

Classification of human activity and social signals: a meta-analysis of RESpeck datasets

Passara Chanchotisatien



Master of Science
School of Informatics
University of Edinburgh
2022

Abstract

The RESpeck is a wearable device used to record activity level, breathing rate, and three-dimensional acceleration and angular velocity data of the wearer. In past projects, several labelled human activity and social signals datasets have been collected from healthy volunteers with the RESpeck device. From this, various machine learning methods and models were developed to recognize and classify human activities (standing, sitting, walking, running, etc.) and social signals (talking, coughing, hyperventilating, laughing, eating, etc.) and have been proven to work well in accomplishing these tasks. However, the performance of these methods is still unknown as they have yet to be applied to real-world unlabelled datasets. The three machine learning models that were developed in previous projects consist of an Auxiliary Classifier Generative Adversarial Network (AC-GAN) to classify human activities, 1D CNN models to classify social signals, and a hierarchical step counter model. This study has applied these models to eleven real-world unlabelled RESpeck datasets collected by the Centre of Speckled Computing at the University of Edinburgh's School of Informatics collected with the RESpeck device. Another device used to collect data was the Airspeck-P, which measured PM (PM1, PM10, and PM2.5) exposure levels, humidity, temperature and GPS data. The eleven datasets include INHALE, Daphne AAP, Daphne MCC, Philap, Apollo-C, Dublin, NHS Borders, Apcaps, PEEPS, QIP, and Leon, which consist of subjects with varying levels of activity and morbidity. From the classification results along with recorded data from the RESpeck and Airspeck sensors, we were able to evaluate the existing models' performance on unlabelled real-world datasets, and developed an alternative HAR model using a Random Forest classifier with manual feature extraction to address the shortcomings of the previous HAR model. The data and classification results were visualized, analyzed and interpreted. From this, we were also able to make observations and analyses from the classification results and RESpeck data in the context of the subjects' activity, detected social signals and their health and see whether our results showed any trends among these subjects and draw connections between our results with external literature.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Passara Chanchotisien)

Acknowledgements

I would like to thank

- Professor D.K. Arvind, who supervised this project, for his unwavering dedication and support throughout the project. Through his insightful comments, advice, and encouragement, this project was able to take its current form that is now being presented.
- Sharan Maiya for his much valued advice during weekly project meetings and assistance with helping me become familiar with using the Specklass library.
- Teodora Georgescu for providing helpful documentation on using previously developed machine learning methods.
- My parents for their unconditional support and encouragement.

Table of Contents

1	Introduction	1
1.1	Motivation and Problem Statement	1
1.2	Research Aims	2
1.3	Contributions	2
1.4	Thesis Structure	3
2	Background and Related Works	4
2.1	Human Activity Recognition	4
2.2	Social Signals Classification	4
2.3	Step Counting Systems	5
2.4	Machine Learning and Deep Learning Approaches	5
2.4.1	Ensemble Machine Learning Classifiers	5
2.4.2	Dimensionality Reduction	5
2.4.3	Convolutional Neural Networks (CNN)	6
2.4.4	Generative Adversarial Networks Networks (GAN)	6
2.4.5	Gated Recurrent Unit (GRU) Networks	6
2.5	RESpeck and Airspeck Device	6
2.5.1	RESpeck Device	6
2.5.2	Airspeck Device	7
2.6	Description of RESpeck datasets	7
2.7	Previous work done on RESpeck datasets	9
3	Methodology	10
3.1	Project Methodology Overview	10
3.2	AC-GAN HAR Model	11
3.2.1	Dataset and Data Pre-processing	12
3.2.2	Auxiliary Classifier Generative Adversarial	12

