing data in the test module, we found that the accuracy is 86%. On the other hand, it improved the performance to 88% when both modules have missing data. In the single chest dataset, the performances of random forest and support vector machine had almost similar results during the handling of different percentages of missing data. With 3% of missing data in the test module, it showed a 77% recognition accuracy, but when both modules have missing data, it improved to 80%. The evaluation results demonstrated consistently better performance when we trained the dataset with a missing

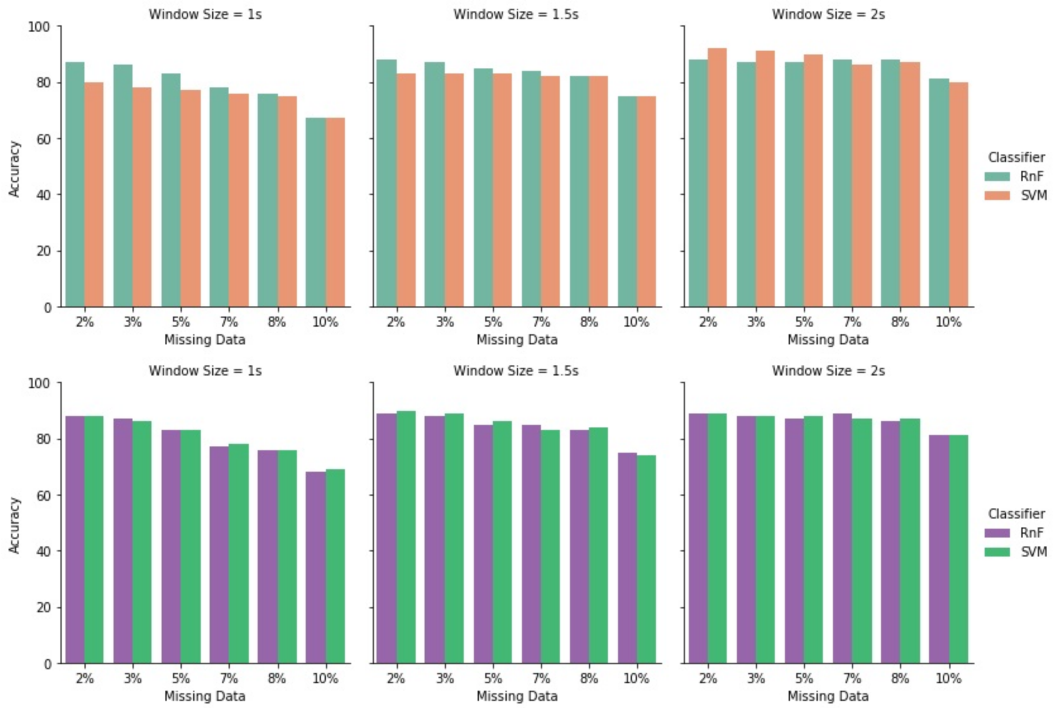


FIGURE 3.12: Evaluation of HASC dataset for window size of 1 s, 1.5 s, and 2 s; and missing percentage of 2%, 3%, 5%, 7%, 8%, and 10% in two scenarios: (a) HASC dataset on traditional approach; (b) HASC dataset on proposed approach.

data environment, and then evaluated with missing data environment at the same time. It has improved the performance compared to the only missing data in the test dataset; however, this enhanced recognition result is not significant. In terms of accuracy and having missing data, we identified that the HASC dataset's performance was superior, because the HASC dataset activity classes do not much vary that much in terms of complexity compared to the single chest dataset.

We tested in different percentages of missing data levels. Table 3.2 represents the comparative result analysis of 5% and 8% missing data in both datasets. We observed that when we added even 5% missing data in the testing module, the recognition rate became 83% by using RnF in HASC dataset, whereas it lowered to 69% in the single chest dataset when only the test module had missing data. By using the SVM classifier too, it was 82% in the HASC dataset, but 68% in the single chest dataset: even the recognition performance in single chest dataset reduced to 60%, when 8% random missing data were present in the test dataset. Performances always improved slightly when using our proposed method when both training and testing modules had missing data in both datasets. On the other hand, for the HASC dataset it was 73% when using RnF and SVM. We evaluated the performances with different window sizes, both datasets performed well with a higher window size in any missing percentages. We evaluated different missing percentages (e.g., 2%, 3%, 5%, 7%, 8%, and 10%) for different window sizes (i.e., 1 s, 1.5 s, and 2 s) – as demonstrated in Figure 3.12. We observed that the window time of 2 s has performed well

TABLE 3.2: Result analysis having 5% and 8% Missing data in both datasets.

	HASC Dataset RnF		HASC Dataset SVM		Single Chest Dataset RnF		Single Chest Dataset SVM	
	5% Missing Data	8% Missing Data	5% Missing Data	8% Missing Data	5% Missing Data	8% Missing Data	5% Missing Data	8% Missing Data
Traditional Method	83%	73%	82%	73%	69%	60%	68%	58%
Proposed Method	84%	77%	85%	77%	70%	62%	72%	60%

in any missing percentage situations. The performances improved for having missing data in both the testing and training modules. We can state that the amount of data is an important issue to estimate the correct activity labels even having a missing data environment. Finally, we noticed that the fairly balanced HASC dataset can handle missing data in a considerable manner, compared to the single chest dataset, which is a more imbalanced dataset. We chose these two different, but genuinely challenging, datasets to manifest the validity and robustness of our approach.

In our approach, the model learns with the specific percentage of missing input examples, and their associated outputs also have different percentages of missing data. In this way, the model initially learns the missing data behavior from the training data. In this study, we explicitly induced different percentages of missing data randomly in the raw sensor data to train the model with missing data. Learning with missing data reinforces the model to regulate missing data during the classification of various activities while having missing data in the test module. However, this method, under realistic circumstances (i.e., real data loss issues in sensor network) where the distribution of the inputs are different between source and target data, approximates the training and testing data distribution totally differently due to random loss happened in training and testing in unknown pattern. Traditional machine learning model is trained and tested on the data, which has same input feature space and same data distribution. If the data distribution is different between train and test data, the performance of the model drops. This approach manifests the plausibility of machine learning model, while it can learn and predict from the identical domain.

3.4.2 Realistic Settings Data Analysis from Real Nursing Care Data

We demonstrated that the approach performs well on the synthetic data and experimental data. Note that sensor-based human activity recognition approaches are still based on limited datasets, by 6 to 10 classes on average, accomplished by few subjects only without any real-life challenging situations, let alone having missing data. Activity recognition based on sensor data is the prominent research area in the ubiquitous computing community. However, there are a very few examples of real-world activity recognition at hospital and nursing homes. Therefore, it is really

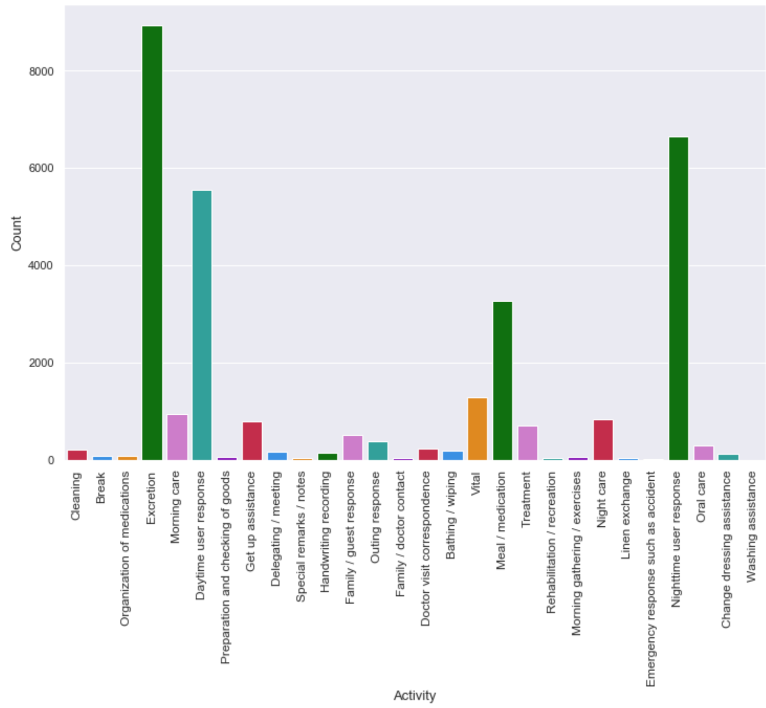


FIGURE 3.13: Nursing care dataset with 21 complex activity classes.

a daunting task to develop more realistic datasets as well as, explore various methods. However, this is the goal of the researchers to achieve. With this reality, we explored our experiments on two really challenging datasets.

Nevertheless, we mention one of our datasets that is explored in a realistic setting on a real-field nursing care dataset [71]. We present this information to show that even in real-life, missing data in sensors are mostly random in nature, which we have explored in this work. In the paper [71], we introduce a system to integrate activity recognition into nursing care record system. The system exploits smartphones to create nurse work and care records, and explores a cloud service to collect sensor data for activity recognition in a real nursing care center for four months. During this experiment, 38,076 activity labels, 2834 h of sensor data, and 46,803 care details were collected. Figure 3.13 shows that this dataset contains 28 complex nursing care activity classes for regular nursing care support, which was challenging to collect, such as vitals (checking), excretion, bathing/wiping, etc. To collect nursing care residents' complex activity data, 27 people, including 23 caregivers and 4 nurses, conducted this experiment in a real nursing care facility center. It required capturing continuous sensor data by using mobile sensors to capture nursing activities; however, there is still a challenge due to the missing samples.

To evaluate the missing pattern from this dataset, we depict Figure 3.14. Here, the abscissa represents the delay or difference of timestamp for receiving samples in millisecond, whereas, the ordinate denotes the record count. In the case of no mission data in the dataset means 200 ms sampling rate for all data. In that case the graph will be one spike in 200 ms sampling rate for all data. If there are no missing data and no variance in the delay then it becomes like one single spike for all data received in 200 ms sampling rate. It is the ideal situation. If we consider the missing value then there will be variance in the 200 ms sampling time but in this time there is small variance in 200 ms sampling time and it is because of jitters. If we have large variance in 200 ms sampling time, then this graph become wider at pick of 200 ms timestamp. When we have one sample missing then it means next sampling time will be 400 ms as 200 ms timestamp is the ideal case for data collection. We have a sample at timestamp 200 ms (this is the absolute time), then 400 ms timestamp as well as 600, 800, and 1000 ms (this carries missing samples information). If one sample is missing at 200 ms timestamp then we can get a record count information in 400 ms or if two samples are missing then we can get missing sample information at the 600 ms timestamp.

Consider the analysis in Figure 3.14, which shows that at 200 ms, samples can be picked, which also have around 1–5 ms. There is no large variance in the 200 ms time. The graph is like a spike graph for this case as all time difference are 200 ms and it is the top sampling timestamp during data collection. We have this kind of small samples because we can see the some of the pick at the very small number/small difference (400, 600, and 800 ms). If we observe the 400 ms sampling data amount then we can observe small number of missing samples and if we observe 600 ms sampling data amount then it means that continuously two succeeding samples are missed. Sometimes this missing pattern are random at any timestamp and sometimes it is continuously missing in two succeeding samples. So, we can see from realistic situation that randomly and continuous missing happens during any consecutive timestamp in the practical application.

To count the missing ratio, we found that missing ration between 400 and 600 ms is 1.15% and missing ratio between 600 and 800 ms is 1.27%, while it is 0.54% during 800 and 1000 ms sampling duration. It is not a large missing ratio in this real application setting. The reason is that our entire networking system and battery lives of sensors were properly checked and enriched with the best-possible manner, so that no missing data can happen. However, in real-life, the causes of missing data may happen and we can face more challenges. After this realistic dataset analysis for missing pattern and missing sample ratio, we can ascertain that our proposed approach can be applicable for realistic settings as well to improve activity recognition. If the missing pattern is random and the missing ratio is approximately 2%, then using our proposed approach, we achieved enriched accuracy up to 88% (for the HASC dataset) and 84% (for the single chest dataset) by using the RnF classifier, and 86% and 80% for these datasets, respectively, by exploiting SVM classifier.

3.5 Summary

Human activity recognition is a very important research area. In this work, we explored the human activity recognition based on wearable sensors and we addressed

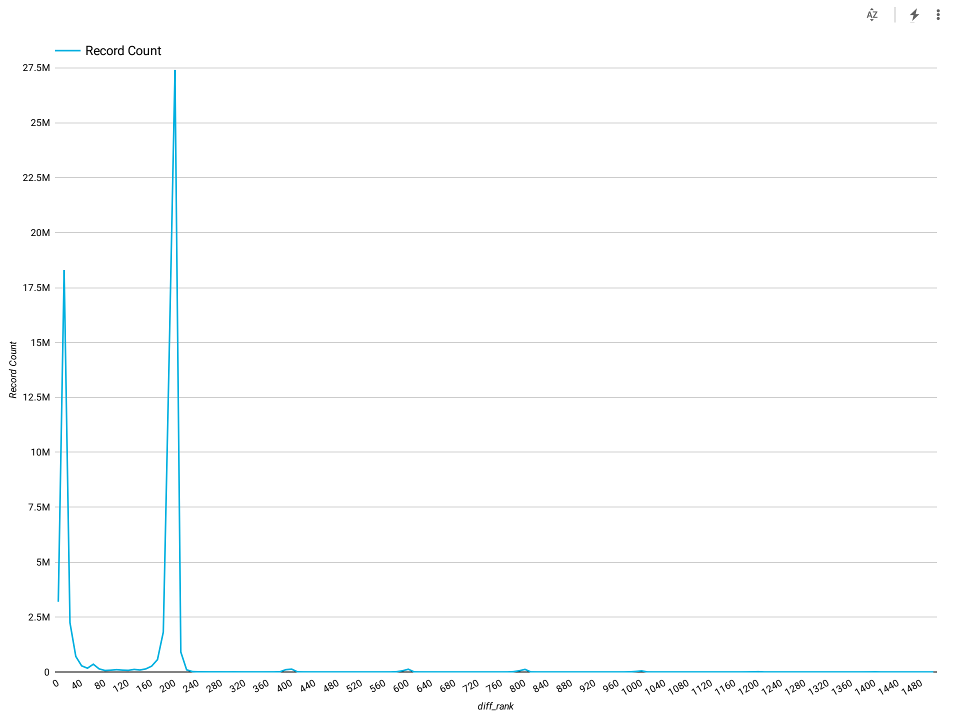


FIGURE 3.14: Demonstration of samples analysis from a real-life nursing care data (x-axis is the timestamp difference for number of samples received in millisecond vs. record count).

a very important challenge, that is, the presence of missing data in the activity recognition without data imputation. We rigorously studied the missing data issue in this domain and proposed an approach to handle the missing data. In our method, we explored different types of twenty-one statistical features that were computed for the three axes of the accelerometer. The features are: minimum value, maximum value, mean, median absolute deviation, variance, skewness, and kurtosis. These features are fed into a smart classifier to recognize different activity classes. Explored classifiers are random forest and support vector Machine. We created various scenarios for experimentation, having various levels of missing data in a random manner, as well as creating different window sizes with varying sliding widths in the time-series data.

Initially, we developed three simulation data patterns with and without any missing data. Based on the successful analysis, we then explored activities from two very challenging benchmark datasets for our experiment. These are the HASC dataset and the single chest dataset. Under various experiments of random missing data levels, we recognized activity classes. From our experimental analysis, we can conclude that the recognition results improve by our proposed method when missing data patterns are available in both training and testing modules without recovering the raw data.

Chapter 4

Activity Recognition Using LoRaWAN (Long Range Wide Area Network) for Nursing Care

4.1 Introduction

Sensors-based human activity recognition system can help to provide information about an early indication of decline in health for healthcare service[2]. Wearable sensors can be extremely useful in providing accurate and reliable information on people's activities and behaviors, thereby ensuring a safe and sound living environment [160]. Machine-to-machine communication has progressed rapidly in healthcare settings . IoT-themed applications in care-giving center need to improve for providing smart care facilities in near future. But healthcare IoT is not without its obstacles.

IoT applications have specific requirements such as long range, low data rate, low energy consumption, and cost-effectiveness. The broadly used short-range radio technologies (e.g., ZigBee, Bluetooth) are not adapted for scenarios that require long range transmission. Therefore, IoT applications requirements have driven the emergence of a new wireless communication technology: low power wide area network (LPWAN). LPWAN is increasingly gaining popularity in industrial and research communities because of its low power, long range, and low-cost communication characteristics. It provides long-range communication up to 10–40 km in rural zones and 1–5 km in urban zones [31]. Therefore, these promising aspects of LPWAN have prompted recent experimental studies on the performance of LPWAN in outdoor and indoor environments [54, 13]. In summary, LPWAN is highly suitable for IoT applications in healthcare monitoring service. LoRa (Long Range) Sensor which is called LoRaWAN is one of the promising technologies among LPWA technologies. These sensors open up exciting new areas for data mining research and data mining applications. In this paper, we explore the use of one of these sensors, in order to identify the human activity.

[2]This Section is reproduced from the article T. Hossain, T. Tazin, M.A.R. Ahad, and S. Inoue, "*Activity Recognition by Using LoRaWAN Sensor*", ACM Int'l Conf. Pervasive and Ubiquitous Computing (Ubicomp'18), 2018. with permission of the editor

In order to address the activity recognition task by using LoRaWAN sensor accelerometer data we used supervised machine learning. We first collected accelerometer data from the users. We then aggregated this raw time series accelerometer data instances, where each example is labeled with the activity that occurred while that data was being collected. We then built predictive models for activity recognition using four classification algorithms. Based on our study, no work has been done on activity recognition using LoRaWAN sensor. Figure 1 describes the basics of activity recognition from sensor data. The objectives of this paper are as follows: To explore LPWA technology for sensing experiment; to use LoRaWAN Sensors for human activity recognition; and to identify supervised classifier to achieve best recognition performance.

IoT-themed applications in care-giving center need to improve for providing smart care facilities in near future. However, IoT applications have specific requirements such as long range, low data rate, low energy consumption, and cost-effectiveness. It provides long-range communication up to 10-40 km in rural zones and 1-5 km in urban zones [13, 54, 31, 44]. LoRa (Long Range) WAN (LoRaWAN) is a very promising LPWA technology. In this paper, we explore the use of LoRaWAN sensors in order to identify the human activity.

The scalability of the LoRaWAN network with multiple end nodes is a challenge to keep all nodes active at the same time. Packet loss and packet receive ratio is a crucial factor in healthcare services. In this study, we explore LoRaWAN sensors to verify the activity recognition. In single device data collection, we observe no packet loss happens. Packet loss happens when multiple devices are connected at the same time. We study the relation between packet loss and accuracy in simulation environment. The objectives of this work are:

- To verify the activity recognition by using LoRaWAN;
- To explore the relation between data loss and accuracy; and
- To estimate the data loss in realistic indoors experiment with many sensors.

4.2 LoRaWAN Sensing and Data Communication Overview

LoRaWAN networks are organized in a star-of-stars topology, in which gateway nodes relay messages between end devices and a central network server. End devices send data to gateways over a single wireless hop, and gateways are connected to the network server through a non-LoRaWAN network (e.g., IP over cellular or Ethernet).

4.2.1 LoRaWAN Connection Components for Sensing

First, we have to integrate three integrated parts (Ardunio Lucky Schield, LoRaWAN and Ardunio Uno) for making a compact system. We used Arduino Uno for 3.3-volt

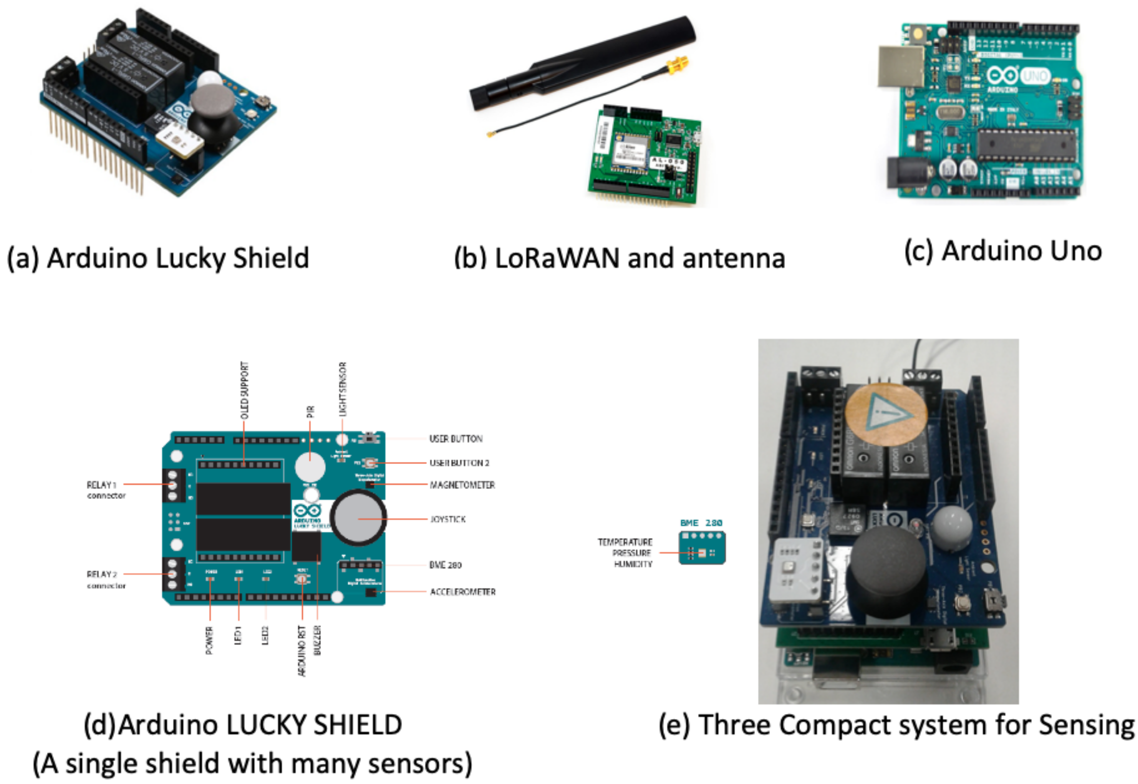


FIGURE 4.1: LoRaWAN connection components for data sensing.

power source which supplies power to LoRaWAN device. By using Arduino Lucky Shield, we can integrate accelerometer and many other sensors with LoRaWAN. Figure 4.1 shows each connection components for complete data sensing and sending through any LoRa based communication system.

4.2.2 LoRaWAN data communication

For LoRaWAN data communication service, we choose SORACOM platform. Figure 4.2 shows the data collection process from LoRaWAN sensors, through LoRaWAN Gateway. After preparation of the sensor setup, we prepare a program for continuous sensing by using LoRa sensors. We write the program among LoRaWAN sensors by using Arduino.