3.2.3	AC-GAN Performance and Implementation on Unlabelled Datasets	13
3.3	HAR with Ensemble Classifiers and Dimensionality Reduction Methods	13
3.3.1	Dataset, Data segmentation, Feature Extraction, and Data Scaling	14
3.3.2	Ensemble Classifiers and Dimensionality Reduction Methods	15
3.3.3	Performance of Models on Labelled Data	15
3.3.4	RF Classifier Implementation on Real-world Datasets	16
3.4	1-D-CNNs Social Signals Classification Models	16
3.4.1	Dataset, Data Pre-processing and Feature Extraction	16
3.4.2	Model Description	16
3.4.3	Model Performance on Labelled and Real-world Data	17
3.5	Hierarchical Step Counting Machine Learning Model	17
3.5.1	Model Structure	18
3.5.2	Frequency and Domain Step Counting Algorithm	18
4	Results and Discussion	19
4.1	Comparison of activity and social signals between datasets	19
4.2	INHALE Analysis	22
4.3	Apollo-C Analysis	25
4.4	Dublin Analysis	28
4.5	Daphne MCC Analysis	31
4.6	NHS Borders Analysis	33
4.7	Apcaps, PEEPS, QIP, and Leon Analysis	35
4.8	Evaluation and Discussion	37
4.8.1	Performance of models on real-world datasets	37
4.8.2	Limitations of the study	38
5	Conclusion and Future Works	39
5.1	Summary	39
5.2	Future Works	40
	Bibliography	41
A	Additional Figures	54

Chapter 1

Introduction

1.1 Motivation and Problem Statement

The RESpeck is a wearable device developed by the Centre for Speckled Computing at the University of Edinburgh's School of Informatics and is used to sense non-gravitational accelerations, and rotational velocities which are analysed to derive breathing rate/flow and intensity of physical activity. In previous projects, a library of several machine learning methods have been developed to perform Human Activity Recognition (HAR) and social signals classification on labelled datasets collected from healthy volunteers (primarily from students in the University) using the RESpeck device. During data collection, subjects were asked to perform a variety of basic day-to-day activities (sitting, standing, walking, running, shuffling, cycling, and lying down in various positions) and social signals (talking, coughing and hyperventilating). Machine learning methods were developed to classify a selection of these activities, social signals, and also count the number of steps walked. The methods include an Auxiliary Classifier Generative Adversarial Network (AC-GAN) [21] for HAR, a series of 1D Convolutional Neural Network (CNN) models [33] for social signals classification, and a three-layer hierarchical deep learning model with a frequency domain step-counting method [82] to determine the cumulative number of steps walked.

In the past, there have been many RESpeck datasets that were collected from a variety of subjects with varying levels of activity and types of morbidity and spread of ages including healthy subjects (from countries in four continents), asthmatic adolescents, adult asthmatics, adult COPD patients, pregnant women, discharged COVID-19 hospitalised patients, and post-operative patients. In addition to RESpeck data, several datasets also collect Airspeck data, which consists of personal exposure levels to PM

particles for each subject. On labelled datasets collected from volunteers, the AC-GAN was able to achieve over 90% accuracy score in classifying human activities, and the 1D social signals models were able to detect coughs with over an 80% accuracy. Though the mentioned existing machine learning methods have been proven to perform generally well on labelled datasets, their performance on real-world unlabelled datasets is still unknown as they have never been applied to real-world datasets before. This project will provide insights on the performance of these models on real-world datasets and whether modifications to the models are needed or whether new methods need to be developed along with performing analysis on RESpeck datasets in terms of activity types, activity levels and social signals.

1.2 Research Aims

The aim of this project is to perform meta-analysis on these real-world RESpeck datasets. Existing machine learning methods for HAR and social signals classification will be applied to real-world unlabelled datasets to gain insight on the performance of these methods and generate observations and insights in the context of the activity types performed, their activity levels, and detected social signals of these real-world subjects through data analysis and visualization from classification results. Furthermore, this project will aim to draw connections between known characteristics of the datasets (such as questionnaire data and a subject's exposure levels to air pollution) and visualized results to gain a deeper understanding of the real-world RESpeck datasets and evaluate the performance of existing methods. Finally, we would also like to know whether any connections can be drawn between the results obtained and findings from previous literature.

1.3 Contributions

The contributions of this project is outlined below:

- Applied existing machine learning methods (HAR, social signals classification, and step counting model) to eleven real-world RESpeck datasets and evaluated the performance of those models based on known characteristics of the datasets.
- Developed, implemented, and evaluated an alternative HAR model which addressed the shortcomings of the AC-GAN.

- Visualized activity level, breathing rate, PM exposure levels, social signals and activities types performed within each dataset, across datasets and for specific participants of interest.
- Analyzed the visualized results and identified overall trends and insight from the datasets in the context of activity, social signals, morbidity and PM exposure.
- Drew connections between our results and observed trends with findings from external literature and explain whether our findings support or contradict the conclusions made by previous studies.

1.4 Thesis Structure

The dissertation is organized as follows. Chapter 2 will provide the background theory and information behind the methods used in this study, along with a description of related works on human activity recognition systems, social classification methods, and step counting algorithms. The chapter will also provide a review of previous works that have been developed and implemented on RESpeck datasets. Chapter 3 outlines the methodology used in the project and the performance of the methods on labelled and unlabelled datasets. Chapter 4 presents the visualized results of the project along with detailed explanations, observations, insights, and a discussion of the performance of the models on unlabelled datasets and limitations in the study. Finally, Chapter 5 concludes the thesis and provides suggestions for potential future works.

Chapter 2

Background and Related Works

2.1 Human Activity Recognition

Human Activity Recognition (HAR) involves classifying and identifying different activities recorded by different sensors. HAR is a challenging task, especially when dealing with sensor data that is often noisy. Additionally, activity signals can vary from one individual to another, making it difficult to generalize and classify more complex dynamic activities. HAR has been studied extensively in the past on different labelled datasets. Studies have come up with a variety of approaches to identify and classify human activities [91][11][74]. Common methods which have proven to work well include machine learning models such as Decision Trees [18], Support Vector Machines (SVM) [62], and Random Forest (RF) [30]. These methods have often worked well when accompanied by the correct manual feature engineering and feature extraction. However, many studies have explored different deep learning approaches as well that have been reported to achieve high performance, including Convolutional Neural Networks [64], Recurrent Neural Networks [46], and Long Short Term Memory [63].

2.2 Social Signals Classification

One of the main tasks in this project involves social signals classification. Examples of social signals include hyperventilation, breathing deeply, coughing, talking, eating, and singing. The presence of abnormal social signals such as constant coughing or hyperventilation can signal to a person's deteriorating health or early symptoms of respiratory illnesses such as asthma or Chronic obstructive pulmonary disease (COPD).

Therefore, there has been many efforts to develop various machine learning [92][86] and deep learning algorithms [8][57][17] to help detect, identify and classify these social signals. Some common examples include using CNNs and RNNs to classify social signals [9][50] and Random Forest to detect talking [36].

2.3 Step Counting Systems

With the rise of electronic gadgets and wearable devices such as the Apple Watch and FitBit, which constantly collect sensor data, there has been much-increased research in developing methods to track the number of steps walked each day. Popular step counting algorithms that have proven to work and is used in many of these devices include above-threshold acceleration [41][52], short-term Fourier transform, wavelet transform [94], and peak detection [15].

2.4 Machine Learning and Deep Learning Approaches

2.4.1 Ensemble Machine Learning Classifiers

Ensemble learning is when a new classifier is obtained through the generation of multiple base classifiers. The new classifier will often perform better than the individual base classifiers. Examples of ensemble classifiers include stacking, bagging, and boosting. Popular ensemble machine learning classifiers include Random Forest [14], XGBoost [23] and LightGBM [53]. Random Forest is considered a bagging algorithm and is an ensemble of decision trees.

2.4.2 Dimensionality Reduction

Dimensionality reduction is a statistical technique used to reduce the number of random variables to help decrease the model's complexity and avoid overfitting [85]. Principal Component Analysis [4], also known as PCA, is the most common dimensionality reduction approach that reduces features from a higher dimension into a lower one. Finding and selecting the fittest principal components can maximize the variance of the previous higher dimension in a newer lower dimension space. Incremental Principal Component [98] uses a different calculation method from PCA by using mini-batching that allows it to reduce the memory required at any one time while closely achieving

the same results as PCA. Sparse Principal Component Analysis [81] is a variant of PCA that attempts to solve one of the disadvantages of PCA by using sparse vectors to represent the new dimension space instead of using dense expressions.

2.4.3 Convolutional Neural Networks (CNN)

A Convolutional Neural Network (CNN) [69] is a neural network composed of at least one convolutional layer. It is often used for processing, classifying, and segmenting image data, and consists of three main layers: a convolutional layer, a pooling layer, and a dense layer. CNNs are used for various tasks as they can extract features without requiring manual intervention.