Our sensing system can sense specific sensor data and send those data to SORACOM cloud storage platform (Fig.4.2). We can collect sensing data from SORACOM by using SORACOM Harvest service [106]. The gateway encodes the LoRaWAN frame in base 64 and encapsulates it in a Java Script Object Notation (JSON) structure with other parameters such as the timestamp, specific sensor data, hardware information, gateway information, received signal strength indicator (RSSI) value and signal to noise ratio (SNR) value. Afterwards, we stored sensor data in the Amazon AWS cloud server.

We collected accelerometer sensor data in the lab environment and environmental sensor data (temperature, humidity, light and PIR) in a nursing care center. Based

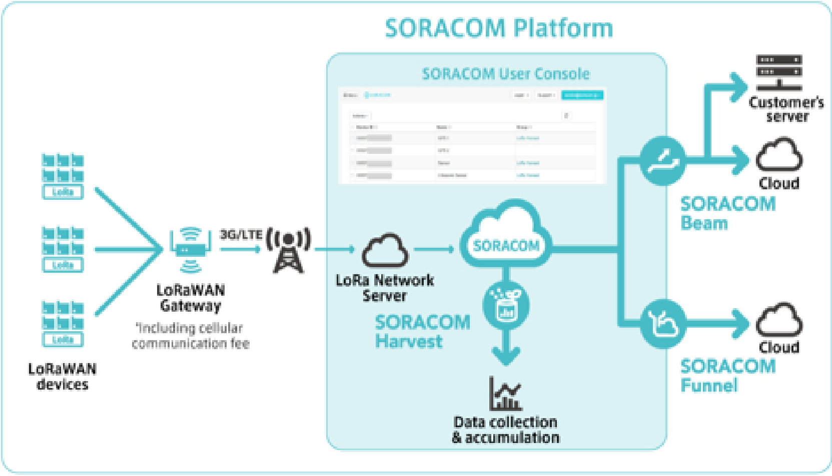


FIGURE 4.2: Flowchart of LoRaWAN cloud server data communication [106].

on our study, we introduced LoRaWAN based sensor for human activity recognition for the first time [64]. No work has been done on activity recognition by using LoRaWAN Sensor. Our experimental result shows promising prospect for LoRaWAN sensor for improving healthcare monitoring service. Based on our findings in the lab experiment, we want to evaluate LoRaWAN sensor data communication performance in a real nursing care field. In the lab study, it has been showed that LoRaWAN sensor data sending ability is much more depending on number of connecting node at the same time.

4.2.3 Indoor Experiment Setup

We collected every second accelerometer sensor data from LoRaWAN sensors in the lab environment[3]. Afterwards, we introduce the setup environment of LoRaWAN sensors in the care-giving facility center in Japan. The facility is a 6-floor building, the first-floor part is the parking lot and the entrance, the second-floor part is the administration office. There are 65 private rooms on the 2nd to 5th floors and residents live. On each floor there is a shared space in the dining room, dining halls, station, waste disposal room, waste laundry room, and there are bathrooms on the 2nd and 4th floors. Figure 4.3 shows a partial view of private rooms and public space in the nursing care center with LoRaWAN setup. We setup 42 LoRaWAN sensors in rooms and public areas. We sense temperature sensor data, light sensor data, PIR sensor data and humidity sensor data in every 3-minute interval from the LoRaWAN sensors. Figure 3 shows the setup image in the facility center. In our setup,

[3]This Section is reproduced from the article T. Hossain, Y. Doi, T. Tazin, M.A.R. Ahad, and S. Inoue,, *"Study of LoRaWAN Technology for Activity Recognition"*, Workshop on Human Activity Sensing Corpus and Applications (HASCA), ACM Int'l Conf. Pervasive and Ubiquitous Computing (Ubicomp'18), 2018. with permission of the editor

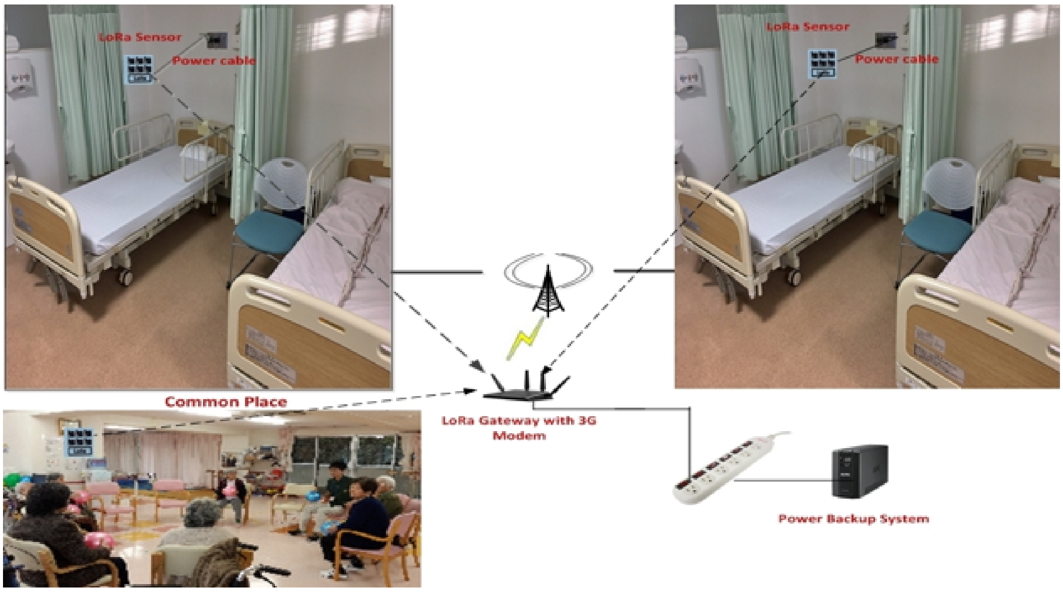


FIGURE 4.3: LoRaWAN setup in the facility center Japan.

our LoRaWAN gateway was in 4th floor and we covered multiple floors LoRaWAN setup in different places through that single gateway.

4.3 Data Analysis and Experimental Results

4.3.1 Proposed Framework for Human Activity Recognition with LoRaWAN

In this sub-section, we introduce the design of the LoRaWAN sensor data communication for continuous activity sensing and recognition by using LoRaWAN accelerometer sensor data. Figure 4.1 shows the core parts of LoRaWAN sensors in our experiment. First, we have to integrate three integrated parts of LoRaWAN devices. We used Arduino Uno, for 3.3-volt power source which supplies power to LoRaWAN device. By using Arduino Lucky Shield, we can integrate accelerometer and many other sensors with LoRaWAN. Figure 4.2 shows the data collection process from LoRaWAN sensors, through LoRaWAN Gateway. After preparation of the sensor setup, we prepare a program for continuous sensing by using LoRa sensors. We write the program among LoRaWAN sensors by using Arduino. We collected every second accelerometer sensor data from LoRaWAN sensors and send those data to SORACOM cloud storage platform (Fig. 4.2). We can collect sensing data from SORACOM by using SORACOM Harvest service. After specific label data collection, we analyze those data by using supervised machine learning approach. In this approach, we collect data from sensors. We segment sensor data for testing and training. Data can be extracted from sensors like an accelerometer. We use windowing approach for feature calculation.

For feature value calculation, we use statistical features - mean, variance and magnitude of data. By the mean value calculation, we summarize the data attributes;

while variance can be used in identifying sharp details of time series data. For classification techniques, K-Nearest-Neighbours (KNN) and Linear Discriminant Analysis (LDA) have been exploited because they are well known supervised machine learning approaches.

For the activity recognition experiment, we collect raw accelerometer data with labels. The measurement time of the segmented data is 8 second. In this experiment, we consider four activity classes: stay, stand, walk and run. For every activity, data instances are collected for analysis. Accelerometer data is stored as a simple comma-separated values format with time stamp and x, y, z-axis acceleration data in the cloud server. We aggregate the time series data into different action labels that summarize the user activity over a time interval. After we collected specific label data for activity recognition, we analyze those data by using supervised machine learning approach. We segment sensor data for testing and training. After, we train data to induce a predictive model for activity recognition. We considered cross-validation approach while computing the feature vectors. Among the collected data, 80% data we used during training and 20% isolated data we keep for testing.

We exploited machine learning approaches to build a model and train the dataset by calculating feature values with a particular time stamp data. For feature calculation, we use x, y, z-axis accelerometer sensor data and magnitude. By using a windowing approach, we calculate the features for 80% overlapping data. For feature value calculation, we use statistical features - mean, variance and magnitude of data. By the mean value calculation, we summarize the data attributes; while variance can be used in identifying sharp details of time series data. For classification techniques, we explore the Linear Discriminant Analysis (LDA), Random Forest (RnF) and K-Nearest Neighbor (KNN). Recognition accuracy is higher by using KNN compare to other algorithm (Fig. 4.4). We achieve recognition accuracy 94.44% by LDA, 84.72% by Random Forest and 98.61% by KNN. Then we impute data loss in our dataset to find out the relation of data loss and accuracy. To evaluate packet loss/missing data environment recognition performance, we simulate missing data environment in our dataset.

We find that the overall recognition performance decreases when packet loss increases. In (Fig. 4.4), we observe that recognition performance is higher by using KNN (98.6%), when there is no packet loss. Recognition performance decreases with 5%, 30%, 50% and 80% packet loss situation gradually. In a simulated environment, we checked the activity recognition performance with 5%, 30%, 50% and 80% data loss environment and have found recognition accuracy of 81.94% LDA, 80.55% RnF and 91.66% by KNN while 5% data are lost (Fig. 4.5). In our simulation environment, we use random packet loss. Missing at random (MAR) is practically realistic in real environment. However, it is a big challenge of LoRaWAN sensors, especially when many sensors are active to transfer data to server at a time, some of them become idle.

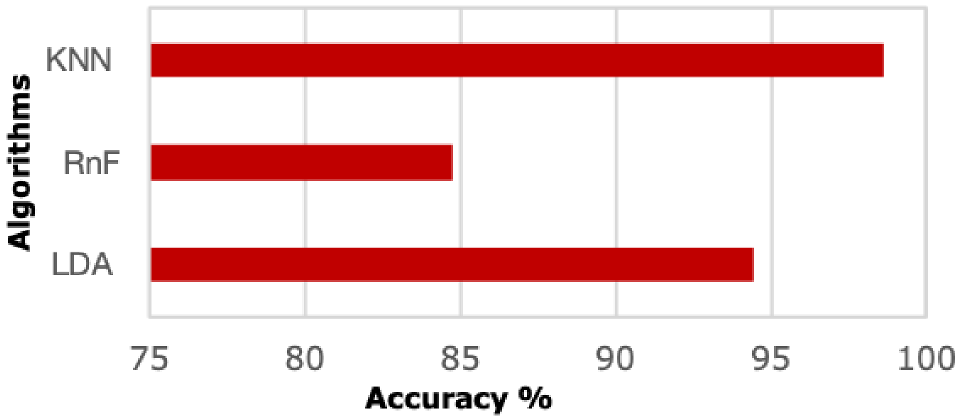


FIGURE 4.4: Activity Recognition Accuracy.

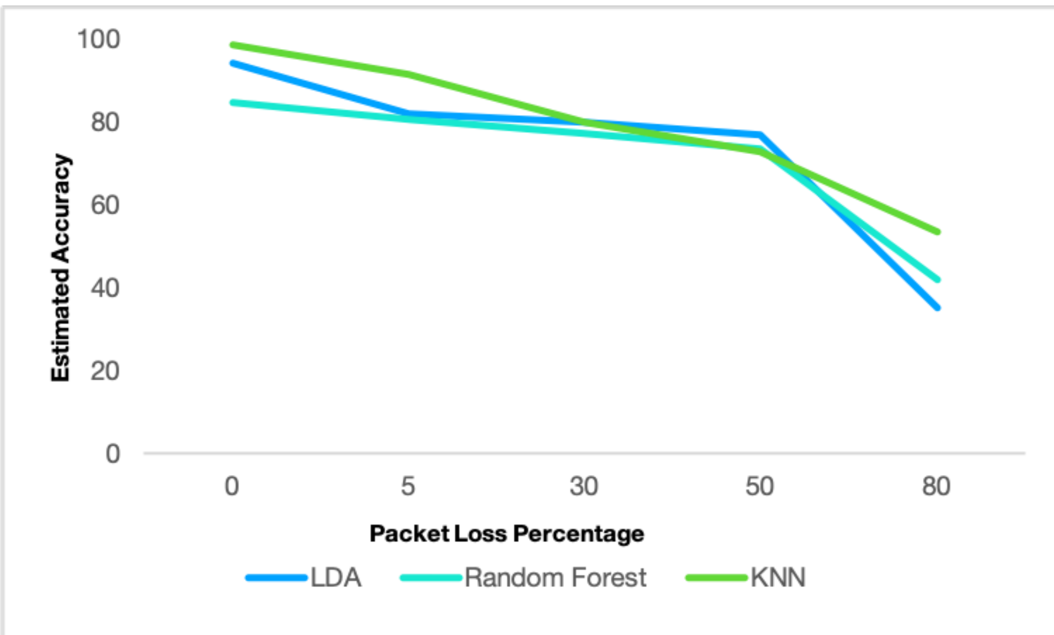
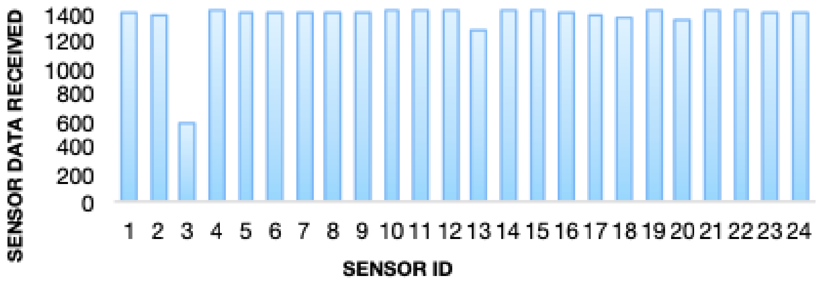
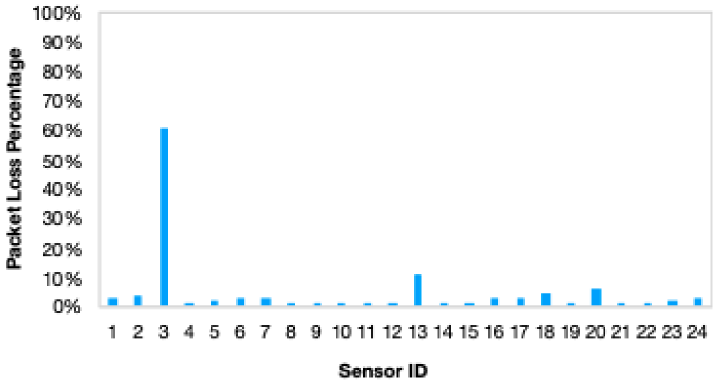


FIGURE 4.5: Activity Recognition performance with different percentage of data loss status.



a) Data received pattern with 24 sensors for a day



b) Data loss pattern with 24 sensors for a day

FIGURE 4.6: Sensor data (a) receive; and (b) loss status.

4.3.2 Experimental Result in Indoor Environment

We introduce the setup environment of LoRaWAN sensors in the care-giving facility center in Japan. We setup 42 LoRaWAN sensors in rooms and public areas. We sense environmental data in every 1-minute from the LoRaWAN sensors. In figure. 4.3 it shows the setup image in the facility center. We integrate 3 parts of LoRaWAN device. We use Arduino Uno (for 3.3-volt power source to power the LoRaWAN device), Arduino Lucky Shield and LoRaWAN. Using a gateway, data communication is done from end devices to the cloud server. After preparation of the sensor setup, we prepare a program for continuous sensing by using LoRaWAN sensors. We can collect sensing data by using SORACOM Harvest service. We analyze data from cloud server.

From this data analyzing part, we find data receive and packet loss ratio from 24 LoRaWAN sensors for a day (Fig.4.6) . Out of the 42 sensors, 18 sensors become idle during the data processing due to unknown reasons. We should get 1440 (1 day = 1x60x24 minute) data from 1 sensor due data collection in every 1-minute, per day. We observe 5% data missing. Figure 4.7, represents 1-month data analysis status from a sensor, where some sensors' data-sending status become low due to congestion of multiple sensors data at the same time. This issue can be explored in future.

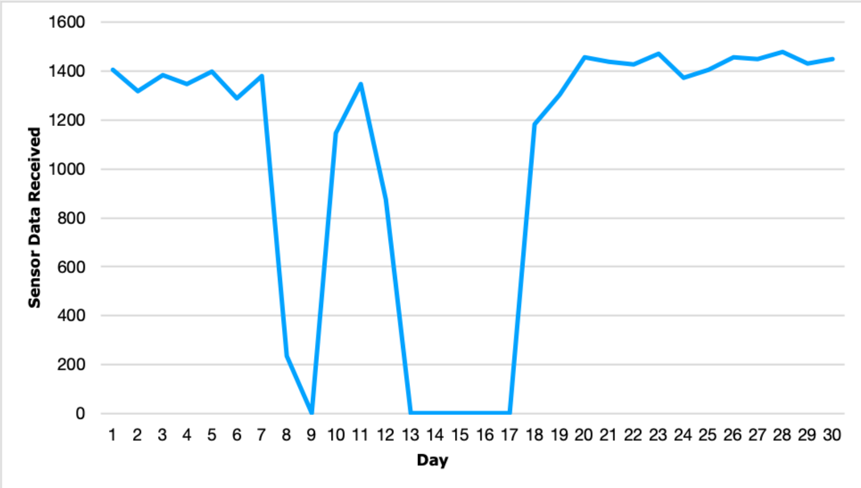


FIGURE 4.7: One-month data analysis from one LoRaWAN sensor.

4.4 Summary

We are exploring this work for scalability analysis of LPWAN technology for real sensing environment where data loss is a vital issue for healthcare monitoring system. We explore the activity recognition by using LoRaWAN sensors. We evaluate the recognition accuracy in data loss environment. From the indoor experiment, we evaluate data loss pattern for 1-day. We observe that during 1-minute interval sensing time, 5% data loss happens by the sensors with single gateway. To improve activity recognition, we need to add smarter features for more classes. In future, by adding second gateway, we will try to increase the number of nodes. It can increase the number of packet receive ratio as well as improve activity recognition performance with less data loss environment.

Chapter 5

Integrating Activity Recognition and Nursing Care Records

5.1 Introduction

In an aging society, the number of nursing care facilities increase, and there is an increased demand for caregiver human resources. As well, it is important to improve the efficiency of nursing care services using information technology.In the field of ubiquitous computing, researches on human activity recognition technology using mobile sensors such as smartphones have been conducted [29]. If this technology is applied to recognition of nursing care activities, nursing care work and records can be created automatically, improving the efficiency of such works. Also, by visualizing the record and looking back, it can also be used as a material for care improvement.

In this work, we introduce a system to integrate activity recognition into nursing care record system. The system uses smartphones to create nurse work and care records and cloud service to collect sensor data for activity recognition. We deployed the system to a nursing care facility for 4 months, where they completely switched from paper records to just using the system in the latter 2 months. The staff, including staffs not familiar with smartphones before, has provided 38,076 activity labels, 2,834 hours of sensor data, and 46,803 care details. These were used as their official care records, without losing the efficiency of their duty. Instead, they improved the amount and speed of recording.

This work also summarizes the analysis of the collected data. We review the dependency of activities to staff users, target residents, and days. We observed that the activities are user and target dependent while not day dependent. We also observed that the care details have a lot of information linked with activity records.

5.2 Requirements of Nursing Care Record Systems

[4]In nursing care facilities, care records must be written for every patient everyday. A care record states the type of care done (*activity*), and its *care details* such as amount of food given or vital information. Nursing care records are used for care coordination, (i.e. what has already been done to the patient and what needs to be done next?), workload planning (i.e. is the workload being equally distributed?), care quality evaluation (are all care activities being done for all patients?), and more unavoidably, for insurance invoices. Documentation accounts for almost one third of nurses workload [60]. Reducing the time spent in writing records can improve nursing home efficiency by allowing more time for patient care. It is said that technology can reduce costs and improve efficiency in health care [35]. For example, digital records can reduce documentation time.

However, introducing IT systems to nursing care facilities is not an easy task. Firstly, unlike hospital records, there is no standard for nursing home care records. Therefore, each facility records in different format and the system needs to adapt to different needs and data requirements. Therefore, the data model needs to be flexible and customizable at each facility. Secondly, the caregivers and nurses of nursing care facilities are not familiar with IT, and a non-negligible number of the staff are senior and have not use smartphones. Therefore, the system needs to be user-friendly for such workers if we want to ask them to use the system. Thirdly, most of nursing care facilities are not connected to the internet, or only the office is connected but they do not cover whole area from room to room, because few elder residents use the internet. Moreover, introducing many wearable or environmental sensors is not realistic from the cost perspective or privacy concerns of residents. Only the use of smartphones or wearable sensors on the caregivers are often allowed. We need easy to install and offline-available system. As such, introducing IT system including activity recognition functionality into nursing care facilities has various challenges.

5.2.1 Label Collection for Activity Recognition

Activity recognition, the automatic identification of what a user is doing based on sensor data, is an active field of research in the ubiquitous computing community [29, 100, 102, 154, 152], but there are few examples of real world activity recognition. Existing studies on real life complex activities, mainly focus on home activities [92, 147, 96]. Examples of using activity recognition at hospitals or nursing homes [115, 15, 74] are limited. Sensors in nursing care facilities have been used for care coordination and planning [128, 142] or are focused on patient sensing [129].

For activity recognition systems aimed at physical activities recognition such as `walking` or `running`, it is possible to create general models that are used as a starting

[4]This Section is reproduced from the article Sozo Inoue, Paula Lago, Tahera Hossain, Tittaya Mairittha, Nattaya Mairittha, *"Integrating Activity Recognition and Nursing Care Records: The System, Deployment, and a Verification Study"*, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT), Vol. 3, No. 86, 2019-09-09 with permission of the editor

model in the application. Such models can be trained with public data. However, in the scenario of complex activities, like `Assist excretion` or `Vital checking`, public data is not available. Moreover, the activity classes are different depending on the users and the business where the application is installed. Therefore, labels must be collected from on-site field.

However, collecting accurate labels (*annotation*) is crucially challenging. A first method is to have the data labeled by third party observers. However, this can be costly and time consuming[34, 74] and it requires the installation of cameras or additional devices [131, 26] if not done in situ. The second method is self-labeling. This can be done by creating diaries at the end of the day [151] or by asking the user to create reports repeatedly throughout the day (Ecological Momentary Assessment) [84, 145]. However, both methods have the risk of missing some of the activity labels [33]. This can often miss the less regular and less frequent activities.