2.4.4 Generative Adversarial Networks Networks (GAN)

Generative Adversarial Networks (GAN) allow for the generation of new data through learning and the patterns or regularities found in input data. A GAN consists of a generator and a discriminator, which participate in a supervised learning problem where the generator produces a target output, and the discriminator will learn to determine whether the output is real or fake.

2.4.5 Gated Recurrent Unit (GRU) Networks

A Gated Recurrent Unit (GRU) [31] network is a type of Recurrent Neural Network (RNN) [105], which deploys a gating mechanism. For specific applications, GRUs will benefit over long short-term memory (LSTM) as in GRU, less memory is used and has shorter operation times than LSTM. In GRU, connections through a sequence of nodes are used to perform machine learning tasks that involve clustering and memory.

2.5 RESpeck and Airspeck Device

2.5.1 RESpeck Device

The RESpeck [34] is a wearable device consisting of a tri-axial accelerometer and gyroscope. The RESpeck device records the breathing level, activity level, linear acceleration, and angular velocities in three dimensions (x , y , z). The RESpeck device is typically worn on the left lower coastal margin (below the ribs) as this position was

considered the most suitable and most effective for measuring the wearer's activity level and breathing rate and is sensitive to chest movements while breathing [34]. Data packets from the sensor are sent to an Android App through Bluetooth Low Energy (BLE) at a sampling rate of 12.5 Hz. An image of the RESpeck device is shown in Figure A.1.

2.5.2 Airspeck Device

There are two kinds of Airspeck [13][12] devices: Airspeck-P and Airspeck-S. The Airspeck-S is a stationary device placed outside to record ambient concentrations of airborne particulates, whereas the Airspeck-P is a wearable device to record real-time personal exposure to PM (PM10, PM2.5, and PM1) along with temperature, humidity, and GPS data. This study will be using data collected from the Airspeck-P. Figure A.2 shows an image of the Airspeck-P device.

2.6 Description of RESpeck datasets

In this project, analysis, HAR and social signals classification has been performed on a total of eleven unlabelled RESpeck datasets: INHALE, DAPHNE-AAP, PHILAP, DAPHNE-DMC, QIP, NHS Borders, Dublin, APOLLO-C, APCAPS, WHO-PEEPS and Leon dataset. All datasets were collected using the RESpeck and Airspeck-P devices.

INHALE dataset. This dataset consists of 10 older adult asthmatics (ages 40 years and above) and 13 healthy subjects of balanced gender and socio-economic status from London, UK. Subjects wear the RESpeck for two weeks with the exception of when they take a shower. Subjects wear the Airspeck Personal when they are outdoors and they either wear them or place it near them when they are indoors. Each subject wears the RESpeck and Airspeck twice- once during the winter months and once during another season.

DAPHNE-AAP dataset The DAPHNE-AAP dataset is collected on asthmatic adolescents (10-18 years of age) in Dehli, India. Subjects are given the RESpeck and Airspeck device which they wear for 48 hours. Subjects go through three cycles of data collection, at least one of which was during the winter months.

PHILAP dataset. The PHILAP dataset is collected by young asthmatics ranging in ages between 9 and 15 years in Delhi, India. The subjects wear the RESpeck and Airspeck sensors for 48 hours. Subjects wear the sensors twice- once during the winter

months and once during another season.

DAPHNE-MCC dataset. The DAPHNE-MCC (Mother-Child cohort) is a dataset consisting of data collected from 58 pregnant women who wear the RESpeck and Airspeck sensors for 24 hours once during each trimester. Out of 58 subjects, 32 subjects have completed data collection for their second trimester and 13 subjects have completed data collection for their third trimester.

QIP dataset. The QIP dataset is a dataset consisting of post-operative patients from the Western General Hospital. The subjects wear the RESpeck device for 3 to 10 days (depending on the type of operation). Some subjects have undergone back surgery, while others had simple procedures. Subjects spend most of their time in bed, but are sometimes moving around. Subjects in this dataset do not wear the Airspeck device.

Dublin dataset. The Dublin dataset is a dataset consisting of 30 asthmatics and 30 COPD patients from Liberec, Dublin and Madrid. The aim of this dataset is to study the impact of second-hand smoking (SHS) on Chronic Obstructive Pulmonary Disease (COPD) and asthmatic subjects. Subjects wear the RESpeck and Airspeck devices on them for 24 hours. During a point in the day, the subjects are brought to an enclosed smoking area where they are exposed to SHS for around 30 minutes to 2 hours. At the end of the 24 hours, the sensors are collected from the subjects. Additional data such as their age, weight, smoking status, whether they live with a smoker and number of smokers present during SHS period are also recorded.