For this problem, we introduce an idea of integrating nursing care record system and activity label collection. Nurse can record the activity labels by themselves as well as inputting care records. Thanks to the integration, activity annotations for each nurse can be obtained from the record system. Furthermore, by integrating activity recognition and nursing care record system, we can also enhance the usage of activity recognition, such as:

- *recognizing the activities* automatically from the sensors and suggesting staff to record them automatically, or,
- by using the sequence of activities day by day, we can also build a prediction model to *predict the near future*, such as tomorrow's care activities for residents, by which staff can prepare for unexpected events of the residents.

5.2.2 Integration Challenge of Nursing Care Record System and Activity Label Collection

While integrating nursing care records and activity label collection is attractive, it introduces some challenges. The first challenge is the accuracy of timestamps. To be accepted by users, an activity recognition system should not be obtrusive [100]. However, in general, care records do not require the timestamps to be precise, such as seconds granularity, but activity labels do. The staff may input the activity labels in a different time than the actual activity, and input them in advance or afterwards. If it happens, challenges of doing activity recognition with inaccurate timestamps as addressed in [148] arise. We may need some techniques to modify or extend the label time durations.

The second challenge is the overlapping between activity labels. To adapt flexible usage of the system, the system should record multiple activities at the same time, because staff may do different type of activities at a time, or the same type of activity to different target residents depending on their business or the type of activities. Moreover, if we modify or extend the label durations mentioned above, the overlap

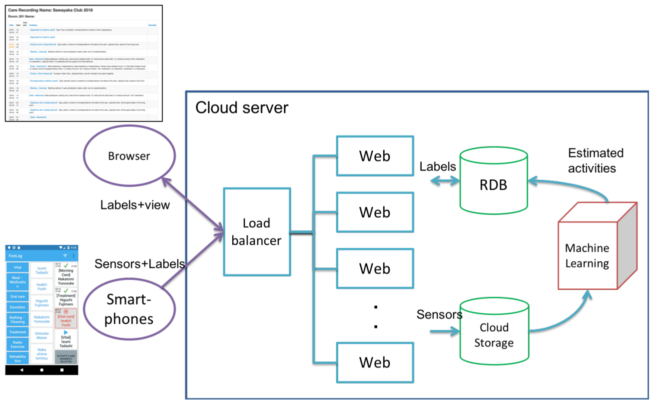


FIGURE 5.1: Care record / activity recognition system configuration.

may happen more frequently. We should be aware of such overlaps, and the activity recognition should handle such overlaps.

The third challenge is the sample imbalances among activity classes[74]. Most of the activities are relatively short compared with the duty time of a day. It makes the samples imbalance among activity classes. Many machine learning algorithms, such as Support Vector Machine, are affected by class imbalances. Moreover, some accuracy metrics such as precision output lower values when imbalance exists. We should use a valid approach for machine learning using imbalanced classes.

Overall, integrating nursing care record system and activity label collection has various challenges, such as flexible data model, user-friendliness for non-IT-familiar users, usage in poor IT infrastructure, self-label collection without much burden, accurate timestamps of labels, overlaps between activities, and class imbalances. We need intensive effort for such application-oriented activity recognition research.

5.2.3 Data Model

Our data model is flexible based on the different requirements of each nursing care facility. Therefore, we designed the data model so that one can retrieve a template form when she/he selects a type of activity, and each activity can be targeted to either of single or multiple residents. The following are the main entities of the data model:

- **Activity class**: These are predefined but customizable, activity types, which can be used for the classes for activity recognition. Moreover, each activity class is associated with a template for record details introduced in the following, thereby a staff can get forms for input from an activity class.
- **Activity**: It resemblances activity instances for each staff. They can be also used as labels for activity recognition after being associated with sensor data. Each activity has information of the subject staff, an activity class, start and finish times, and one or more target residents.

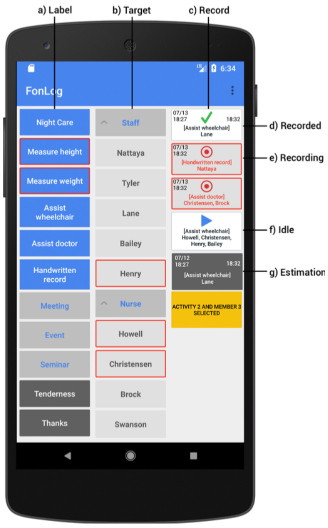


FIGURE 5.2: FonLog: Care record / activity label input application screen.

- **Care detail:** Each activity can have multiple care details. Care details record has various information for target residents such as vital check values and care records. A care detail can be either single choice, multiple choice, integer, number, short text, and long text.

Through these steps, the recorded information can have flexible structure, such that different activities have different input forms, and one activity has multiple targets. This data structure is shared by the smartphone application and cloud server. The developed Android smartphone application called *FonLog*, has a function that when a staff inputs any care record during work, these are relayed to the server on the cloud together with the sensor data in the smartphone (Figure 5.2).

5.3 Deployment in a care facility

Between March and June 2018, we conducted a verification experiment on nursing care records and activity recognition with prior consent of staff and residents at a nursing care facility.

In the experiment, we verify the usability, feasibility, and efficiency of the system. Moreover, we aimed at constructing a dataset for activity recognition described in the following section.

Nursing care records are usually recorded by handwriting at the target facility, but in the experiment, in the first two months in March and April, this system is also used in parallel in the usual way to keep handwritten recording, and in the latter half of May and June, the stability of this system also improved, so we asked the record of this system to stop recording the handwritten record.

The facility is a 6 floor building the first floor part is the parking lot and the entrance, the second floor part is the administration office. There are 65 private rooms on the 2nd to 5th floors and residents live. On each floor there is a shared space in the dining room, dining halls, station, waste disposal room, waste laundry room, and there are bathrooms on the 2nd and 4th floors. We got the agreement of 27 people including 23 caregivers and 4 nurses and conducted experiments.

During the experiment, the staff carried the smartphone during the working hours and had them carried in an arbitrary position such as a pocket. Also, we asked them to record activity labels and care details using nursing care records.

5.3.1 Equipment

We used smartphones Priori 3 LTE from Plus One Marketing. As there were no network facilities such as wireless LAN in the facility, we set up mobile data routers and mobile routers that can be Wi-Fi base stations on each floor. Although it was not possible to cover all the areas, sensor data is stored on the smartphone even if it is not connected to the network, so that no data is lost by this. However, in order to increase the number of simultaneous connections with the smartphone and the bandwidth at the vicinity of the second floor where the smartphone is put at the end of the day, a dedicated wireless router was installed under the mobile router.

5.3.2 Activity Classes

We predefined 28 activity classes after discussions with the staffs and manual documents such as nursing care facilities care records. The activity labels collected are shown in Table 5.1.

TABLE 5.1: Activity classes

Activities of direct care
`1: Vital, 2: Meal / medication, 3: Oral care, 4: Excretion, 5: Bathing / wiping, 6: Treatment, 7: Morning gathering / exercises, 8: Rehabilitation / recreation, 9: Morning care, 10: Daytime user response, 11: Night care, 12: Nighttime user response, 13: Family / guest response, 14: Outing response, 19: Get up assistance, 20: Change dressing assistance, 21: Washing assistance, 27: Emergency response such as accident`
Activities of residence cleaning
`15: Linen exchange, 16: Cleaning,23: Preparation and checking of goods, 24: Organization of medications`
Documentation/Communication activities
`17: Handwriting recording, 18: Delegating / meeting, 22: Doctor visit correspondence, 25: Family / doctor contact`
Other activities
`26: Break, 28: Special remarks / notes`

From the table, we can see that there are categories of direct care activities such as `Vital (checking)`, `Excretion`, and `Bathing/wiping`, maintenance such as `Preparation` `and checking of goods` and `Organization of medications`, communication between others such as `Family/guest response`, `Delegating/meeting`, and `Doctor visit correspondence`.

As mentioned in Section 5.2.3, care detail types are defined for each activity classes. Some care details types are numeric such as `food intake`, `water intake`, some are single choice such as `the places of care support on the body`, and some are text such as `special remarks`. We also found that there are several common care detail types throughout different activity classes such as `excretion assistance` in activities of `Night care` and `Excretion`, and `Special remarks / notes` were defined in most of the activity classes as well as defined as an independent activity class. These are defined by the staff manager according to their usage scenario, and such mixture implies that some activities have similar motions inside, and some may done in parallel by one caregiver, which leads to challenges for activity recognition.

In the next section, we dig in to the detail of the data we obtained. In our extensive dataset, we have many variability and these are realistic dataset. Therefore, to extract important relationship and information from these data, we need to decipher intra-connectivity among various activities based on caregivers/users point of view, residents point of view, date and time wise activity performance as well as time savings by the care record system. In the following section, we are going to give a brief overview of this dataset on this aspect.

5.4 Data Analysis and Results

As a result of the experiment, from the latter 2 month when the system was fully used, we collected 38,076 activity labels, 46,803 record details, and 2,834 hours of sensor data.

In this section, we assess the data obtained during the experience. First, we overview the activity labels and care details, and see the varieties and dependency of activities and care details.

In the following sub-sections, we overview the data from the view points of –

- whether the data are dependent on subject (e.g., staff or target resident), or date, and
- whether the obtained data are meaningful and informative for activity understanding

for both of activities and care details.

5.4.1 Activities

In this subsection, we explore activity classes by depicting several illustrations. In Figure 5.3, there are records of activity classes where the abscissa presents the activity classes and the ordinate shows the records of each class by the caregivers/users.

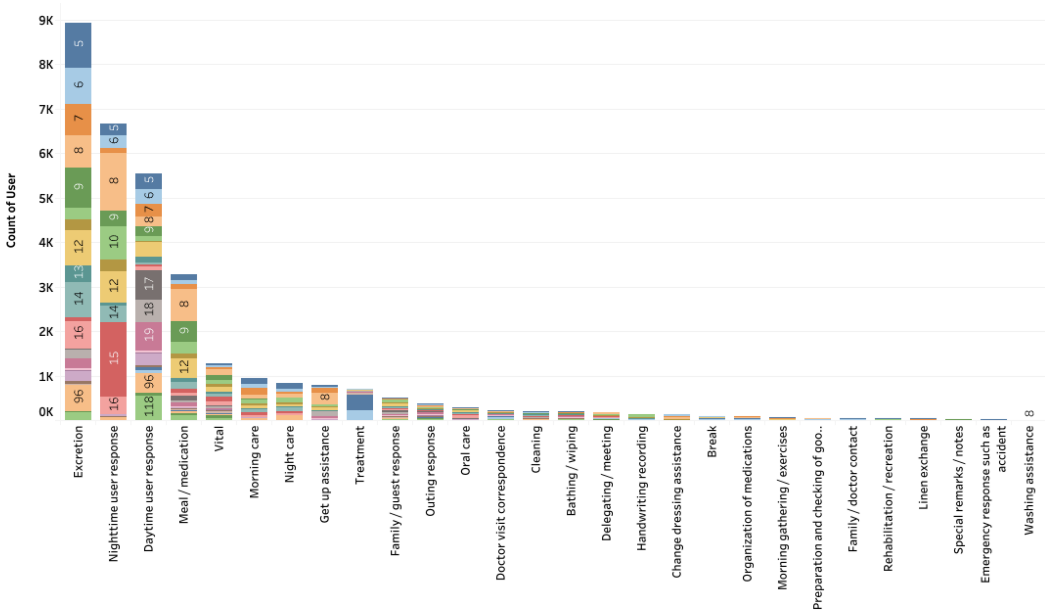


FIGURE 5.3: Number of activities for each activity class recorded by caregivers/users.

We can notice that some classes have higher number of records than other classes. Among them, `Excretion`, `Nighttime user response`, and `Daytime user response` classes have much higher records compare to other classes. One of the main reasons for these discrepancies among classes is that some classes are required on-demand basis, hence, the number of records of those classes are relatively less in number. For example, `Washing assistance`, `Linen exchange`, etc. have very few activity instances. Other activities are mostly done by regular routine. The activity records are also dependent on different caregivers' input. Therefore, we can understand that some caregivers have the higher number of activity-records compared to others.

In Figure 5.4, we analyze the activity records by different caregivers. Here, we decipher to realize whether these records are dependent on caregivers' input. In this figure, the abscissa has the ID of each caregiver. Therefore, we can understand that caregiver8 has the highest number of activity-records, e.g., `Nighttime user response` has the higher records than `Excretion` and so on. It implies that these two activities are more dominant than most of the other activities, as recorded by the caregivers. We also need to evaluate the inter-relationships among various activities, so that the activities and their patterns can be well-understood.

On the other hand, Figure 5.5 demonstrates the activity records on the residents, by the caregivers. In this figure, the horizontal axis shows the ID for the residents, and the vertical axis depicts the number of records. We can notice that among the residents, four of them have significantly large number of records. It has been observed that data records are mostly varied from one resident to another. Some residents has more activity records, whether, others have few number of activities. It implies that these four residents performed more activities than others. Apart from this finding,

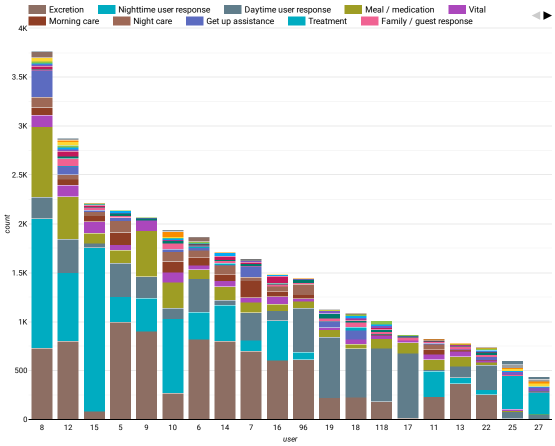


FIGURE 5.4: Activity records by caregivers/users. In the abscissa,the numbers are the ID of caregivers. For example, Caregiver8 has recorded the highest number of activities.

it is also evident that `Excretion` and `Food and medication` are two most widely explored activities for the residents. From the activity records by caregivers (as shown in Figure 5.3), we can get the direct relationship of higher instances of activities, e.g., `Excretion`.

To demonstrate any temporal relationship on various activities, we produce Figure 5.6. We can observe various activities in different dates. From Figure 5.6, we can say that this dataset is not date-dependent. Hence, there is no temporal relationship among activities. In another note, we find differences in terms of day-time activities and night-time activities. These day/night activities are varied in different periods by different patients. From these analyses, we can estimate the activity level of a resident and his/her corresponding health status. If someone has more night-activities, it implies that s/he has more health problems. It requires more staff-calls, therefore, more records on a day by a single resident. These findings can be crucial elements for healthcare study related to activity recognition.

5.4.2 Care Details

For each activity classes, different types of care records are recorded. Under these care records, there are different values for each care record.

In Figure 5.7, we have summarized some specific care records. We find that some care records are associated with multiple activity classes. For example, the `Excretion assistance` care records are associated with `Night care` and `Excretion` classes. Sometimes, the care record values are recorded as assistance types, or as amount of value, or location of body, etc. The care record values are dependent on care record types. In Figure 5.7, we have sorted a few kind of care records and their associated values. For example, `Excretion` activity class has care record type `Excretion assistance`, `Excretion method`, `Voided volume`, and `Defecation`

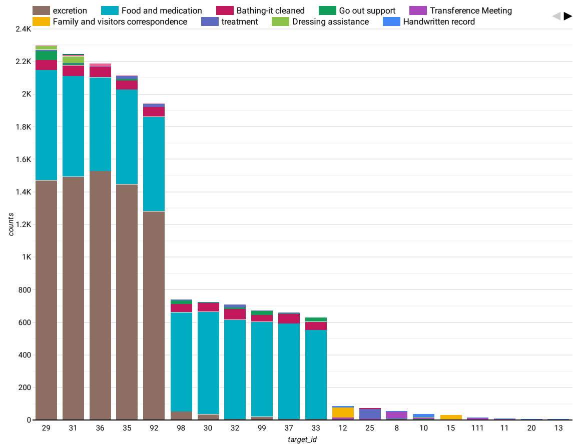


FIGURE 5.5: Activity records on the residents by the caregivers. Here, `Excretion` and `Food and medication` activities are more prevalent.

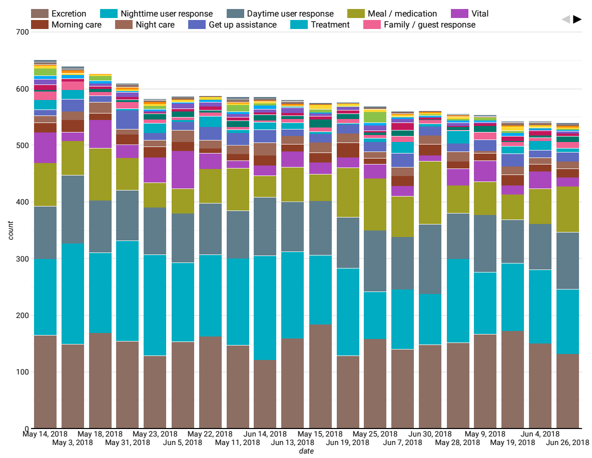


FIGURE 5.6: Demonstration of activity records in different dates.

`amount`. Under `Excretion assistance`, values are recorded by the following different values: all assistance, some assistance, and voice over induction. Similarly, for `Voided volume` care record, values are stored as – medium weight, a large amount, no selection, and small quantity. When the care record is `Food intake`, then values are amount of food intake by a resident.

From these care records and values, we can estimate several interesting information on the activity performances of the residents. In this line, in Figure 5.8, we have counted a specific care record of some residents. From this analysis, we can observe that residents with IDs – 36,35,92,31 and 29 require more `Excretion assistance` support. This demonstrates a clear connectivity between the Figure 5.5 and 5.8. In both images, resident with IDs: 29, 31, 35, 36, and 92 have more records and activities. In the similar fashion, we can also analyze a resident's meal assistance, bathing

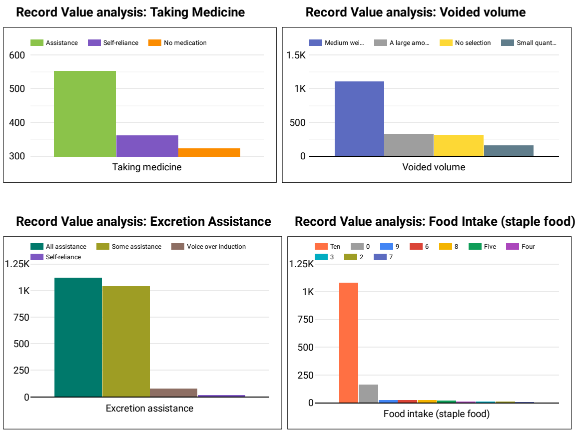


FIGURE 5.7: Record details with different values

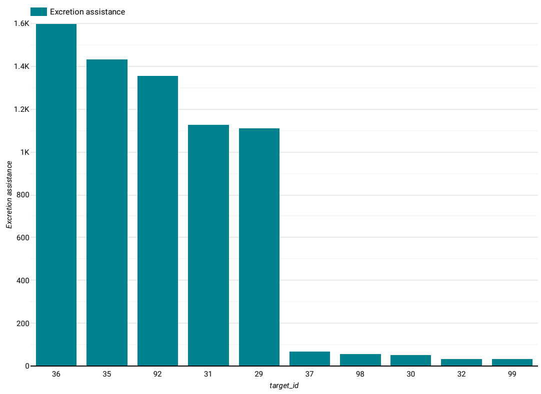


FIGURE 5.8: Measurement of `Excretion assistance` care records for residents. The abscissa represents some residents with their IDs.

assistance, taking medicine assistance and other assistance. These cues are detrimental to predict resident's other activities associated in regular basis. From this care record data, we can find out essential support requirements and their other activity performance in any nursing care facility.

These findings can be essential elements for activity recognition, having many intra-class relationships and sub-divisions among activities. We need to decipher more intra-connectivity among various activities.

5.5 Summary

From this data, we revealed the varieties, nature and dependency of activities and care details which can be a measure of any healthcare outcome. We analyze intra-class relationship of real life nursing care activity data which has important information to revealed the nature of real life data. We investigated the dependency of activities to staff users, target residents, and days. Using the obtained data, we revealed the nature of the data, including dependency of activities to several factors, and the nature of timestamps of self-labeling. We investigated "staff activity" for support service requirements as well as "elderly people activity" from this data. This work also help to examined elderly people health status from the data and types of activities recorded by staff.

Chapter 6

Exploring Human Activities Using eSense Earables

6.1 Introduction

In the healthcare monitoring center, it is important to detect the behavioral activities of elderly people related to head and mouth[5]. It is quite difficult to detect some kind of mouth-related activities like speaking, eating, laughing, etc. through regular wearable devices. These activities have an impact to measure the well-being of a person and socialization ability. In assisted living and healthcare centers, caregivers also need to record and manage activities of elderly people like the amount of food intake, swallowing ability, etc. These are measurement criteria for good health for elderly people. Taking this into consideration, in this paper we mainly focus on detecting head- and mouth-related activities with some other regular physical activities using eSense device. It can be used for monitoring and analyzing personal scale activities [86, 66]. eSense is a new device from Nokia Bell Lab, UK [86]. We acknowledge them for providing the devices to explore in various applications and methods. In this paper, hence, we explored as in the paper. eSense is a very light-weight, wireless device. It performs all required functionalities of earbud like listening to music, receiving phone calls, etc. Besides it contains accelerometer and gyroscope sensors which enables it to track head and mouth related activities drinking, eating, shaking, speaking, etc. These types of activities cannot be detected using smartphone sensors. eSense device can detect minute head and neck movements that have potential applications in clinical medicine related to head and neck injury. Besides, with the help of conversational activity monitoring capabilities of this device, social interactions can be detected. As a result, it will help to treat different types of mental health conditions. In a nutshell, eSense has potential usability in the areas of computational social science, healthcare, and well-being.

eSense is a multi-modal device. It has a 6-axis inertial measurement unit (IMU) with Bluetooth unit. Using eSense devices, it is possible to collect real-time data

[5]This Section is reproduced from the article Tahera Hossain, Md Shafiqul Islam, Md Atiqur Rahman Ahad, Sozo Inoue, *"Human Activity Recognition using Earable Device"*, Proceedings of the ACM International Conference. Pervasive and Ubiquitous Computing (Ubicomp'19), September 2019.

– audio, motion, and proximity. From eSense we can collect three-axis accelerometer and gyroscope data to detect various human activities. It is also possible to collect audio data from eSense which can be impactful to distinguish some similar pattern of mouth movement. There are various types of popular machine learning methods for sensor data exploration. Feature selection has a significant impact on several machine learning approaches for measuring accuracy [155]. For deep learning methods like Convolutional Neural Network (CNN), manual feature extraction is not required [85]. CNN can learn to identify complex patterns. In this paper, our goal is to detect head- and mouth-related human activities with other regular activity through eSense. We explored both traditional machine learning approaches and CNN to classify selected activities.

For detecting human activities different types of sensor data (accelerometer, gyroscope, etc.) are collected from the smartphone, smartwatch, Inertial Measurement Unit (IMU), etc. These devices are placed in different body parts i.e., wrist, leg, waist for detecting activities like walking, sitting, running, etc. From these devices head and mouth-related activities i.e., eating, nodding, drinking, shaking, etc. cannot be detected properly. eSense helps to detect such activities. It is an earbud that contains an accelerometer, gyroscope sensors. From these sensor data such head and mouth-related activities can be detected. So initially we have taken the challenge to detect some of the useful gesture types like speaking, head shaking, nodding, eating, walking, etc. Besides we have developed our own Android application "eSenseLog" to connect the smartphone to eSense device to fetch accelerometer, gyroscope, and audio data. We also presented a relative comparison of various models' performance for detecting head and mouth-related activities from the only accelerometer and both accelerometer, gyroscope data. There are other activities like measuring VO2 max, monitoring heart rate, chewing, swallowing, etc. These activities are important in nursing care. In the future, we will detect these activities. Our current work will play a vital role in detecting such activities.

The objectives of this paper are as follows: collecting accelerometer and gyroscope data from eSense device through Bluetooth and mobile application; detecting head- and mouth-related human activities alongside some other regular activities; using traditional machine learning and deep learning classifiers to detect activities and compare performances[6]. We compare the result performance among machine learning technique and deep learning technique as well as we evaluate the result while having only accelerometer sensor data and having accelerometer and gyroscope sensor data together. Having both sensors' data demonstrate significantly higher recognition results (e.g., for Linear Discriminant Analysis (LDA): 4%, for Support Vector Machine (SVM): 6%, for Random Forest (RnF): 3.5%, and for Convolutional Neural Network (CNN): 3%).

[6]This Section is reproduced from the article Md Shafiqul Islam, Tahera Hossain, Md Atiqur Rahman Ahad and Sozo Inoue, *"Exploring Human Activities Using eSense Earable Device"*, Activity and Behavior Computing, Smart Innovation, Systems and Technologies, Vol. 204, Springer, 2021.

6.2 Proposed Framework and System

The eSense device can be used to track head- and mouth-related activities like nodding, eating, swallowing, speaking, etc. It can be used for detecting other activities like walking, staying, etc. There are accelerometer and gyroscope sensors in eSense device. So these sensors data can be used to recognize human activities. Besides the audio data from eSense device can be used to track other activities like gossiping, laughing, etc. After connecting the eSense device to a mobile device, sensor data can be collected. The eSense device is regarded as a peripheral device and the mobile is regarded as a host device after establishing the connection between the mobile and the eSense device. Through an Android application, the sensor data and audio data are collected. Our developed Android application continuously maintains a connection with eSense device and helps to collect data from the device. In this section, we present our developed Android application, system architecture, and framework.

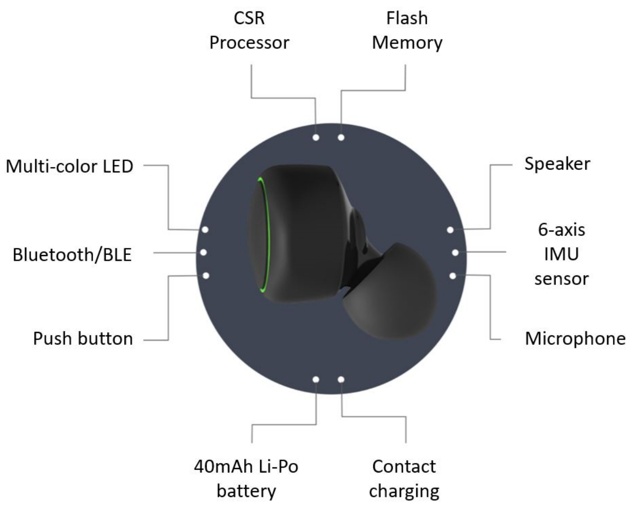


FIGURE 6.1: Overview of eSense device [86]

In Fig. 6.1, an overview of eSense device is presented. Advanced Audio Distribution Profile (A2DP) is used in the firmware of eSense device. A2DP is used for mono channel recording and high-definition audio streaming [86]. A thin middleware is used in eSense to support application development in Android and iOS platforms. The Node.js middleware allows developers to connect eSense devices with desktop applications, configure and ingest sensory data in real-time.

The size of both eSense device and standard wireless earbud is equal and it contains a battery, electronics, etc. The eSense device contains a custom-designed PCB (Printed Circuit Board) whose size is 15 x 15 x 3 mm [86]. The eSense device has System-on-Chip (SoC), an Inertial Measurement Unit (IMU). The IMU contains a three-axis gyroscope and accelerometer. It also contains a two-state button; a circular LED, a digital motion processor, and battery-charging circuitry. The eSense device is powered by a 40-mAh LiPo battery. The weight is each earbud is 20g and the size is 18 x 20 x 20 mm. There is no flash memory in eSense device.

The right earbud does not have a Bluetooth Low Energy (BLE) interface. On the other hand, the left earbud has a BLE interface that can be used to configure various aspects of the IMU sensor and collect accelerometer and gyroscope data. By default, It transmits periodic BLE advertisement packets. The interval is between 625ms and 750ms. The advertisement packets contain the Complete List of 16-bit Service Class UUIDs (Universally Unique Identifier). All the standard UUIDs are used in eSense [8] device.

eSense is a Bluetooth Low Energy (BLE) device. It is based on a specification called "General ATTribute profile" (GATT). The specifications to send and receive attributes between a server and a client are defined by GATT. Attributes are short pieces of data. In this case, eSense is the client and the connected host device (mobile) is the server. The Attribute Protocol (ATT) is used as the base to build GATT on top of ATT. GATT consists of "characteristics", "services", "profiles", "descriptors". A profile consists of one or more services.

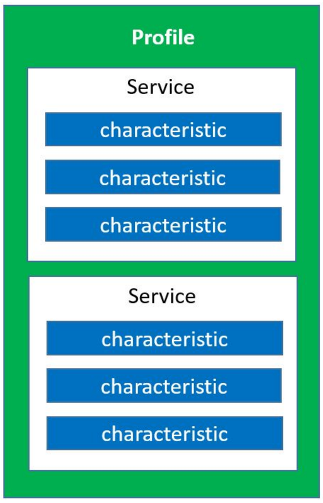


FIGURE 6.2: GATT diagram.

In Fig. 6.2, the GATT diagram is explained. Each service consists of one or more characteristics that encapsulate data. In eSense, there are such services that provide accelerometer, gyroscope data. A characteristic has a single value and 0-n descriptors. The descriptors describe the characteristic's value. For example, a descriptor might specify a unit of measure which is specific to the characteristic's value, a human- readable description, etc. It is similar to a class. When the connection gets established between a Bluetooth Low Energy (BLE) device to a host device like a smartphone, then the Generic Attribute Profile (GATT) is used. This general specification is used for sending and receiving 'attributes' over a BLE link. Attributes are short pieces of information.

In Fig. 6.3, an overall system architecture is explained. It is for the BLE device and a host device. The eSense has accelerometer and gyroscope sensors. Each sensor (accelerometer or gyroscope) sends some characteristic data. This data contains a single value and 0-n descriptors. The descriptor describes the characteristic's

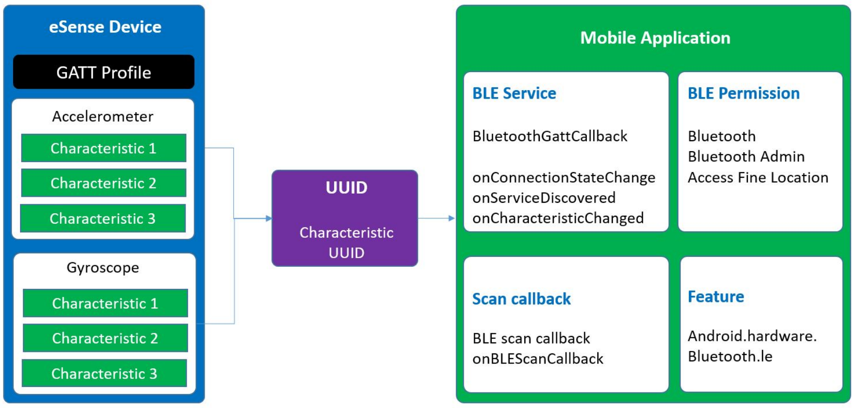


FIGURE 6.3: A general system architecture.

value. Each GATT profile contains one or more services. The service has some characteristics. A Universally Unique Identifier (UUID) is used to identify each attribute uniquely. The UUID has a string ID which size is 128-bit.

The host device (for example, a mobile phone) scans for nearby devices to connect. There are some callback methods, which are used in the application. These callback methods are used in different scenarios. After establishing the connection between mobile and eSense device, *onConnectionStateChanged*() method is called. The *onCharacteristicChanged*() callback method is responsible when data is sent from the eSense device to a mobile device. The whole process runs in service to make it runnable in the background. Inside the application, different types of callback methods and normal methods are used for establishing a connection, sending and receiving data.

We have developed an Android application named 'eSenseLog' for collecting audio and sensor data from the eSense device (6.4 demonstrates a screenshot of it). There are several features of this application. A user can connect eSense device with mobile phone by clicking the 'Connect' button as shown at the top-left corner of Fig. 6.4. If the device gets connected, its name and connected status will be shown.

In Fig. 6.4, the overview of our developed eSenseLog application is presented. After connecting the device, if someone wants to collect any activity data, he needs to click an activity button from the list of activity buttons. There is a toggle 'start' and 'stop' button icon. After selecting the activity, if a user clicks 'start' toggle button, the data collection will be started. The elapsed time will be shown under the toggle button. If the user clicks the toggle button again, data collection will be stopped and saved in the mobile storage. The sensor data will be saved as '*.csv' format and the audio data will be saved as '*.3gpp' format. In the app, there is an activity history list from where the user can easily track the collected activities and their durations.

In Fig. 6.5 our proposed framework is shown. The audio and sensor data from the accelerometer and gyroscope will be collected simultaneously. These data are saved in the mobile phone's internal storage. However, if there is not enough space and if any external storage (e.g., a microSD card) is available then the data will be

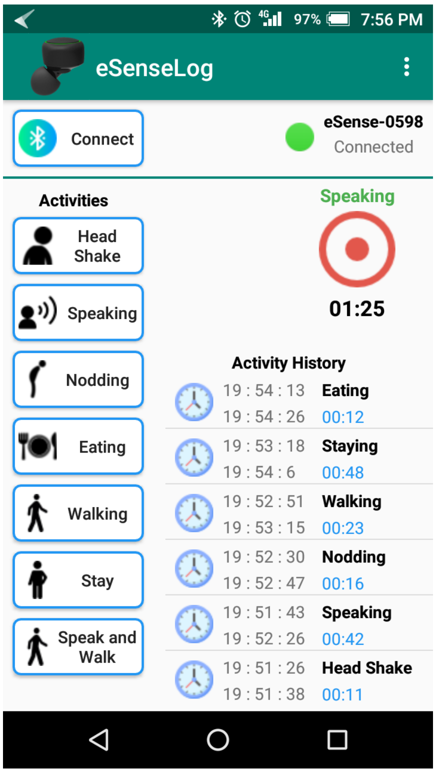


FIGURE 6.4: Overview of eSenseLog Android application.

saved in external storage. When the data collection is finished, a user can fetch the data from the mobile device and then apply any machine learning algorithm on the collected data for recognizing activities.

6.3 Experimental Result and Analysis

We performed experiments for seven activities. Four activities are related to head and mouth (namely, eating, speaking, headshaking, and head nodding). Three regular activities (i.e., walk, stay, and speaking while walking). These activities are

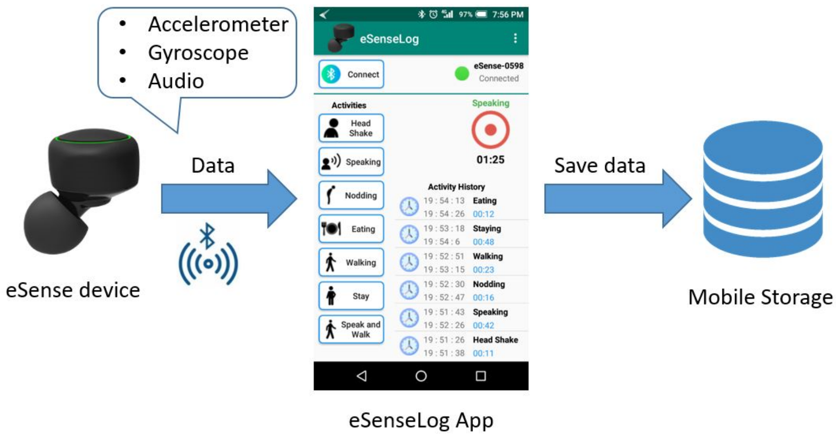


FIGURE 6.5: The proposed framework for this work.

detected from eSense device's accelerometer and gyroscope data. Each activity was performed for 3 minutes. One-person data collection time was 21 minutes for all activities in total. We have collected these activities data from 6 persons (having age range from 25 to 35 years, 4 males and 2 females). There are around 105881 records for these seven activities. The sampling rate was 50Hz while the data was collected from 3-axial accelerometer and gyroscope sensors of the eSense device.

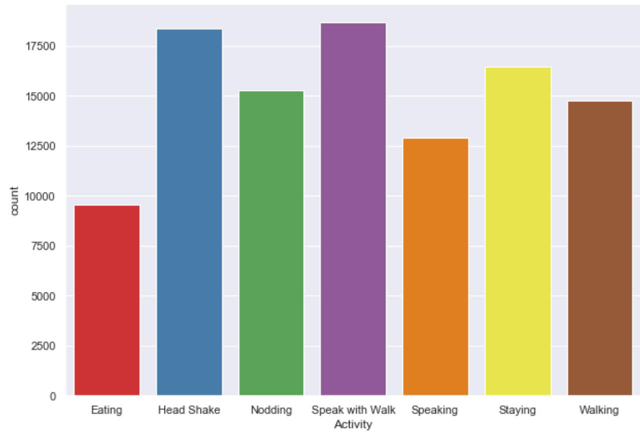


FIGURE 6.6: Bar plot for Activities in the dataset

In Fig. 6.6, the activity count from the collected data is represented. We collected data from multiple persons to split train and test data. Based on different users, we separated training and test data with different trials. These are not continuous classes. K-fold cross-validation technique was used to separate train and test data. We split the training dataset into k-folds based on the completely separate person data for training and a different person for testing. During the training procedure, we use first kth folds for model training and holdout another fold to use in testing. We use this k-fold splitting technique because in our dataset there are 6 persons' data and we want to isolate test and train data in different persons with different trials. We have used all person's data for training and testing. Normally training a model needs lots of data. We have used 1 person out cross-validation. We did not use only one person's data for training and then testing on the rest of the person's data. Rather, we used 5 persons' data for training and 1 person's data for testing.

We did not use any kind of feature selection techniques. We have calculated the magnitude value of raw accelerometer and gyroscope data for extracting features. Some of the extracted time-domain statistical features are maximum value, minimum value, mean, standard deviation, etc. We used a sliding window approach with 50% overlap. We used several traditional machine-learning algorithms, i.e., Random Forest (RnF), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) with radial basis function kernel, K-Nearest Neighbor (KNN). For the deep learning model, we used 1D CNN. For 1D CNN model we did not ex- tract any features as CNN automatically extract features. Among the four traditional machine-learning methods, KNN achieved the highest accuracy of 92.43% when we used only accelerometer sensor data. On the other hand, SVM performs better while we used both accelerometer and gyroscope sensor data together for evaluation.

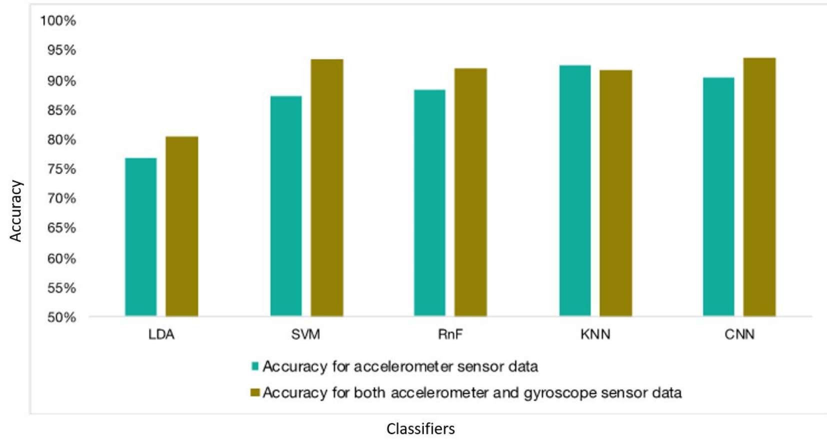


FIGURE 6.7: Accuracies comparison obtained from accelerometer sensor data and both accelerometer and gyroscope sensor data

In Fig. 6.7, the performance comparison among different classifiers is explained. Afterward, we used 1D CNN model. CNN model achieved the highest accuracy (93.76%) among all other classifiers, with the combined accelerometer and gyroscope data. In the CNN model, 2 convolutional layers were used. Each convolutional layer has 64 feature maps and the kernel size is 1x3. A max-pooling layer is used after the second convolutional layer. The pool size of the max-pooling layer is 1x2. Then two fully connected (dense) layers are used after the max-pooling layer. Finally, a soft-max layer is employed. In this model, we used ReLu activation function and Adam optimizer. These are very well-established methods for optimization and activation. In this model, there is no need to extract hand-crafted features from the sensors' data, because the CNN model can do it as per the model.

The comparison of accuracy from different classifiers by using accelerometer sensor data, as well as, by using both accelerometer and gyroscope sensor data are demonstrated in Fig. 7. On the other hand, with accelerometer-only data, the achieved recognition results are 76.75% by LDA, 87.39% by SVM, 88.23% by RnF, 92.43% by KNN, and 90.55% by CNN. We can notice from these achieved results in five classifiers that when we exploit both sensors' data, the results are superior to using accelerometer-only sensor data. Only one exception is the case with KNN. Using KNN classifier, the results are slightly better with accelerometer-only data. However, the other four cases, having both sensors' data demonstrate significantly higher recognition results (e.g., for LDA: 4%, for SVM: 6%, for RnF: 3.5%, and for CNN: 3%). Hence, we can conclude that both sensors are more suitable for better recognition results in this case.

In Fig. 6.8, we observe that only head- and mouth-related activity perform better results with the eSense. In this case, major miss-classification occurred among 'Eating' and 'Speaking' classes as these classes have closely similar patterns of muscle movement of the mouth.

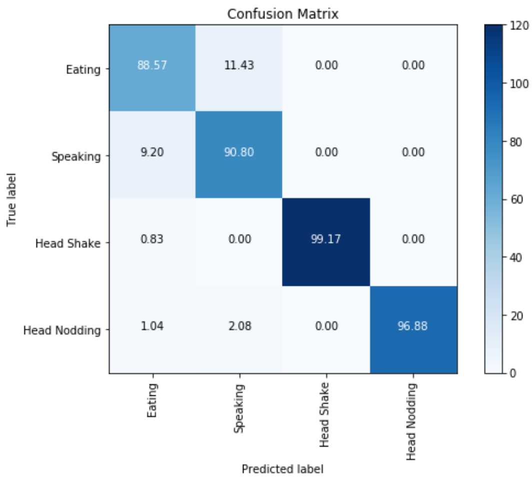


FIGURE 6.8: Confusion matrix only head-mouth activities by RnF

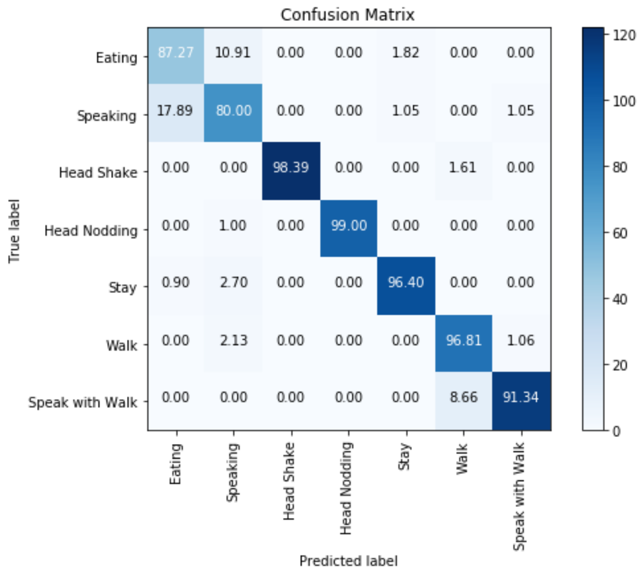


FIGURE 6.9: Confusion matrix for SVM with regular activities

From Fig. 6.9, it is found that if regular activity classes (i.e., 'Walk', 'Stay' and 'Speak with Walk') are added with the previous four head- and mouth-related activity classes, then the performance of the classifiers drops. From the confusion matrix, we noticed that the 'Speaking' class is confused with 'Eating' class 'Eating' class is mostly confused with 'Stay' class (Fig.6.9). Both 'Eating' and 'Stay' actions need less movement and these activities can be performed simultaneously. So there is a good chance of getting confused with each other for these activities. Besides, 'Eating' and 'Speaking' activities require the mouth's movement. Therefore, these 2 activities have similar patterns. Moreover, 'Speaking' activity data was recorded in the sitting state (no movement) for this study. This is the main reason for the confusion between activities. Sometimes, 'Head shake', 'Nodding' activities also get confused with 'Stay' class. The reason is, these activities were also performed while sitting on a chair (Fig. 6.10).

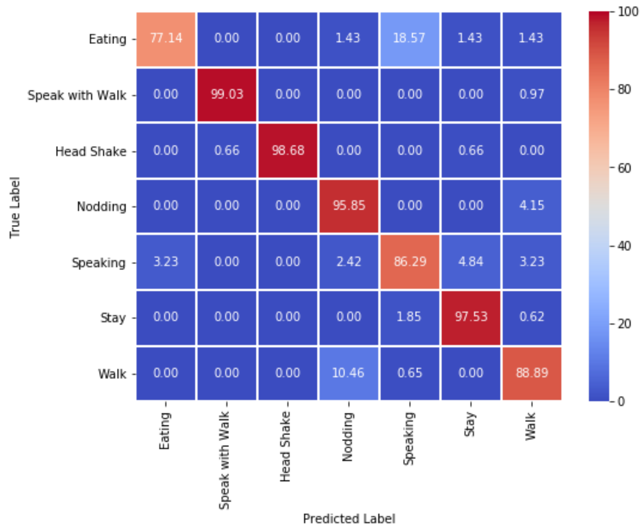


FIGURE 6.10: Confusion matrix for CNN classifier with other regular activities

From Fig.6.10 it is seen that the recognition results for the head and mouth related activities are better than the regular activities (i.e., 'Walk', 'Stay' and 'Speak with Walk').

6.4 Summary

Wearables sensors are used to recognize activities of daily life. Accelerometer and gyroscope sensors are mostly used for this purpose. There are several types of behavioral activities (i.e., nodding, drinking, eating, shaking, etc.) which are related to head and mouth. Detecting these activities accurately using only accelerometer and gyroscope sensors data quite challenging. Also, the sensor position plays a vital role to collect proper data for the head- and mouth-related behavioral activities. In this regard, an earable device like 'eSense' can play a vital role to detect head- and mouth-related activities. In this paper, we used several machine-learning methods to evaluate the performance of detecting head- and mouth-related activities as well as other regular activities from sensor data of eSense device. We compared the recognition performance of several classifiers using only accelerometer data and using both accelerometer and gyroscope sensor data. We used both traditional and deep learning methods for classification. We used Random Forest, Support Vector Machine, Linear Discriminant Analysis, K-Nearest Neighbor, and 1D Convolutional Neural Network. Having both accelerometer and gyroscope data, Convolutional Neural Network achieves the best accuracy (93.76%) among all classifiers, which is slightly better than Support Vector Machine accuracy (93.34%) among many machine learning classifiers. Our results demonstrate that using both accelerometer and gyroscope significantly improves recognition performance. However, there exist some misclassifications among different activities that require body movement and which don't require any kind of body movement. Also, head- and mouth-related activities demonstrate slightly better results than normal activities. Therefore, using our proposed framework, it will be easier for collecting head- and mouth-related activity data. Any machine-learning model can be applied to this data to detect head- and mouth-related activities. It will create a new scope in research and applications using earable device.

Chapter 7

A Bayesian Approach for Quantifying Data Scarcity when Modeling Human Behavior via Inverse Reinforcement Learning

7.1 Introduction

Computational behavioral models formalize complex human actions and contexts in which the behaviour is located. Computational modeling methods [112] allow such models to be explored by simulating situations in which individuals find themselves and predicting the actions they take in those situations. This facilitates the understanding of complex human activities and their environments through domain experts and researchers.

Inverse Reinforcement Learning (IRL) [117] has emerged as a viable alternative computational modeling approach for capturing, exploring, and predicting human behaviors [19, 20, 167]. For modeling human behavior, IRL leverage previously collected behavioral data to estimate reward function (e.g., preference that individuals have for various situations and actions) that could have led to this behavior. The reward function, rather than the policy, is the most succinct, robust, and transferable definition of the task, since it quantifies how good or bad certain actions are [117]. As well as a list of weights describing how important each behavior is [164]. IRL does probability of action given states based on the value of that action in that state and what the value of being that state. And all value is actually about the future that what more people can accomplish. In human behavior, what people are going to do next is depend on what they can able to accumulate not based on what they have done. That is the promising aspect of using IRL for modeling human behavior.

After having a promising computational model quality and amount of training data is often the single most dominant factor that determines the performance of a model. Although large quantities of high quality training data would almost always result in an accurate human behavior IRL model, it may be difficult to collect behavior cases to estimate parameters of such models. Failure to collect adequate data (under collection) would eventually lead to high model parameter uncertainty. More than required data (over collection) puts an unnecessary burden on the participants and wastes scarce resources and time. Therefore, it is important for both the conservation of resources and the accurate estimation of model parameters to have a well-informed idea of *how much data to collect*.

We look for quantitatively approach to know the amount of data to collect by using the concept of *uncertainty quantification (UQ)* [52]. The uncertainty of model parameters and how the uncertainty varies with different dataset sizes could be calculated by UQ. Increasing the number of data samples could reduces the uncertainty of model parameters. To obtain an accurate estimation of behavior model parameters, we present a UQ-based approach for estimating how much behavior data to collect. We do this for a particular IRL algorithm, MaxCausalEnt [166], which has been used to model human behaviors [19, 20] as a Markov Decision Process (MDP) [125]. We address the issue of epistemic uncertainty (i.e., model parameter uncertainty for lack of training data). Our purpose is not to determine and announce the optimum number of data points for any particular application query. Our approach helps model designers to make an informed decision about the trade-off between the resources needed to collect a variety of data points and uncertainty of model parameters that will cause from that data size.

We recast our SSD [105] problem as a Bayesian experimental [32, 114, 68] simulation-based design. The key idea behind our approach is that it is possible to update the probability of model parameters given training data from from prior to posterior through Bayesian inference [23, 24, 87, 141, 153]. We hypothesize that the probability density function (PDF) of the posterior will be narrower (on average) than that of the prior. We use Kullback-Leibler (KL) divergence [82] to quantify the shrinkage of posterior against prior. It will continue to increase as the dataset size increases. This shrinkage can then be investigated by the model designers to decide whether to add further behavior instances. We demonstrate the applicability of our method in to two separate situations: 1) *pre-hoc* experiment design—before data collection, using simulated data generated from prior distributions and the IRL model to estimate the number of required behavior instances to collect, and 2) *post-hoc* dataset analysis—with real behavior instances data collected with real people, to decide if the existing dataset size is acceptable. We used a real-life example of Multiple Sclerosis (MS) patient behavior data for modeling behaviors, a chronic central nervous system autoimmune disease [59].

In our *pre-hoc* experiment design, we generate synthetic datasets based on the parameter priors with different underlying true model parameters and dataset sizes. Then using Markov chain Monte Carlo (MCMC) [53, 28], draw samples from the posterior distributions and calculate the difference between the prior and the posterior using KL divergence for each synthetic dataset. Afterwards, we show how the expected KL divergence among synthetic datasets changes as the dataset size increases. We present a function to find the optimum size of the dataset by optimizing the estimated utility to collect a certain amount of data and comparing it with the cost of collecting data. For the *post-hoc* analysis, we repeat our experiment with a current MS dataset [91, 89, 90], with its underlying true model parameters, to obtain the KL divergence curve and compare it with the *pre-hoc* curve.

Overall, our approach serves as a decision-making support tool, but does not say how a decision should be made. For illustration, we presented an example of a particular hypothetical cost scenario and decision-making rule for MS behavioral modeling in our pre-hoc design, in which our approach indicated 935 samples is optimal. The same method is then applied to the collected MS dataset to verify our sample size determination approach. Our findings show that the predicted KL divergence of the synthetic dataset in the *pre-hoc* experimental design has a similar pattern to that of the MS dataset in the *post-hoc* dataset study. Therefore, the conclusion about the appropriate size of the dataset that we draw can then be applied to the MS dataset at the pre-hoc stage and recommends data collection with 228 more data samples.

Our research helps model designers to conduct a more in-depth, principled analysis of data set size effects on IRL model parameters. It enables the determination of sample size, one of the hallmarks of scientific science, as a precursor for developing more accurate models of human behavior.

7.2 Method

Based on principles of uncertainty quantification (UQ), we implement a Bayesian method to quantify the effects of dataset size (i.e., number of data samples collected) on parameter estimation for a particular human behavior model based on IRLs [19]. We apply our method to a particular human behavior model [19] that uses the MaxCausalEnt algorithm [166] to estimate parameters of the model from data from behavior instances. We begin by explaining the IRL model's classical construction, followed by a presentation of the Bayesian framework [87, 141, 153].

7.2.1 Modeling Human Behavior with Markov Decision Process

Markov Decision Processes (MDP) can represent a decision process following a graphical structure of states, actions, rewards and stochastic transition between states, this can be helpful to model different context and the preference of action people have for different actions in different situations. Following [19], we formalize purposeful behaviors as a Markov Decision Process (MDP) [125], represented by a tuple:

$$\mathcal{M}_{MDP} = \left\{\mathcal{S},\ \mathcal{A},\ R(s,a),\ P_0(s),\ P(a \mid s),\ P(s' \mid s,a)\right\}. \tag{7.1}$$

Here $s \in \mathcal{S}$ is the state from a finite state space representing different situations that a human participant (i.e., agent) can be in, and $a \in \mathcal{A}$ is the action from a finite action space that the participant can take. The reward function $R(s,a)$ defines the reward that the participant incurs when performing an action a while in the state s. The reward function therefore reflects the preference of the participant to be in different situations and to take different actions in those situations.

The initial state probability $P_0(s)$ captures the probability that a state initiates a behavior instance. The conditional probability of actions given states $P(a \mid s)$ represents the probability that the participant will choose to perform a particular action a in a particular state s. The action-dependent transitional probability $P(s' \mid s, a)$ designates the probability of transition into the next state s' after the participant performs action a in state s. The transitional probability thus captures how the actions that participants choose in various situations influence their surroundings.

7.2.2 Estimating Model Parameters via MaxCausalEnt IRL Algorithm

In an IRL problem setting, $P_0(s)$, $P(s' \mid s, a)$, $P(a \mid s)$, and $R(s, a)$ are all unknown, and the goal is to estimate them from an available dataset of observed behavior instances, where each behavior instance sample is defined as a sequence of state and action pairs over time:
$\left\{\left(s_1^{(i)}, a_1^{(i)}\right), \left(s_2^{(i)}, a_2^{(i)}\right), \ldots, \left(s_T^{(i)}, a_T^{(i)}\right)\right\}$, where $i = 1, \ldots, n_d$ denotes the ith sequence sample of length T in the dataset for a total of n_d samples.

Following the framework introduced in [19], $P_0(s)$ and $P(s' \mid s, a)$ can be estimated by constructing two separate Bayesian networks [25] from the dataset samples. Furthermore, we can define a linear parametric reward function:

$$R(s, a) = \theta^T \mathcal{F}_{s,a}, \tag{7.2}$$

where $\mathcal{F}_{s,a}$ is a general feature vector and θ is a vector of unknown weight parameters representing the strength of preferences for individual features.

MaxCausalEnt IRL algorithm [166] then estimates the parameter θ of the reward function $R(s, a)$. MaxCausalEnt IRL algorithm solves the optimization problem using the Stochastic Gradient Descent (SGD) [25] over the reward function parameter θ. At each gradient step, it iteratively computes action-based value function $Q_\theta^{\text{soft}}(s, a)$ that represents the expected value of performing a specific action a in a specific state s, and state-based value function $V_\theta^{\text{soft}}(s)$ that represents the expected value of being in a specific state s:

$$Q_\theta^{\text{soft}}(s_t, a_t) = \sum_{s_{t+1}} P(s_{t+1} \mid s_t, a_t)\, V_\theta^{\text{soft}}(s_{t+1}) \tag{7.3}$$

$$V_\theta^{\text{soft}}(s_t) = \operatorname*{softmax}_{a_t} \left\{Q_\theta^{\text{soft}}(s_t, a_t), c\right\} + \theta^T \mathcal{F}_{s_t, a_t} \tag{7.4}$$

where $\text{softmax}_x\{f(x), c\} := \frac{1}{c} \ln \sum_x e^{cf(x)}$ and c is a hyperparameter. These two value functions are then used to compute the policy *via* following equation:

$$P(a_t \mid s_t) = \frac{1}{Z(s_t)} e^{c\left(Q_\theta^{\text{soft}}(s_t, a_t)\; - \;V_\theta^{\text{soft}}(s_t)\right)}, \tag{7.5}$$

where $Z(s_t) = \sum_{b \in \mathcal{A}} e^{c\left(Q_\theta^{\text{soft}}(s_t,b) - V_\theta^{\text{soft}}(s_t)\right)}$ ensures the overall expression is a proper probability mass function. The SGD algorithm uses the stochastic policy $P(a_t \mid s_t)$ in a forward pass to update the estimated expected feature counts $\mathbb{E}_{P(s,a)}\left[\mathcal{F}_{s,a}\right]$, and updates the parameter θ using the following equation until convergence:

$$\Delta\theta = \mathbb{E}_{P(s,a)}\left[\mathcal{F}_{s,a}\right] - \mathbb{E}_{\widetilde{P}(s,a)}\left[\mathcal{F}_{s,a}\right]. \tag{7.6}$$

Overall in the MaxCausalEnt IRL algorithm [166], the dataset, and the the number of sequence samples n_d in the dataset, affects the estimation of $P_0(s)$, $P(s' \mid s, a)$, and $\mathbb{E}_{\widetilde{P}(s,a)}\left[\mathcal{F}_{s,a}\right]$. While MaxCausalEnt IRL provides an efficient method for estimating θ, it produces a single point estimate only and does not offer a measure of uncertainty or confidence surrounding the estimated value due to n_d.

7.2.3 Bayesian Inference for Quantifying Parameter Uncertainties

Bayesian inference [141, 153] is a framework that uses probability and statistical formalism to rigorously measure uncertainty. It is a natural mechanism for integrating sparse, noisy, and incomplete data from various sources. The Bayesian inference for the IRL problem is referred to as the Bayesian IRL (BIRL), initially proposed by [127]. Under the Bayesian framework, we treat θ as a (continuous) random vector with an associated probability density function (PDF). Given a dataset with n_d behavior instances (sequence samples) $D = \left\{\left(s_t^{(i)}, a_t^{(i)}\right)\right\}$, $t = 1, \ldots, T$, $i = 1, \ldots, n_d$, the uncertainty of our unknown model parameter θ is updated via Bayes' rule

$$p(\theta \mid D) = \frac{P(D \mid \theta) p(\theta)}{P(D)}, \tag{7.7}$$

where $p(\theta \mid D)$ is the posterior PDF[1], $p(\theta)$ is the prior PDF, $P(D \mid \theta)$ is the likelihood function (i.e., the probability density of having observed the state and action trajectories in D if the true feature weights were θ), and $P(D)$ is the Bayesian evidence (a normalization constant). For a Bayesian inference problem, we generally have the ability to evaluate the prior and likelihood PDFs. Usually the model designer or a domain expert would select the prior, which represents their knowledge or belief about θ before having seen any data. The likelihood can be directly derived using the Markovian structure of the MDP model with the assumption of independence among different behavior instances [127].

7.2.4 Quantifying Data Scarcity

Following the Bayesian framework that we introduced in the previous section, we quantify data scarcity based on the extent of uncertainty reduction we achieve in learning θ. Specifically, we employ the Kullback-Leibler (KL) divergence [93] from

[1] We use lower case $p(\cdot)$ for PDF of a continuous random variable or vector, and upper case $P(\cdot)$ for probability mass function of a discrete random variable or vector.

the prior to the posterior:

$$D_{\mathrm{KL}}(p(\theta \mid D) \,||\, p(\theta)) = \int_{\Theta} p(\theta \mid D) \ln \left[\frac{p(\theta \mid D)}{p(\theta)} \right] d\theta. \tag{7.8}$$

The KL divergence is non-negative, and equals zero if and only if $p(\theta \mid D) = p(\theta)$. We need to take the expectation over all possible realizations of D, to arrive at the final *expected KL divergence* (or expected information gain (EIG)), where we explicitly show the dependence on n_d. In general, the KL divergence has no closed-form and must be approximated numerically. We adopt a Monte Carlo estimator [130] using posterior samples generated from the aforementioned MCMC algorithm:

$$\mathrm{EIG}(n_d) \approx \frac{1}{LM} \sum_{j=1}^{L} \sum_{k=1}^{M} \left[\ln p(\theta^{(j,k)} \mid D^{(j)}) - \ln p(\theta^{(j,k)}) \right], \tag{7.9}$$

where L and M are the Monte Carlo sample sizes in this estimator. Intuitively, one may expect EIG to increase as more data becomes available (larger n_d), but the additional benefit from each new sample may diminish as the overall dataset grows. Finding the critical point where the rate of benefit is no longer worthwhile can be used to guide the decision-making of sample size determination.

7.3 Sample Size Determination for a Model of Multiple Sclerosis Patient Behavior

We demonstrate our method for evaluating the sample size determination of a behavioral model of multiple sclerosis people (MS) [168]. People with MS suffer physical disability and constant pain, tiredness, depressed mood, and cognitive difficulties. Symptoms are linked to various negative effects, including unemployment, disability, social impairment, dissatisfaction with life, interference with everyday activities, worsening in general mental and physical health, and loss of group integration. Thus, approaches that address the most significant or impactful symptoms, such as pain and fatigue, may enhance people's health outcomes and their quality of life. Such a model of behavior is of great interest to clinicians to test the hypothesis that the everyday living behaviors of participants that we can feel, remember, and capture can predict their discomfort, exhaustion, and overall well-being. Our aim is therefore to provide guidance to model designers in determining how many samples (i.e., how many behavioral instances of individuals with MS) to collect for modeling people's behavior. We then illustrate the effects of dataset size under two scenarios:

- *pre-hoc* experiment design, conducted in the planning stage before any data collected, to guide the estimation of how many samples to collect; and

- *post-hoc* dataset analysis, performed when data is already collected, to determine if there are sufficient samples in the current dataset and whether more data is needed.

To model behaviors of people with MS, we follow the modeling approach from [19] and use the MDP framework described in Section 7.2. For this investigation, we have consulted a domain specialist from our institute who is a Physical Medicine and Rehabilitation Research Non-Clinical Psychologist specialized in MS. Here we explain how we model various aspects of the actions of people with MS, and in the following sections we provide information about how we estimate the parameters of the model in both *pre-hoc* experiment design and *post-hoc* dataset analysis.

7.3.1 States and Actions

Here, we explicate our state space $\mathcal{S}$ and action space $\mathcal{A}$ by defining state and action features (Tables 7.1 and 7.2). These features in turn define feature vectors $\mathcal{F}_{s,a}$. State features (Table 7.1) identify the demographics of participants (gender and age), contextual details, such as the time of day (wake, morning, afternoon, evening and bed), self-reported symptoms and health measures (i.e., momentary evaluation of pain and fatigue at each time of day, self-reported positive effects and well-being (PAW) [137]). Action features (Table 7.2) reflect the participant's objective measure of activity strength and frequency (e.g. as calculated by the ActiGraph watch) and whether or not the participant has completed a momentary evaluation of their symptoms and their PAW (i.e. end-of-day functional outcome) at each time of the day.

Feature	**Description**
Gender	Gender of the patient {Male, Female}
Age	Age of the patient {Younger than 30, Between 30 to 60, 60 and Older}
Current Daytime Interval	Time of the day {Wake, Morning, Afternoon, Evening, Bed}
Current Pain	Current interval pain score {Low, Medium, High, Not Recorded}
Current Fatigue	Current interval fatigue score {Low, Medium, High, Not Recorded}
Last Activity Bouts	Last interval activity bouts based on average activity bouts per 15s from an ActiGraph watch {Low, Medium, High, Not Recorded}
Last Activity Pace	Last interval pace (determines whether last activity was performed with/without breaks) {Low, Medium, High, Not Recorded}
End-of-Day Positive Affect and Well-being	Positive impact on bed interval signifies how much positive impact (sense of well-being, feeling hopeful and satisfying, cheerful, etc.) the patient had on that particular day (recorded at bedtime only) {None (Not Applicable), Moderate, Mild, Normal, Not Recorded}

TABLE 7.1: State features that define the different situations the a participant can be in.

Feature	Description
Current Activity Bouts	Current interval activity bouts based on average activity bouts per 15s from an ActiGraph watch {Low, Medium, High, Not Recorded}
Current Activity Pace	Current interval pace (determines whether current activity was performed with/without breaks) {Low, Medium, High, Not Recorded}
Record Next Pain	Status of next state pain {Recorded, Not Recorded}
Record Next Fatigue	Status of next state fatigue {Recorded, Not Recorded}
Record Next Positive Affect and Well-being	Status of next state positive affect and well-being {Not Applicable, Recorded, Not Recorded}

TABLE 7.2: Action features representing different actions that participant can perform in each situation.

7.3.2 Behavior Instances

We define a behavior instance as a sequence of state and actions that captures situations that a particular participant found themselves in and the actions they performed in those situations each day. Thus, we treat each participant's day as *one sample* one of n_d samples. Because each behavior instance reflects behaviors of a single participant, *Gender* and *Age* variables remain the same throughout a behavior instance.

7.3.3 Estimating Initial State Probabilities

Each participant's behavior instance starts with an initial state s with probability $P_0(s)$ and $Current\_Daytime\_Interval = Wake$. We construct a Bayesian network [25] to estimate initial state probabilities $P_0(s)$. We consider 6 features defining each state: age, gender, pain level, fatigue level, last activity bout, and last pace. *Age* and *Gender* both influence *Pain* and *Fatigue*, which are mutually independent, as are *Age* and *Gender*. On the other hand, we can observe that *LastAcbout* and *LastPace* only influence level of *Fatigue* and they (*LastAcbout* and *LastPace*) are mutually independent. Thus, we compute the initial state probability as follows:

$$\begin{aligned}
P_0(s) &= P(Age, Gender, Pain, Fatigue, LastAcbout, LastPace) \\
&= P(Pain|Age, Gender) \times P(Fatigue|Age, Gender, LastAcbout, LastPace) \\
&\quad \times P(Age) \times P(Gender) \times P(LastAcbout) \times P(LastPace).
\end{aligned} \tag{7.10}$$

where $s \in \mathcal{S}_0$ and $\mathcal{S}_0$ represents all possible initial states.

7.3.4 Estimating State Transition Probabilities

Each state transition is driven by the action that the participant performs and changes in participants' symptoms irrespective of their actions, as influenced by their demographics, time of day, and previously reported symptoms and healthcare outcomes. To capture the probability of these transitions, we built another Bayesian network to estimate state transition probabilities $P(s' \mid s, a)$. Here, we used all of the features from current state s and action a to estimate the joint probability of pain, fatigue, and positive affect and well-being in state s', which corresponds to the probability of next state s'. This is because all of the other state features are deterministic: $DaytimeInterval$ transitions are fixed, Age and $Gender$ stay the same at each transition, and value of $LastActivityBouts$ in s' is the same as in action a.

7.3.5 Estimating Action Probabilities

In this model, the participants decide on their next action based on a reward function $R(s, a)$ (Equation 7.2), which represents the preference that people with MS have for certain situations (e.g., specific pain and fatigue levels) and performing certain actions in those situations. Given the model parameters θ, we use the MaxCausalEnt IRL algorithm [166] (Equation 7.5) to estimate the conditional probability of participants' actions given their current situation $P(a \mid s)$. This stochastic policy captures the probability of each participant action according to their preference for various features of states and actions.

7.3.6 *Pre-hoc* Experimental Design and *Post-hoc* Dataset Analysis

To demonstrate and validate our method, we used an existing MS dataset [91, 89] containing 749 behavior instances collected from a total of 107 participants with MS. We split this dataset into two subsets to illustrate *pre-hoc* experimental design and *post-hoc* dataset analysis: 1) *pre-collected* dataset with 6 participants resulting in 42 behavior instances (i.e., samples), and 2) *post-collected* dataset with 101 participants, or 707 samples. To avoid any statistical contamination (i.e., 'cheating') between pre-hoc experimental design and post-hoc dataset analysis, we kept these two subsets strictly apart.

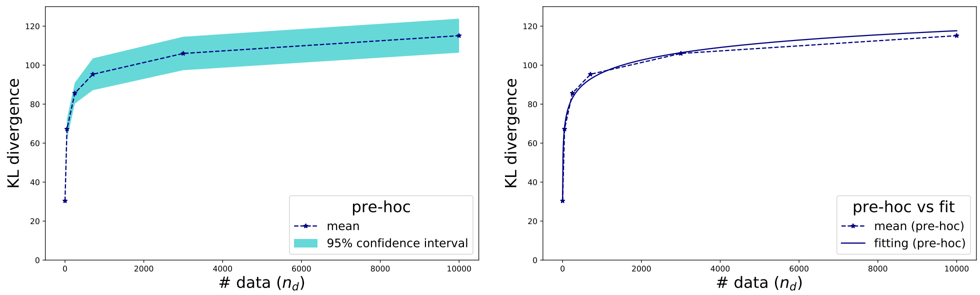
We performed *pre-hoc* experimental design for determining the sample size for the main data collection using only the 42 pre-collected samples (and without using or seeing any of the 707 post-collected samples). We then mimicked the main data collection using the remaining 707 samples and conducted a *post-hoc* dataset analysis to update the model uncertainty and provide guidance on whether additional data samples are needed.

7.4 Results

The results of the *pre-hoc* experiment design and the *post-hoc* dataset analysis for the MS dataset are discussed. We compute the expected information gain of model parameters as a function of dataset sample size ($\text{EIG}(n_d)$). We then illustrate how an optimal sample size can be determined from an example decision-making rule.

7.4.1 *Pre-hoc* Experiment Design

Here, we estimate how many samples to collect for different costs of data collection. We computed $\mathrm{EIG}(n_d)$ using the Monte Carlo estimator in Equation 7.9, with $L = 60$ and $M = 92,000$ (92 MCMC chains in parallel, each with 1,000 samples). Figures 7.1a shows the mean of KL divergence curves ($\mathrm{EIG}(n_d)$) and their 95% confidence intervals (i.e., ± 1.96 standard deviations). The figure offers a quantitative summary of the tradeoff between information gain and number of data samples. We observe a sharp initial increase of the EIG followed by a gradual flattening of the curve—a "kneebend"-like transition—which is consistent with our intuition of diminishing return as the total dataset size grows. Furthermore, the EIG curve appears logarithmic. Fitting EIG to a logarithmic function, we obtain $\widehat{\mathrm{EIG}}(n_d) = 9.35 \ln(n_d) + 31.51$ and appears to provide excellent agreement with the computed points (Figure 7.1b); this fitted function can then be used to guide future interpolation and extrapolation studies.



(A) Pre-hoc (B) Pre-hoc vs fitting

FIGURE 7.1: KL divergence versus n_d for *pre-hoc* experiment design. The blue dashed line represents the mean of KL divergence curves (i.e., the EIG), the light blue shadow represent the 95% confidence interval (± 1.96 standard deviations), and the blue solid line is the logarithmic fit.

Next, we present an example of how this EIG curve can be used to evaluate the size of the sample in the *pre-hoc* experiment design. We emphasize that the methods set out in Section 7.2 are a general decision-making process, not a decision-making mechanism. To determine the optimum dataset sample size, we demonstrate a particular example to compute the optimal dataset sample size n_d^* that maximizes an utility composed of a simple weighted sum between EIG and cost of data acquisition [103]:

$$n_d^* = \arg\max_{n_d} \{\mathrm{EIG}(n_d) - c \cdot n_d\} \tag{7.11}$$

where c denotes the ratio of unit cost for collecting one behavior instance to one nat of information gain [40] (e.g., in units of [$/nat]). Put another way, c is the unit cost of information: how much you need to pay for information learned from the data in the experiment? Figure 7.2 shows the optimal sample size n_d^* versus c using the fitted $\widehat{\text{EIG}}$. As expected, n_d^* decreases as c increases, and appears linear on a log-log plot. For example, if $c = 0.01$ [$/nat] for our MS problem setting, then the optimal sample size would be $n_d^* = 935$.

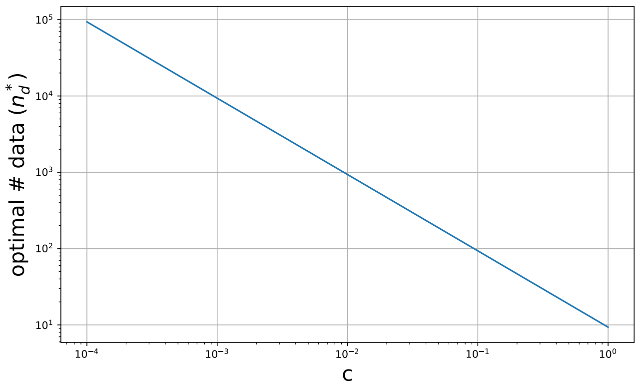


FIGURE 7.2: Optimal dataset sample size versus the unit cost of information c (*pre-hoc*).

7.4.2 Post-hoc Dataset Analysis

Here, our goal is two fold: 1) to validate our *pre-hoc* experimental design, and 2) to estimate if an existing dataset contains enough samples for different costs of additional data collection. Computing for *post-hoc* dataset analysis, Figure 7.3 shows the actual information gain in purple based on the 707 post-collected data instances collected in real life. The uncertainty in Figure 7.3 is due to shuffling the order of data points (e.g., for $n_d = 50$ there are many different ways to choose 50 from the total of 707) and randomness from MCMC sampling. Indeed, in Figure 7.3, both curves share the same trend, the *post-hoc* curve (in pink) has lower uncertainty compared to the *pre-hoc* curve (in blue), and a *pre-hoc* decision based on the its *pre-hoc* curve is very representative of the actual dataset collected afterwards.

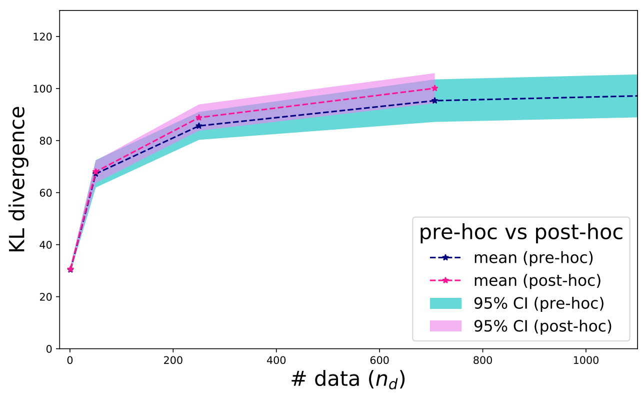


FIGURE 7.3: KL divergence versus data number of synthetic data and true data (comparison between *pre-hoc* design and *post-hoc* analysis).

7.5 Discussion

The results in the previous section serve as a diagnostic tool for sample size determination. In particular, the results of *pre-hoc* allow one to determine how many new samples to be collected, and the results of *post-hoc* provide an evaluation of whether the current dataset is sufficient. In our demonstration, *post-hoc* also acted as a confirmation of the overall process, where the real knowledge gain pattern is captured within the *pre-hoc* EIG. These tools are especially useful when the cost of data is expensive, when careful decision-making about how much data to collect can lead to substantial resource savings, as well as the alleviation of potentially undue pressures on participants to undertake the data collection process.

We reiterate that our framework and methodology offers *support* for decision-making, but do not make decisions on the collection of sample sizes. A decision includes many additional considerations, such as the basic objectives of the model and the use of data, the monetary cost and scheduling of experiments, the importance of information and expertise, the implications and threats, the regulatory and policy criteria, and even the degree of risk-aversion of the decision-maker. It is exceedingly difficult and beyond the scope of our work to integrate these components systematically and comprehensively, as sought in the study of decision-making theory (e.g., [23, 121]).

Instead, we provide a certain example where *if* a decision-maker come up with the information of how much each data costs and how much they are willing to pay for one unit of information gained from the data, then our framework can return the optimum number of data samples n_d^*. The methods that we provide are therefore the starting point for a potentially complicated decision-making process by providing a quantitative evaluation of the information quality provided by the data.

Chapter 8

Discussion

In this thesis, it is a step towards addressing the grand challenges of establishing practical foundation for work in nursing domain which will open the possibilities and scopes to explore and understand human activities and behavior by using technology to support their life. The work on exploring and understanding complex and regular human activity has direct implications on study of intelligibility of technology for well-being and healthcare domain. In this regard, it is essential to understand the unique nature of human activities as the information extracted from raw activity data has been shown to be crucial to functional and behavioral health monitoring. Data mining and machine learning methods have proved to be successful in extracting information and finding, inferring activity from data. Activities related to fitness are common among young adults. And it also helps them to keep track of their fitness on a regular basis. The activities that every individual does on a regular basis are functional activities such as *preparing food*, *talking with others*, *washing hands*, *brushing teeth*, *wearing coats*, *shoes*, *answering the door*, *writing a check*, *taking medicine* etc. The inference and assessment of such functional and behavioral activities helps to decipher personal health and well-being.

Thus the overall contribution of this thesis is to introducing different computational approaches in various modalities for better functional and behavioral activities understanding in nursing care domain by addressing the challenges related to data collection, complex activity understanding and behavior modeling. Through this thesis, we targeted a few problems and challenges related to better activity understanding in nursing care domain. This thesis introduce various methods and frameworks for improving domain specific activity understanding specially for nursing care domain. The proposal and framework from this thesis will contribute better understanding of daily life activity in nursing care center. Also, the proposed sample size determination method in this thesis will contribute as a decision making support for model designers to better understanding the optimum amount of dataset size for better modeling the patient behavior. Hence, the summary contribution of this thesis are:

Method for improving Sensor-Based Activity Recognition in Missing Data Scenario.

In this thesis, our proposed method for improving sensor-based activity recognition in missing data scenario will be helpful to improve activity recognition while having missing data without any data recovery. We examined the method for different missing percentages, varied window sizes, and diverse window sliding widths. Learning with missing data reinforces our model to regulate missing data during the classification of various activities that have missing data in the test module. This approach demonstrates the plausibility of the machine learning model, as it can learn and predict from an identical domain. For the missing data pattern, we considered data to be missing in a random pattern, which is a realistic missing pattern for sensor data collection.

For evaluating our method's acceptability, we developed a synthetic dataset to empirically evaluate the performance and show that the method can effectively improve the recognition accuracy from 80.8% to 97.5%. Based on the successful analysis, we then explored activities from two very challenging benchmark datasets for our experiment. These are the HASC dataset and the single chest dataset. Under various experiments of random missing data levels, we recognized activity classes. From our experimental analysis, we can conclude that the recognition results improve by our proposed method when missing data patterns are available in both training and testing modules without recovering the raw data.

Proposed an activity recognition framework by exploiting LoRaWAN (Long Range Wide Area Network) protocol for Nursing Care

We propose a framework for activity recognition by using LoRaWAN protocol and network to send various sensor data. This work explores exciting new opportunities to significantly increase the sensing range with the introduction of LoRaWAN for Nursing Care Center. After having a activity recognition framework in laboratory settings experiment. We examined the performance of the LoRaWAN in both laboratory environment and real nursing care environment. We evaluate LoRaWAN accelerometer sensors data for human activity recognition by using our proposed framework. We explore the Linear Discriminant Analysis (LDA), Random Forest (RnF) and K-Nearest Neighbor (KNN) for classification. We achieve recognition accuracy of 94.44% by LDA, 84.72% by RnF and 98.61% by KNN.

In LoRaWAN technology, the amount of sensor nodes connected with a single gateway have an impact on the performance of sensors ultimate data sending capability in terms of data loss. We investigated this issue from real nursing care data to check the feasibility of using LoRaWAN sensor for future healthcare monitoring center. In a real nursing care we did 4-month experiment to collect sensor data. During this experiment, 42 LoRaWAN sensors environmental sensors data is collected to know the data loss ratio. We observe 5%

data loss happened by the sensors with a single gateway. In a simulated environment, we checked the activity recognition performance with 5%, 30%, 50% and 80% data loss environment and have found recognition accuracy of 81.94% LDA, 80.55% RnF and 91.66% by KNN while 5% data are lost. Through our proposed framework it can open a new opportunity to significantly increase the sensing range in nursing care center by LoRaWAN.

Revealed the nature of nursing care activities from real-life data

In this thesis, we suggest a framework that combines activity records and activity label records that are routinely used by workers in the nursing sector. In nursing care or any domain specific, activity recognition is challenging because the nature of activities is not clear. Hence, this thesis revealed the nature of real settings activity data for nursing care domain. We have collected 4-months nursing care data from real nursing care center in Japan. It has been shown that by using our system it can possible to increase the number of label collections by easy recording. We have collected 38,076 activity labels, 46,803 record details, and 2834 hours of sensor data during this experiment.

From this data, we revealed the varieties, nature and dependency of activities and care details which can be a measure of any healthcare outcome. We analyze intra-class relationship of real life nursing care activity data which has important information to revealed the nature of real life data. We investigated the dependency of activities to staff users, target residents, and days. Using the obtained data, we revealed the nature of the data, including dependency of activities to several factors, and the nature of timestamps of self-labeling. We investigated "staff activity" for support service requirements as well as "elderly people activity" from this data. This work also help to examined elderly people health status from the data and types of activities recorded by staff.

Propose a framework to detect head and mouth related personal scale behavior activities which activities are difficult to detect with usual wearable sensors with other regular activities

We can measure nurse activities, but patient's are also very important to make complete records, therefore in this thesis we propose a framework using eSense earable to detect head and mouth behavioral activities with good accuracy. We develop a smartphone application for data collection from the eSense. Through eSense earables, we performed experiments for seven activities. Four activities are related to head and mouth (namely, *eating*, *speaking*, *headshaking*, and *head nodding*). Three regular activities (i.e. *walk*, *stay*, and *speaking while walking*). We have found that the accuracy 93.34% by SVM, 91.92% by RnF, 91.64% by KNN, and 93.76% by CNN. Detecting these activities accurately using only accelerometer and gyroscope sensors data quite challenging. By using our propose framework, it will be possible to detect complex head-mouth related behavioral activity along with regular activity in nursing care center.

A Bayesian approach for quantifying data scarcity when modeling human behavior via Inverse Reinforcement Learning (IRL)

We proposed a Bayesian approach for quantifying data scarcity to modeling human Behavior via Inverse Reinforcement Learning. Our proposed sample size determination method can be a decision making support to model designers to estimate the number of samples (dataset size as part of data collection) for modeling human behavior. After having a promising computation model, it is always important to know the amount of data needed to train the model. It is necessary to have a well-informed idea of how much data to collect for both resource conservation and to obtain an accurate model parameter estimate. In the thesis, we proposed a sample size determination approach based on uncertainty quantification (UQ) for a specific Inverse Reinforcement Learning (IRL) model of human behavior.

The main insight behind our method is that the probability of model parameters given training data can be updated from prior to posterior through Bayesian inference. For illustration, we provided an example with a specific hypothetical cost scenario and decision-making rule for MS (Multiple Sclerosis) behavioral modeling, under which our method indicated 935 samples is optimal. Our approach can be use as a tool to help decision-making for model designers to selecting optimum amount of data when modeling patient behavior.

In future technology will help people to be productive, comfortable, healthy, and safe. Thus human-data supported interfaces will create technology that will automatically reason and explain common user behaviors, infer their priorities, predict future user behaviour, and even will able to coach users to strengthen their behaviors. The user interface of medical informatics helps physicians understand patient data, identify chronic problems, and find the right care that is tailored to the patient based on the patient's behaviours in almost every aspect of their lives.

Chapter 9

Conclusion and Future Work

9.1 Conclusion

The nursing domain is one of the domains that can benefit greatly from daily life human activity understanding research. Though daily life human activity understanding in the real field is immensely challenging due to various levels of granularity of human action and data collection from wild. Therefore, we contribute to addressing a few real-life activity recognition challenges in this thesis. The thesis aims to incorporate various computational methods, addressing the problems related to data collection, complex activity understanding and behavioral modeling. We focus on activity recognition by introducing different computational approaches that recover missing data and utilize spatial and temporal contexts of the data to achieve higher recognition accuracy in various modalities. In this thesis we proposed a method to improve activity recognition while having missing data without any data recovery. We considered data to be missing in a random pattern, which is a realistic missing pattern for sensor data collection. We developed a synthetic dataset to empirically evaluate the performance and show that the method can effectively improve the recognition accuracy from 80.8% to 97.5%. Afterward, we tested our approach with activities from two challenging benchmark datasets: the human activity sensing consortium (HASC) dataset and single chest-mounted accelerometer dataset. This approach can be a rudimentary step so that, based on any realistic model, we can explore human activity recognition in genuine nursing home or healthcare facilities, where missing data are more prevalent. Later, we propose a framework for activity recognition by using LoRaWAN protocol and network to send various sensor data. After having a activity recognition framework in laboratory settings experiment. We examined the performance of the LoRaWAN in both laboratory environment and real nursing care environment. This work explores exciting new opportunities to significantly increase the sensing range with the introduction of LoRaWAN for Nursing Care Center. Afterwards, in this thesis, we suggest a framework that combines activity records and activity label records that are routinely used by workers in the nursing sector. In nursing care facility, activity recognition is challenging because the nature of activities is not clear. We analyze intra-class relationship of real life nursing care activity data which has important information to revealed the nature of real life data and will be helpful for any application settings in nursing care domain. We have found that in real-life settings, some activities are more dominant than most of the other activities, as recorded by the caregivers. Also some activities are most widely explored activities for the residents based on the demand basis. We can estimate the activity level of a resident and his/her corresponding health status from this data. These cues are detrimental to predict resident's other activities associated in regular basis. From this care record data, we can find out essential support requirements and their other activity performance in any nursing care facility. Later, we proposed a framework to detect head and mouth related complex personal scale behavioral activities (like speaking, eating, head movement etc.) Detecting these activities accurately using smartphone sensors data is quite challenging. In our evaluation, we have used both machine learning and deep learning models to evaluate the results. Having both accelerometer and gyroscope data, Convolutional Neural Network achieves the best accuracy (93.76%) among all classifiers, which is slightly better than Support Vector Machine accuracy (93.34%) among many machine learning classifiers. By using our propose framework, it will be possible to detect complex head-mouth related behavioral activity along with regular activity in nursing care center. Finally, in this thesis, we proposed a sample size determination method based on uncertainty quantification(UQ) for a specific Inverse Reinforcement Learning(IRL) model of human behavior where computational behavioral models formalize complex human actions. The main insight behind our method is that the probability of model parameters given training data can be updated from prior to posterior through Bayesian inference. For illustration, we provided an example with a specific hypothetical cost scenario and decision-making rule for MS (Multiple Sclerosis) behavioral modeling, under which our method indicated 935 samples is optimal.

This thesis contribution can help to create various tools to aid stakeholders, such as domain experts and end users, in exploring human behavior understanding. In this regard it is important to understand existing challenges for well-defined environments to perform specialized tasks. Future work should explore various methods to direct stakeholders to use a mixed-initiative learning approach to apply their understanding of the complexities of the world.

9.2 Understanding Limitations and Challenges

To understanding daily life functional and behavioral activities, there are few work of actually conducting research on complicated activities understanding at work sites. Since it may interrupt nurses or caregivers' daily duties. So it is difficult to progress domain specific daily complex activity understanding research in a vast way. While in the laboratory experimental area, we can set up pre-defined activity groups and properly set up measuring instruments such as video cameras or sensors. A controlled experiment using motion capture and mobile sensors showed that nursing behaviors can be identified with high accuracy by motion sensing. However, in real life, it is difficult to manage all these parameters in the right way. Patients also do not want to wear sensors for behavioral data collection all day long. Video-based systems may be an option, but privacy concerns have also been a major concern for researchers over the last decade. In spite of the fact that privacy and security issues are removed, the wearable Sensors also present a range of distinctive obstacles, such as collecting activity labels, intra-class variability, timestamps are not accurate, class imbalance, intra-class dependency, sensing devices, sensor device positioning. It is also difficult to find the exact beginning and ending point of the activity episode, given that the sensors typically have a higher sampling frequency. Thus, evaluating the recognition of real-life activity is not straightforward. Timestamp extension can be improved by considering real activity average duration. There is also a method for representing time extension probabilities and optimizing the probability of label timestamps with the EM algorithm [148]. However, because these approaches use repeated optimization calculations not only for training but also for testing, training before convergence is not realistic. Moreover, how long extensions we can give for initial value for optimization is not clear [148]. It's still an open issue how to find right timestamps for self-labels. For the development of this domain specific activity understanding research, public data sets for the community are necessary. Data is required for evaluating the proposals for learning and testing in future.

9.3 Future Research Perspectives

Given the growth in the world's elderly population, smart nursing care center or assisted living center are a demand of much-needed solution. There are many challenges, to monitor and help elderly people by using technology in a smart nursing care center. This thesis is a contribution to improve few challenges in care system related to data collection, complex activity understanding and activity recognition by introducing different computational approaches that recover missing data and utilize spatial and temporal contexts of the data to achieve higher recognition accuracy in various modalities. There are still some problems in making smart nursing care center a reality for the elderly people. One key future research direction is studying the effect of the location of the sensor as elderly people do not want to wear sensors for behavioral data collection all day long. It is necessary to develop and implement custom sensors for this domain, and studying the best features in this specific domain.

In the future to directly extend the missing data analysis, it is required to explore for more features to enhance the recognition results of missing data at random order; however, modeling of the realistic nature of missing data under all major circumstances can be explored by network-based researchers. This can be a rudimentary step so that, based on any realistic model, we can explore human activity recognition in genuine nursing home or healthcare facilities, where missing data are more prevalent. The impact of missing data for each class wise is necessary to investigate. As well as smarter datasets are required having realistic missing data from various sensors and there explorations.

Other challenges related to the extension of this work refer to continue LoRaWAN technology in healthcare domain. It has has promising prospect due to its long range, low data rate communication and low power communication. If we can use this technology for human behavior analysis from accelerometer and environmental sensors data then it will be promising prospect for healthcare settings. We have found the considerably good recognition performance for activity sensing data. Real nursing care center evaluation time it has been observed that some sensors got idle and sensor data loss is high. We used single gateway which located in different floor then sensors location. It has impact on RSSI(Received Signal Strength Indicator) value and SNR (Signal-to-noise ratio) value. It is necessary to evaluate signal quality evaluation from different location and distance from data receiver gateway with the LoRaWAN technology. In future, it is necessary to extend performance evaluation experiment. In that case, density of the gateway should consider as a primary metric for data evaluation performance. Also, maximum connected node capacity per gateway is another important factor for data handling capacity for performance evaluation of LoRaWAN based system.

Also, in our propose framework to detect head and mouth related personal scale behavior activities, we observed some misclassifications among different activities that require body movement and which don't require any kind of body movement. Also, head- and mouth-related activities demonstrate slightly better results than normal activities. Therefore, using our proposed framework, it will be easier for collecting head- and mouth-related activity data. Any machine-learning model can be applied to this data to detect head- and mouth-related activities. It will create a new scope in research and applications using earable device. Using our proposed system, we also collected audio data. In the future, researchers can use audio data to reduce misclassification related to *'Speaking'* and *'Eating'* activities, which have some similar patterns due to the muscle's movement in the face. Also increase the number of activities including complicated activities will be impactful for real nursing care/healthcare centers.

Another key future research direction for evaluating and conduct the sample size determination experiments in various domains which may reveal insights for other domain requirement. It can provide perspectives on how various domain-specific factors influence the design of their respective data collection studies. Such work could lead to the development of benchmark datasets using IRL across domains for modeling human behavior.

Bibliography

[1] M.A.R Ahad. 2020. Vision and sensor based human activity recognition: challenges ahead. In *Advancements in Instrumentation and Control in Applied System Applications*. Srijan Bhattacharya, editor. Hershey, PA: IGI Global, Oxford. Chapter 2, 17–35.

[2] Md. Atiqur Rahman Ahad. 2011. *Computer Vision and Action Recognition - A Guide for Image Processing and Computer Vision Community for Action Understanding*. *Atlantis Ambient and Pervasive Intelligence*. Volume 5. Atlantis Press. ISBN: 978-94-91216-19-0. DOI: 10.2991/978-94-91216-20-6. https://doi.org/10.2991/978-94-91216-20-6.

[3] Md. Atiqur Rahman Ahad. 2013. Motion history images for action recognition and understanding. In *Springer Briefs in Computer Science.* Springer-Verlag London.

[4] Md. Atiqur Rahman Ahad, Anindya Das Antar, and Masud Ahmed. 2021. *IoT Sensor-Based Activity Recognition - Human Activity Recognition*. *Intelligent Systems Reference Library*. Volume 173. Springer. ISBN: 978-3-030-51378-8. DOI: 10.1007/978-3-030-51379-5. https://doi.org/10.1007/978-3-030-51379-5.

[5] Md. Atiqur Rahman Ahad, Anindya Das Antar, and Masud Ahmed. 2020. Sensor-based human activity recognition: challenges ahead, (July 2020), 175–189. DOI: https://doi.org/10.1007/978-3-030-51379-5_10.

[6] Md. Atiqur Rahman Ahad, Anindya Das Antar, and Omar Shahid. 2019. Vision-based action understanding for assistive healthcare: a short review. In (June 2019).

[7] M. Ahmed, A. Das Antar, and M. A. R. 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)*, 140–145. DOI: 10.1109/ICIEV.2019.8858582.

[8] 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. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable*

Computers (UbiComp-ISWC '20). Association for Computing Machinery, Virtual Event, Mexico, 378–383. ISBN: 9781450380768. DOI: 10.1145/3410530.3414611. https://doi.org/10.1145/3410530.3414611.

[9] 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. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (UbiComp-ISWC '20). Association for Computing Machinery, Virtual Event, Mexico, 378–383. ISBN: 9781450380768. DOI: 10.1145/3410530.3414611. https://doi.org/10.1145/3410530.3414611.

[10] C. Alippi, G. Boracchi, and M. Roveri. 2012. On-line reconstruction of missing data in sensor/actuator networks by exploiting temporal and spatial redundancy. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 1–8. DOI: 10.1109/IJCNN.2012.6252689.

[11] Md Atiqur Rahman Ahad, Anindya Das Antar, and Omar Shahid. 2019. Vision-based action understanding for assistive healthcare: a short review. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[12] Abeer Badawi, Ahmad Al-Kabbany, and Heba Shaban. 2020. Sensor type, axis, and position-based fusion and feature selection for multimodal human daily activity recognition in wearable body sensor networks. *Journal of Healthcare Engineering*, 2020, (June 2020). DOI: 10.1155/2020/7914649.

[13] A. M. Baharudin and W. Yan. 2016. Long-range wireless sensor networks for geo-location tracking: design and evaluation. In *2016 International Electronics Symposium (IES)*, 76–80. DOI: 10.1109/ELECSYM.2016.7860979.

[14] Ahmad Muzaffar Baharudin and Wanglin Yan. 2017. Long-range wireless sensor networks for geo-location tracking: design and evaluation. English. In *Proceedings - 2016 International Electronics Symposium, IES 2016*. 18th International Electronics Symposium, IES 2016 ; Conference date: 29-09-2016 Through 30-09-2016. Institute of Electrical and Electronics Engineers Inc., (February 2017), 76–80. DOI: 10.1109/ELECSYM.2016.7860979.

[15] G. Bahle, A. Gruenerbl, P. Lukowicz, E. Bignotti, M. Zeni, and F. Giunchiglia. 2014. Recognizing hospital care activities with a coat pocket worn smartphone. In *6th International Conference on Mobile Computing, Applications and Services*, 175–181. DOI: 10.4108/icst.mobicase.2014.257777.

[16] M. Balazinska, A. Deshpande, M. J. Franklin, P. B. Gibbons, J. Gray, M. Hansen, M. Liebhold, S. Nath, A. Szalay, and V. Tao. 2007. Data management in the worldwide sensor web. *IEEE Pervasive Computing*, 6, 2, 30–40. DOI: 10.1109/MPRV.2007.27.

[17] D. Bankov, E. Khorov, and A. Lyakhov. 2017. Mathematical model of lorawan channel access. In *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 1–3. DOI: 10.1109/WoWMoM.2017.7974300.

[18] D. Bankov, E. Khorov, and A. Lyakhov. 2016. On the limits of lorawan channel access. In *2016 International Conference on Engineering and Telecommunication (EnT)*, 10–14. DOI: 10.1109/EnT.2016.011.

[19] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K. Dey. 2016. Modeling and understanding human routine behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). Association for Computing Machinery, San Jose, California, USA, 248–260. ISBN: 9781450333627. DOI: 10.1145/2858036.2858557. https://doi.org/10.1145/2858036.2858557.

[20] Nikola Banovic, Anqi Wang, Yanfeng Jin, Christie Chang, Julian Ramos, Anind Dey, and Jennifer Mankoff. 2017. Leveraging human routine models to detect and generate human behaviors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, Denver, Colorado, USA, 6683–6694. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025571. http://doi.acm.org/10.1145/3025453.3025571.

[21] Ling Bao and Stephen Intille. 2004. Activity recognition from user-annotated acceleration data. In volume 3001. (April 2004), 1–17. ISBN: 978-3-540-21835-7. DOI: 10.1007/978-3-540-24646-6_1.

[22] Abdelkareem Bedri, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, Thad Starner, and Gregory Abowd. 2017. Earbit: using wearable sensors to detect eating episodes in unconstrained environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1, 3, (September 2017). DOI: 10.1145/3130902. https://doi.org/10.1145/3130902.

[23] James O. Berger. 1985. *Statistical Decision Theory and Bayesian Analysis*. *Springer Series in Statistics*. Springer New York, New York, NY. ISBN: 978-1-4419-3074-3. DOI: 10.1007/978-1-4757-4286-2.

[24] Jose M. Bernardo and Adrian F. M. Smith. 2000. *Bayesian Theory*. John Wiley & Sons, New York, NY.

[25] Christopher Bishop. 2006. *Pattern Recognition and Machine Learning*. *Information Science and Statistics*. Springer-Verlag New York.

[26] U. Blanke, B. Schiele, M. Kreil, P. Lukowicz, B. Sick, and T. Gruber. 2010. All for one or one for all? combining heterogeneous features for activity spotting. In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 18–24. DOI: 10.1109/PERCOMW.2010.5470597.

[27] Leo Breiman. 2001. Random forests. *Mach. Learn.*, 45, 1, (October 2001), 5–32. ISSN: 0885-6125. DOI: 10.1023/A:1010933404324. https://doi.org/10.1023/A:1010933404324.

[28] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng, editors. 2011. *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC. ISBN: 978-1-4200-7941-8. DOI: 10.1201/b10905.

[29] Andreas Bulling, Ulf Blanke, and Bernt Schiele. 2014. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Comput. Surv.*, 46, 3, (January 2014), 33:1–33:33. ISSN: 0360-0300. DOI: 10.1145/2499621. http://doi.acm.org/10.1145/2499621.

[30] Pierluigi Casale, Oriol Pujol, and Petia Radeva. 2012. Personalization and user verification in wearable systems using biometric walking patterns. *Personal and Ubiquitous Computing - PUC*, 16, (June 2012), 1–18. DOI: 10.1007/s00779-011-0415-z.

[31] Marco Centenaro, Lorenzo Vangelista, Andrea Zanella, and Michele Zorzi. 2015. Long-range communications in unlicensed bands: the rising stars in the iot and smart city scenarios. *IEEE Wireless Communications*, 23, (October 2015). DOI: 10.1109/MWC.2016.7721743.

[32] Kathryn Chaloner and Isabella Verdinelli. 1995. Bayesian Experimental Design: A Review. *Statistical Science*, 10, 3, 273–304. ISSN: 0883-4237. DOI: 10.1214/ss/1177009939.

[33] Larry Chan, Vedant Das Swain, Christina Kelley, Kaya de Barbaro, Gregory D. Abowd, and Lauren Wilcox. 2018. Students' experiences with ecological momentary assessment tools to report on emotional well-being. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2, 1, (March 2018), 3:1–3:20. ISSN: 2474-9567. DOI: 10.1145/3191735. http://doi.acm.org/10.1145/3191735.

[34] Yung-Ju Chang, Gaurav Paruthi, and Mark W. Newman. 2015. A field study comparing approaches to collecting annotated activity data in real-world settings. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). ACM, Osaka, Japan, 671–682. ISBN: 978-1-4503-3574-4. DOI: 10.1145/2750858.2807524. http://doi.acm.org/10.1145/2750858.2807524.

[35] Basit Chaudhry, Jerome Wang, Shinyi Wu, Margaret Maglione, Walter Mojica, Elizabeth Roth, Sally C Morton, and Paul G Shekelle. 2006. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Annals of internal medicine*, 144, 10, 742–752.

[36] Girija Chetty, Matthew White, and Farnaz Akther. 2015. Smart phone based data mining for human activity recognition. *Procedia Computer Science*, 46, 1181 –1187. Proceedings of the International Conference on Information and Communication Technologies, ICICT 2014, 3-5 December 2014 at Bolgatty

Palace Island Resort, Kochi, India. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2015.01.031. http://www.sciencedirect.com/science/article/pii/S1877050915000320.

[37] Geeta Chhabra, Research Scholar, Vasudha Vashisht, and Jayanthi Ranjan. 2019. A classifier ensemble machine learning approach to improve efficiency for missing value imputation, (April 2019).

[38] T. Choudhury, G. Borriello, S. Consolvo, D. Haehnel, B. Harrison, B. Hemingway, J. Hightower, P. P. Klasnja, K. Koscher, A. LaMarca, J. A. Landay, L. LeGrand, J. Lester, A. Rahimi, A. Rea, and D. Wyatt. 2008. The mobile sensing platform: an embedded activity recognition system. *IEEE Pervasive Computing*, 7, 2, 32–41. DOI: 10.1109/MPRV.2008.39.

[39] Diane Cook and Maureen Schmitter-Edgecombe. 2009. Assessing the quality of activities in a smart environment. *Methods of information in medicine*, 48, (June 2009), 480–5. DOI: 10.3414/ME0592.

[40] Thomas A. Cover and Joy A. Thomas. 2006. *Elements of Information Theory*. (2nd edition). John Wiley & Sons, Hoboken, NJ.

[41] Zhicheng Cui, Wenlin Chen, and Yujie He. 2015. Optimal action extraction for random forests and boosted trees. In (August 2015), 179–188. DOI: 10.1145/2783258.2783281.

[42] A. Das Antar, M. Ahmed, and M. A. R. Ahad. 2019. Challenges in sensor-based human activity recognition and a comparative analysis of benchmark datasets: a review. In *2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, 134–139. DOI: 10.1109/ICIEV.2019.8858508.

[43] A. G. D'Sa and B. G. Prasad. 2019. A survey on vision based activity recognition, its applications and challenges. In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, 1–8. DOI: 10.1109/ICACCP.2019.8882896.

[44] Ahmed. S. Elselini, Hamed. R. Eleribi, Mahdi Sanaani, and Abdelhakim Alwerfalli. 2020. A performance study of an iot system using lora access network technology. In *Proceedings of the 6th International Conference on Engineering MIS 2020* (ICEMIS'20). Association for Computing Machinery, Almaty, Kazakhstan. ISBN: 9781450377362. DOI: 10.1145/3410352.3410807. https://doi.org/10.1145/3410352.3410807.

[45] Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio. 2010. Why does unsupervised pre-training help deep learning? *J. Mach. Learn. Res.*, 11, (March 2010), 625–660. ISSN: 1532-4435.

[46] Macarena Espinilla, J. Medina, and Chris Nugent. 2018. Ucami cup. analyzing the uja human activity recognition dataset of activities of daily living. *Proceedings*, 2, (October 2018), 1267. DOI: 10.3390/proceedings2191267.

[47] I. Ezzine and L. Benhlima. 2018. A study of handling missing data methods for big data. In *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, 498–501. DOI: 10.1109/CIST.2018.8596389.

[48] Md. Ahasan Atick Faisal, Md. Sadman Siraj, Md. Tahmeed Abdullah, Omar Shahid, Farhan Fuad Abir, and M. A. R. Ahad. 2020. A pragmatic signal processing approach for nurse care activity recognition using classical machine learning. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (UbiComp-ISWC '20). Association for Computing Machinery, Virtual Event, Mexico, 396–401. ISBN: 9781450380768. DOI: 10.1145/3410530.3414337. https://doi.org/10.1145/3410530.3414337.

[49] J. Favela. 2013. Behavior-aware computing: applications and challenges. *IEEE Pervasive Computing*, 12, 3, 14–17. DOI: 10.1109/MPRV.2013.44.

[50] A. K. Fletcher, S. Rangan, and V. K. Goyal. 2004. Estimation from lossy sensor data: jump linear modeling and kalman filtering. In *Third International Symposium on Information Processing in Sensor Networks, 2004. IPSN 2004*, 251–258. DOI: 10.1109/IPSN.2004.239018.

[51] Pedro J. García-Laencina, José-Luis Sancho-Gómez, and Aníbal R. Figueiras-Vidal. 2010. Pattern classification with missing data: a review. *Neural Comput. Appl.*, 19, 2, (March 2010), 263–282. ISSN: 0941-0643. DOI: 10.1007/s00521-009-0295-6. https://doi.org/10.1007/s00521-009-0295-6.

[52] Roger Ghanem, David Higdon, and Houman Owhadi, editors. 2017. *Handbook of uncertainty quantification*. Springer International Publishing, Cham. ISBN: 978-3-319-12384-4. DOI: 10.1007/978-3-319-12385-1. arXiv: 1507.00398.

[53] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, New York, NY.

[54] W. Guibene, J. Nowack, N. Chalikias, K. Fitzgibbon, M. Kelly, and D. Prendergast. 2017. Evaluation of lpwan technologies for smart cities: river monitoring use-case. In *2017 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 1–5. DOI: 10.1109/WCNCW.2017.7919089.

[55] Wael Guibene, Johannes Nowack, Nikolaos Chalikias, Kevin Fitzgibbon, Mark Kelly, and David Prendergast. 2017. Evaluation of lpwan technologies for smart cities: river monitoring use-case. In (March 2017), 1–5. DOI: 10.1109/WCNCW.2017.7919089.

[56] Z. Hajihashemi and M. Popescu. 2013. Detection of abnormal sensor patterns in eldercare. In *2013 E-Health and Bioengineering Conference (EHB)*, 1–4. DOI: 10.1109/EHB.2013.6707389.

[57] Mihail Halatchev and Le Gruenwald. 2005. Estimating missing values in related sensor data streams. In (January 2005), 83–94.

[58] Tian Hao, Chongguang Bi, Guoliang Xing, Roxane Chan, and Linlin Tu. 2017. Mindfulwatch: a smartwatch-based system for real-time respiration monitoring during meditation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1, 3, (September 2017). DOI: 10.1145/3130922. https://doi.org/10.1145/3130922.

[59] Stephen L Hauser and Jorge R Oksenberg. 2006. The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration. *Neuron*, 52, 1, 61–76.

[60] Ann L. Hendrich, Marilyn Chow, Boguslaw A Skierczynski, and Zhenqiang Lu. 2008. A 36-hospital time and motion study: how do medical-surgical nurses spend their time? *The Permanente journal*, 12 3, 25–34.

[61] Hongguo Zhang and Liang Yang. 2014. An improved algorithm for missing data in wireless sensor networks. In *International Conference on Software Intelligence Technologies and Applications International Conference on Frontiers of Internet of Things 2014*, 346–350. DOI: 10.1049/cp.2014.1587.

[62] T. Hossain, H. Goto, M. A. Rahman Ahad, and S. Inoue. 2018. A study on sensor-based activity recognition having missing data. In *2018 Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, 556–561. DOI: 10.1109/ICIEV.2018.8641012.

[63] Tahera Hossain, Md Atiqur Rahman Ahad, and Sozo Inoue. 2020. A method for sensor-based activity recognition in missing data scenario. *Sensors (Basel, Switzerland)*, 20, 14. ISSN: 1424-8220. DOI: 10.3390/s20143811. https://europepmc.org/articles/PMC7412080.

[64] Tahera Hossain, Md Atiqur Rahman Ahad, Tahia Tazin, and Sozo Inoue. 2018. Activity recognition by using lorawan sensor. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (UbiComp '18). Association for Computing Machinery, Singapore, Singapore, 58–61. ISBN: 9781450359665. DOI: 10.1145/3267305.3267652. https://doi.org/10.1145/3267305.3267652.

[65] Tahera Hossain, Yusuke Doi, Tahia Tazin, Md Atiqur Rahman Ahad, and Sozo Inoue. 2018. Study of lorawan technology for activity recognition. In

Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp '18). Association for Computing Machinery, Singapore, Singapore, 1449–1453. ISBN: 9781450359665. DOI: 10.1145/3267305.3267510. https://doi.org/10.1145/3267305.3267510.

[66] Tahera Hossain, Md Shafiqul Islam, Md Atiqur Rahman Ahad, and Sozo Inoue. 2019. Human activity recognition using earable device. 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* (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, London, United Kingdom, 81–84. ISBN: 9781450368698. DOI: 10.1145/3341162.3343822. https://doi.org/10.1145/3341162.3343822.

[67] Rima Houari, Ahcène Bounceur, Abdelkamel Tari, and M Kecha. 2014. Handling missing data problems with sampling methods. In (June 2014), 99–104. DOI: 10.1109/INDS.2014.25.

[68] Xun Huan and Youssef M. Marzouk. 2013. Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232, 1, 288–317. ISSN: 00219991. DOI: 10.1016/j.jcp.2012.08.013.

[69] S. Inoue, T. Mairittha, N. Mairittha, and T. Hossain. 2019. Integrating activity recognition and nursing care records: the system, experiment, and the dataset. In *2019 Joint 8th International Conference on Informatics, Electronics Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision Pattern Recognition (icIVPR)*, 73–78. DOI: 10.1109/ICIEV.2019.8858584.

[70] Sozo INOUE, Tatsuya ISODA, Mako SHIROUZU, Yasuhiko SUGIYAMA, Yasunobu NOHARA, and Naoki NAKASHIMA. 2016. Predicting daily nursing load from nurses' activity logs and patients' medical records. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (UbiComp '16). Association for Computing Machinery, Heidelberg, Germany, 89–92. ISBN: 9781450344623. DOI: 10.1145/2968219.2971454. https://doi.org/10.1145/2968219.2971454.

[71] Sozo Inoue, Paula Lago, Tahera Hossain, Tittaya Mairittha, and Nattaya Mairittha. 2019. Integrating activity recognition and nursing care records: the system, deployment, and a verification study. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3, 3, (September 2019). DOI: 10.1145/3351244. https://doi.org/10.1145/3351244.

[72] Sozo Inoue and Xincheng Pan. 2016. Supervised and unsupervised transfer learning for activity recognition from simple in-home sensors. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (MobiQuitous 2016). ACM, Hiroshima, Japan, 20–27. ISBN: 978-1-4503-4750-1. DOI: 10.1145/2994374.2994400. http://doi.acm.org/10.1145/2994374.2994400.

[73] Sozo Inoue and Xincheng Pan. 2016. Supervised and unsupervised transfer learning for activity recognition from simple in-home sensors. In (November 2016), 20–27. DOI: 10.1145/2994374.2994400.

[74] 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* (UbiComp '15). ACM, Osaka, Japan, 1269–1280. ISBN: 978-1-4503-3574-4. DOI: 10.1145/2750858.2807533. http://doi.acm.org/10.1145/2750858.2807533.

[75] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37* (ICML'15). JMLR.org, Lille, France, 448–456.

[76] S.M.; Islam M.Z.; Serikawa S.; Ahad M.A.R Islam Z.Z.; Tazwar. 2017. Automatic fall detection system of unsupervised elderly people using smartphone. In *In Proceedings of the 5th IIAE International Conference on Intelligent Systems and Image Processing*. ICISIP2017, Hawaii, USA.

[77] Patimakorn Jantaraprim, Pornchai Phukpattaranont, Chusak Limsakul, and Booncharoen Wongkittisuksa. 2009. Evaluation of fall detection for the elderly on a variety of subject groups. In *Proceedings of the 3rd International Convention on Rehabilitation Engineering Assistive Technology* (i-CREATe '09). Association for Computing Machinery, Singapore. ISBN: 9781605587929. DOI: 10.1145/1592700.1592713. https://doi.org/10.1145/1592700.1592713.

[78] N. Jaques, S. Taylor, A. Sano, and R. Picard. 2017. Multimodal autoencoder: a deep learning approach to filling in missing sensor data and enabling better mood prediction. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, 202–208. DOI: 10.1109/ACII.2017.8273601.

[79] Imen Jegham, Anouar Ben Khalifa, Ihsen Alouani, and Mohamed Mahjoub. 2020. Vision-based human action recognition: an overview and real world challenges. *Digital Investigation*, 32, (March 2020), 200901. DOI: 10.1016/j.fsidi.2019.200901.

[80] Jeya Vikranth Jeyakumar, Eun Sun Lee, Zhengxu Xia, Sandeep Singh Sandha, Nathan Tausik, and Mani Srivastava. 2018. Deep convolutional bidirectional lstm based transportation mode recognition. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (UbiComp '18). Association for Computing Machinery, Singapore, Singapore, 1606–1615. ISBN: 9781450359665. DOI: 10.1145/3267305.3267529. https://doi.org/10.1145/3267305.3267529.

[81] 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* (MM '15). Association for Computing Machinery, Brisbane, Australia, 1307–1310. ISBN: 9781450334594. DOI: 10.1145/2733373.2806333. https://doi.org/10.1145/2733373.2806333.

[82] Miodrag Lovric, editor. 2011. *Kullback-leibler divergence*. *International Encyclopedia of Statistical Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 720–722. ISBN: 978-3-642-04898-2. DOI: 10.1007/978-3-642-04898-2_327.

[83] 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* (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, London, United Kingdom, 736–740. ISBN: 9781450368698. DOI: 10.1145/3341162.3344859. https://doi.org/10.1145/3341162.3344859.

[84] Ashish Kapoor and Eric Horvitz. 2008. Experience sampling for building predictive user models: a comparative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08). ACM, Florence, Italy, 657–666. ISBN: 978-1-60558-011-1. DOI: 10.1145/1357054.1357159. http://doi.acm.org/10.1145/1357054.1357159.

[85] Nobuo Kawaguchi, Nobuhiro Ogawa, Yohei Iwasaki, Katsuhiko Kaji, Tsutomu Terada, Kazuya Murao, Sozo Inoue, Yoshihiro Kawahara, Yasuyuki Sumi, and Nobuhiko Nishio. 2011. Hasc challenge: gathering large scale human activity corpus for the real-world activity understandings. In *Proceedings of the 2nd Augmented Human International Conference* (AH '11). Association for Computing Machinery, Tokyo, Japan. ISBN: 9781450304269. DOI: 10.1145/1959826.1959853. https://doi.org/10.1145/1959826.1959853.

[86] Fahim Kawsar, Chulhong Min, Akhil Mathur, and Alessandro Montanari. 2018. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, 17, (July 2018), 83–89. DOI: 10.1109/MPRV.2018.03367740.

[87] Marc. C. Kennedy and Anthony O'Hagan. 2001. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 3, 425–464. DOI: 10.1111/1467-9868.00294. http://www3.interscience.wiley.com/journal/118994758/abstract.

[88] L. Kong, M. Xia, X. Liu, M. Wu, and X. Liu. 2013. Data loss and reconstruction in sensor networks. In *2013 Proceedings IEEE INFOCOM*, 1654–1662. DOI: 10.1109/INFCOM.2013.6566962.

[89] Anna L Kratz, Tiffany J Braley, Emily Foxen-Craft, Eric Scott, John F Murphy III, and Susan L Murphy. 2017. How do pain, fatigue, depressive, and cognitive symptoms relate to well-being and social and physical functioning in the daily lives of individuals with multiple sclerosis? *Archives of physical medicine and rehabilitation*, 98, 11, 2160–2166.

[90] Anna L Kratz, Susan L Murphy, and Tiffany J Braley. 2017. Ecological momentary assessment of pain, fatigue, depressive, and cognitive symptoms reveals significant daily variability in multiple sclerosis. *Archives of physical medicine and rehabilitation*, 98, 11, 2142–2150.

[91] Anna L Kratz, Susan L Murphy, and Tiffany J Braley. 2017. Pain, fatigue, and cognitive symptoms are temporally associated within but not across days in multiple sclerosis. *Archives of physical medicine and rehabilitation*, 98, 11, 2151–2159.

[92] Narayanan C. Krishnan and Diane J. Cook. 2014. Activity recognition on streaming sensor data. *Pervasive and Mobile Computing*, 10, 138 –154. ISSN: 1574-1192. DOI: https://doi.org/10.1016/j.pmcj.2012.07.003. http://www.sciencedirect.com/science/article/pii/S1574119212000776.

[93] Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics*, 22, 1, 79–86.

[94] Paula Lago, Sayeda Shamma Alia, Shingo Takeda, Tittaya Mairittha, Nattaya Mairittha, Farina Faiz, Yusuke Nishimura, Kohei Adachi, Tsuyoshi Okita, François Charpillet, and Sozo Inoue. 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* (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, London, United Kingdom, 746–751. ISBN: 9781450368698. DOI: 10.1145/3341162.3345577. https://doi.org/10.1145/3341162.3345577.

[95] Paula Lago, Claudia Jiménez-Guarín, and Claudia Roncancio. 2015. Contextualized behavior patterns for ambient assisted living. 9277, (January 2015), 132–145. DOI: 10.1007/978-3-319-24195-1_10.

[96] Paula Lago, Fréderic Lang, Claudia Roncancio, Claudia Jiménez-Guarín, Radu Mateescu, and Nicolas Bonnefond. 2017. The contextact@a4h real-life dataset of daily-living activities. In *Modeling and Using Context*. Patrick Brézillon, Roy Turner, and Carlo Penco, editors. Springer International Publishing, Cham, 175–188. ISBN: 978-3-319-57837-8.

[97] Paula Lago, Frédéric Lang, Claudia Roncancio, Claudia Jiménez-Guarín, Radu Mateescu, and Nicolas Bonnefond. 2017. The contextact@a4h real-life dataset of daily-living activities. In (May 2017), 175–188. ISBN: 978-3-319-57836-1. DOI: 10.1007/978-3-319-57837-8_14.

[98] Paula Lago, Moe Matsuki, and Sozo Inoue. 2020. Achieving single-sensor complex activity recognition from multi-sensor training data. *CoRR*, abs/2002.11284. arXiv: 2002.11284. https://arxiv.org/abs/2002.11284.

[99] P. Lameski, A. Dimitrievski, E. Zdravevski, V. Trajkovik, and S. Koceski. 2019. Challenges in data collection in real-world environments for activity recognition. In *IEEE EUROCON 2019 -18th International Conference on Smart Technologies*, 1–5. DOI: 10.1109/EUROCON.2019.8861964.

[100] O. D. Lara and M. A. Labrador. 2013. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys Tutorials*, 15, 3, 1192–1209. ISSN: 1553-877X. DOI: 10.1109/SURV.2012.110112.00192.

[101] Oscar Lara and Miguel Labrador. 2013. A survey on human activity recognition using wearable sensors. *Communications Surveys Tutorials, IEEE*, 15, (January 2013), 1192–1209. DOI: 10.1109/SURV.2012.110112.00192.

[102] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. 2006. A practical approach to recognizing physical activities. In *Pervasive Computing*. Kenneth P. Fishkin, Bernt Schiele, Paddy Nixon, and Aaron Quigley, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–16. ISBN: 978-3-540-33895-6.

[103] Dennis V. Lindley. 1997. The choice of sample size. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 46, 2, 129–138. ISSN: 00390526, 14679884. http://www.jstor.org/stable/2988516.

[104] Nattaya Mairittha, Tittaya Mairittha, and Sozo Inoue. 2018. A mobile app for nursing activity recognition. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (UbiComp '18). Association for Computing Machinery, Singapore, Singapore, 400–403. ISBN: 9781450359665. DOI: 10.1145/3267305.3267633. https://doi.org/10.1145/3267305.3267633.

[105] Scott E. Maxwell, Ken Kelley, and Joseph R. Rausch. 2008. Sample size planning for statistical power and accuracy in parameter estimation. *Annual Review of Psychology*, 59, 1, 537–563. PMID: 17937603. DOI: 10.1146/annurev.psych.59.103006.093735. eprint: https://doi.org/10.1146/annurev.psych.59.103006.093735. https://doi.org/10.1146/annurev.psych.59.103006.093735.

[106] SORACOM Air for LoRaWAN. 2018. Soracom. (2018).

[107] World Health Organization. 2019. Ageing. (2019).

[108] Daniel Roggen Md Atiqur Rahman Ahad Sozo Inoue and Kaori Fujinami. 2021. *Activity and Behavior Computing*. *2190*. Springer.

[109] Masud Ahmed Md. Atiqur Rahman Ahad Anindya Das Antar. 2020. In *IoT Sensor-Based Activity Recognition*. Springer, 189.

[110] Paula Lago Md Atiqur Rahman Ahad and Sozo Inoue. 2021. *Human Activity Recognition Challenge*. *2190*. Springer.

[111] Kais Mekki, Eddy Bajic, Frederic Chaxel, and Fernand Meyer. 2019. A comparative study of lpwan technologies for large-scale iot deployment. *ICT Express*, 5, 1, 1 –7. ISSN: 2405-9595. DOI: https://doi.org/10.1016/j.icte.2017.12.005.

[112] Roderick Melnik. 2015. Universality of mathematical models in understanding nature, society, and man-made world, 1–16. DOI: 10.1002/9781118853887.ch1.

[113] M. Moayedi, Y. K. Foo, and Y. C. Soh. 2010. Adaptive kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements. *IEEE Transactions on Signal Processing*, 58, 3, 1577–1588. DOI: 10.1109/TSP.2009.2037853.

[114] Peter Müller. 2005. Simulation Based Optimal Design. *Handbook of Statistics*, 25, 509–518. ISSN: 01697161. DOI: 10.1016/S0169-7161(05)25017-4. http://linkinghub.elsevier.com/retrieve/pii/S0169716105250174.

[115] F. Naya, R. Ohmura, F. Takayanagi, H. Noma, and K. Kogure. 2006. Workers' routine activity recognition using body movements and location information. In *2006 10th IEEE International Symposium on Wearable Computers*, 105–108. DOI: 10.1109/ISWC.2006.286351.

[116] P. Neumann, J. Montavont, and T. Noël. 2016. Indoor deployment of low-power wide area networks (lpwan): a lorawan case study. In *2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 1–8. DOI: 10.1109/WiMOB.2016.7763213.

[117] Andrew Y. Ng and Stuart J. Russell. 2000. Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning* (ICML '00). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 663–670. ISBN: 1558607072.

[118] Nhan Nguyen, Duong Bui, Phuc Truong, and Gu-Min Jeong. 2018. Position-based feature selection for body sensors regarding daily living activity recognition. *Journal of Sensors*, 2018, (September 2018), 1–13. DOI: 10.1155/2018/9762098.

[119] Henry Nweke, Teh Wah, Mohammed Al-Garadi, and Uzoma Alo. 2018. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: state of the art and research challenges. *Expert Systems with Applications*, 105, (April 2018). DOI: 10.1016/j.eswa.2018.03.056.

[120] Maria Pampaka, Graeme Hutcheson, and Julian Williams. 2016. Handling missing data: analysis of a challenging data set using multiple imputation. *International Journal of Research & Method in Education*, 39, 1, 19–37. DOI: 10.1080/1743727X.2014.979146.

[121] Giovanni Parmigiani and Lurdes Y. T. Inoue. 2009. *Decision Theory: Principles and Approaches*. John Wiley & Sons, Inc., West Sussex, United Kingdom, 372. ISBN: 9780471496571. http://books.google.com/books?id=mnjGCYqWj7EC{\&}pgis=1.

[122] J. Petajajarvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo. 2015. On the coverage of lpwans: range evaluation and channel attenuation model for lora technology. In *2015 14th International Conference on ITS Telecommunications (ITST)*, 55–59. DOI: 10.1109/ITST.2015.7377400.

[123] Samira Pouyanfar, Saad Sadiq, Yilin Yan, Haiman Tian, Yudong Tao, Maria Presa Reyes, Mei-Ling Shyu, Shu-Ching Chen, and S. S. Iyengar. 2018. A survey on deep learning: algorithms, techniques, and applications. *ACM Comput. Surv.*, 51, 5, (September 2018). ISSN: 0360-0300. DOI: 10.1145/3234150. https://doi.org/10.1145/3234150.

[124] O. M. Prabowo, K. Mutijarsa, and S. H. Supangkat. 2016. Missing data handling using machine learning for human activity recognition on mobile device. In *2016 International Conference on ICT For Smart Society (ICISS)*, 59–62. DOI: 10.1109/ICTSS.2016.7792849.

[125] Martin L. Puterman. 2014. Markov decision processes: discrete stochastic dynamic programming, 684 Pages.

[126] Zhixin Qi, Hongzhi Wang, Jianzhong Li, and Hong Gao. 2018. Impacts of dirty data: and experimental evaluation, (March 2018).

[127] Deepak Ramachandran and Eyal Amir. 2007. Bayesian inverse reinforcement learning. In *IJCAI*. Volume 7, 2586–2591.

[128] Marilyn Rantz, Marjorie Skubic, Greg Alexander, Myra Aud, Bonnie Wakefield, Colleen Galambos, Richelle Koopman, and Steven Miller. 2010. Improving nurse care coordination with technology. *Computers, Informatics, Nursing*, 28, 6, 325–332. ISSN: 1538-2931. DOI: 10.1097/NCN.0b013e3181f69838.

[129] Marilyn J Rantz, Marjorie Skubic, Steven J Miller, Colleen Galambos, Greg Alexander, James Keller, and Mihail Popescu. 2013. Sensor technology to support aging in place. *Journal of the American Medical Directors Association*, 14, 6, 386–391.

[130] Christian P. Robert and George Casella. 2004. *Monte Carlo Statistical Methods*. Springer New York, New York, NY. DOI: 10.1007/978-1-4757-4145-2.

[131] D. Roggen, K. Forster, A. Calatroni, T. Holleczek, Y. Fang, G. Troster, A. Ferscha, C. Holzmann, A. Riener, P. Lukowicz, G. Pirkl, D. Bannach, K. Kunze, R. Chavarriaga, and J. d. R. Millan. 2009. Opportunity: towards opportunistic activity and context recognition systems. In *2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks Workshops*, 1–6. DOI: 10.1109/WOWMOM.2009.5282442.

[132] Daniel Roggen, Alberto Calatroni, Kilian Förster, Gerhard Tröster, Paul Lukowicz, David Bannach, Alois Ferscha, Marc Kurz, Gerold Hölzl, Hesam Sagha, Hamidreza Bayati, José del R. Millán, and Ricardo Chavarriaga. 2011. Activity recognition in opportunistic sensor environments. *Procedia Computer Science*, 7, 173 –174. Proceedings of the 2nd European Future Technologies Conference and Exhibition 2011 (FET 11). ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2011.09.003. http://www.sciencedirect.com/science/article/pii/S1877050911005631.

[133] Aaqib Saeed, Tanir Ozcelebi, and Johan Lukkien. 2018. Synthesizing and reconstructing missing sensory modalities in behavioral context recognition. *Sensors*, 18, 9, 2967. ISSN: 1424-8220. DOI: 10.3390/s18092967. http://dx.doi.org/10.3390/s18092967.

[134] Hesam Sagha, Ricardo Chavarriaga, and Jose del R. Millan. 2010. A probabilistic approach to handle missing data for multi-sensory activity recognition. In (January 2010).

[135] Swapnil Sayan Saha, Shafizur Rahman, Zarin Rezwana Ridita Haque, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2019. Position independent activity recognition using shallow neural architecture and empirical modeling. 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* (UbiComp/ISWC '19 Adjunct). Association for Computing Machinery, London, United Kingdom, 808–813. ISBN: 9781450368698. DOI: 10.1145/3341162.3345572. https://doi.org/10.1145/3341162.3345572.

[136] Swapnil Sayan Saha, Shafizur Rahman, Miftahul Jannat Rasna, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2018. Supervised and neural classifiers for locomotion analysis. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (UbiComp '18). Association for Computing Machinery, Singapore, Singapore, 1563–1570. ISBN: 9781450359665. DOI: 10.1145/3267305.3267524. https://doi.org/10.1145/3267305.3267524.

[137] John M Salsman, David Victorson, Seung W Choi, Amy H Peterman, Allen W Heinemann, Cindy Nowinski, and David Cella. 2013. Development and validation of the positive affect and well-being scale for the neurology quality of life (neuro-qol) measurement system. *Quality of Life Research*, 22, 9, 2569–2580.

[138] Christian Schüldt, Ivan Laptev, and Barbara Caputo. 2004. Recognizing human actions: a local svm approach. In volume 3. (September 2004), 32 –36 Vol.3. ISBN: 0-7695-2128-2. DOI: 10.1109/ICPR.2004.1334462.

[139] M. R. Seye, B. Ngom, B. Gueye, and M. Diallo. 2018. A study of lora coverage: range evaluation and channel attenuation model. In *2018 1st International Conference on Smart Cities and Communities (SCCIC)*, 1–4. DOI: 10.1109/SCCIC.2018.8584548.

[140] L. Shao, F. Zhu, and X. Li. 2015. Transfer learning for visual categorization: a survey. *IEEE Transactions on Neural Networks and Learning Systems*, 26, 5, 1019–1034.

[141] D. S. Sivia and J. Skilling. 2006. *Data Analysis: A Bayesian Tutorial*. (2nd edition). Oxford University Press, New York, NY, 246.

[142] Valeria Soto-Mendoza, José-Antonio García-Macías, Edgar Chávez, Anna Isabel Martínez-García, Jesus Favela, Patricia Serrano-Alvarado, and Maythé R. Zúñiga-Rojas. 2015. Design of a predictive scheduling system to improve assisted living services for elders. *ACM Transactions on Intelligent Systems and Technology*, 6, (June 2015). DOI: 10.1145/2736700.

[143] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15, 1, (January 2014), 1929–1958. ISSN: 1532-4435.

[144] M. Stikic, D. Larlus, S. Ebert, and B. Schiele. 2011. Weakly supervised recognition of daily life activities with wearable sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 12, 2521–2537. DOI: 10.1109/TPAMI.2011.36.

[145] Maja Stikic and Bernt Schiele. 2009. Activity recognition from sparsely labeled data using multi-instance learning. In *Location and Context Awareness*. Tanzeem Choudhury, Aaron Quigley, Thomas Strang, and Koji Suginuma, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 156–173. ISBN: 978-3-642-01721-6.

[146] Vamsi Talla, Mehrdad Hessar, Bryce Kellogg, Ali Najafi, Joshua R. Smith, and Shyamnath Gollakota. 2017. Lora backscatter: enabling the vision of ubiquitous connectivity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1, 3, (September 2017). DOI: 10.1145/3130970. https://doi.org/10.1145/3130970.

[147] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. 2004. Activity recognition in the home using simple and ubiquitous sensors. In *International conference on pervasive computing*. Springer, 158–175.

[148] Takamichi Toda, Sozo Inoue, and Naonori Ueda. 2016. Mobile activity recognition through training labels with inaccurate activity segments. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services* (MOBIQUITOUS 2016). Association for Computing Machinery, Hiroshima, Japan, 57–64. ISBN: 9781450347501. DOI: 10.1145/2994374.2994378. https://doi.org/10.1145/2994374.2994378.

[149] Yonatan Vaizman, Nadir Weibel, and Gert Lanckriet. 2018. Context recognition in-the-wild: unified model for multi-modal sensors and multi-label classification. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1, 4, (January 2018). DOI: 10.1145/3161192. https://doi.org/10.1145/3161192.

[150] 2011. *Human activity recognition from wireless sensor network data: benchmark and software*. Volume 4. (May 2011), 165–186. ISBN: 978-90-78677-42-0. DOI: 10.2991/978-94-91216-05-3_8.

[151] K. Van Laerhoven, D. Kilian, and B. Schiele. 2008. Using rhythm awareness in long-term activity recognition. In *2008 12th IEEE International Symposium on Wearable Computers*, 63–66. DOI: 10.1109/ISWC.2008.4911586.

[152] Robert-Andrei Voicu, Ciprian Dobre, Lidia Bajenaru, and Radu-Ioan Ciobanu. 2019. Human physical activity recognition using smartphone sensors. *Sensors*, 19, 3, 458.

[153] Udo Von Toussaint. 2011. Bayesian inference in physics. *Reviews of Modern Physics*, 83, 943–999. ISSN: 00346861. DOI: 10.1103/RevModPhys.83.943.

[154] A. Wang, G. Chen, J. Yang, S. Zhao, and C. Chang. 2016. A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sensors Journal*, 16, 11, 4566–4578. ISSN: 1530-437X. DOI: 10.1109/JSEN.2016.2545708.

[155] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. 2019. Deep learning for sensor-based activity recognition: a survey. *Pattern Recognition Letters*, 119, 3 –11. Deep Learning for Pattern Recognition. ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2018.02.010. http://www.sciencedirect.com/science/article/pii/S016786551830045X.

[156] N. Weber and A. Härmä. 2017. Unsupervised learning in human activity recognition : a first foray into clustering data gathered from wearable sensors. In.

[157] Karl Weiss, Taghi Khoshgoftaar, and DingDing Wang. 2016. A survey of transfer learning. *Journal of Big Data*, 3, (December 2016). DOI: 10.1186/s40537-016-0043-6.

[158] Falin Wu, Hengyang Zhao, Yan Zhao, and Haibo Zhong. 2015. Development of a wearable-sensor-based fall detection system. 2015, (January 2015). ISSN: 1687-6415. DOI: 10.1155/2015/576364. https://doi.org/10.1155/2015/576364.

[159] Zhixian Yan, Vigneshwaran Subbaraju, Dipanjan Chakraborty, Archan Misra, and Karl Aberer. 2012. Energy-efficient continuous activity recognition on mobile phones: an activity-adaptive approach. In (June 2012). DOI: 10.1109/ISWC.2012.23.

[160] J. Yin, Q. Yang, and J. J. Pan. 2008. Sensor-based abnormal human-activity detection. *IEEE Transactions on Knowledge and Data Engineering*, 20, 8, 1082–1090. DOI: 10.1109/TKDE.2007.1042.

[161] Yongshuai Shao, Zhe Chen, Fangfang Li, and Chong Fu. 2016. Reconstruction of big sensor data. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, 1–6. DOI: 10.1109/CompComm.2016.7924653.

[162] Yan Yu, Feng Han, Yuequan Bao, and Jinping Ou. 2015. A study on data loss compensation of wifi-based wireless sensor networks for structural health monitoring. *IEEE Sensors Journal*, 16, (January 2015), 1–1. DOI: 10.1109/JSEN.2015.2512846.

[163] YuanYuan Li and L. E. Parker. 2008. Classification with missing data in a wireless sensor network. In *IEEE SoutheastCon 2008*, 533–538. DOI: 10.1109/SECON.2008.4494352.

[164] Yunxiu Zeng, Kai Xu, Quanjun Yin, L. Qin, Yabing Zha, and William Yeoh. 2018. Inverse reinforcement learning based human behavior modeling for goal recognition in dynamic local network interdiction. In *AAAI Workshops*.

[165] Fusang Zhang, Zhaoxin Chang, Kai Niu, Jie Xiong, Beihong Jin, Qin Lv, and Daqing Zhang. 2020. Exploring lora for long-range through-wall sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4, 2, (June 2020). DOI: 10.1145/3397326. https://doi.org/10.1145/3397326.

[166] Brian D. Ziebart, J. Andrew Bagnell, and Anind K. Dey. 2010. Modeling interaction via the principle of maximum causal entropy. In *Proceedings of the 27th International Conference on International Conference on Machine Learning* (ICML'10). Omnipress, Haifa, Israel, 1255–1262. ISBN: 978-1-60558-907-7. http://dl.acm.org/citation.cfm?id=3104322.3104481.

[167] 2008. *Navigate like a cabbie: probabilistic reasoning from observed context-aware behavior*. *Proceedings of the 10th International Conference on Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 322–331. ISBN: 9781605581361. https://doi.org/10.1145/1409635.1409678.

[168] Tjalf Ziemssen, Raimar Kern, and Katja Thomas. 2016. Multiple sclerosis: clinical profiling and data collection as prerequisite for personalized medicine approach. *BMC neurology*, 16, 1, 124.