NHS Borders dataset. In this dataset, 20 COPD patients from NHS Borders, UK, wear the RESpeck device for three months. Patients are asked to perform pulmonary rehab (PR) exercises as often as they can during this period. The COPD patients are divided into three groups consisting of COPD patients with a recent exacerbation, COPD patients with anxiousness, and COPD patients who attend virtual PR sessions. The aim of the dataset is to monitor changes in the patients' activity and social signals over time as they continue to perform PR exercises. Other data that is collected on the patients include data from rehab sessions and a symptoms diary.

APOLLO-C dataset. The APOLLO-C dataset is collected on recently-recovered COVID-19 patients. Subjects are given the RESpeck and Airspeck devices to wear over 24 hours. These subjects will be asked to participate in data collection for 14 times (24 hours each time) over a duration of 6 weeks. The purpose of this dataset is to study and observe the changes in the subjects' activity, social signals, and the effect of air pollution on their recovery over time after having recovered from COVID-19.

APCAPS dataset. The APCAPS dataset consists of over 100 healthy subjects

living in rural areas. Subjects wore the RESpeck device for 24 hours and are expected to be quite active during this period.

WHO-PEEPS dataset. The WHO-PEEPS dataset was collected on healthy individuals who are UN employees in Dehli. Subjects wear the RESpeck and Airspeck device for 24 hours.

Leon dataset. The Leon dataset collected RESpeck and Airsepck data from healthy municipal government employees. The subjects walked or cycled the same 2-3 routes every week day between 8 and 10 am.

2.7 Previous work done on RESpeck datasets

Much previous work has been done on RESpeck datasets in the past regarding developing different machine learning methods to perform HAR, social signals classification, and step-counting. In [39], Teodora Georgescu developed a model using the Random Forest classifier to classify coughing episodes in real time. Following that, Nikita Nikolajev [67] explored various traditional machine learning methods, specifically XGB OvR and neural networks such as MobileNet, to perform multi-class social signal classification where the social signals being classified were laughing, coughing, talking, eating, and breathing and was able to achieve an accuracy score of 76%. Subsequently, Celina Dong stated in [32] that when classifying social signals, a class called 'Other' should be added to capture the remaining signals. Therefore, in [33], a two-class, three-class, and five-class 1D CNN model was developed to classify the following social signals: coughing, talking, breathing, hyperventilating, and other.

For HAR, Stylianos developed a novel AC-GAN model [21] to classify 11 types of activities and achieved an accuracy of 91% on labelled data. Later, to further improve classification performance, various deep learning models, including the AC-GAN were tested on 25 HZ RESpeck data [22]. However, this model is not used as most data is collected at 12.5 Hz. Finally, to count the number of incremental steps, Shuai Shi developed an end-to-end real-time step counting algorithm from RESpeck data [82]. In this project, we will be applying methods from Charalambous' AC-GAN [21], Dong's social signals models [33], and Shi's step counter [82].

Chapter 3

Methodology

3.1 Project Methodology Overview

This project involves applying existing machine learning and deep learning methods that have been demonstrated to work on existing labelled RESpeck datasets collected on volunteers to real-world unlabelled RESpeck datasets and evaluating the performance of these models. An alternative machine learning model for HAR was also proposed, which reduces the number of misclassifications found in the existing HAR model. This project also aims to perform a meta-analysis of the aforementioned real-world RESpeck datasets in terms of their activities and detect social signals through data analysis and visualization of classification results of the machine learning models. The diagram illustrating the overview of the project methodology can be seen in Figure 3.1. Firstly, the raw RESpeck data is input into the HAR model. We explore two HAR models in this project: the AC-GAN, developed in [21], and another developed by the author in this project, which implements a Random Forest classifier with manual feature extraction. After that, the HAR model will output the activity classifications with the log probabilities, which will then be given to a Finite State Machine (FSM). The FSM improves the HAR classifications by removing impossible activity transitions. An example of an impossible activity transition is, for instance, going from 'lying down' to 'cycling' without any movement in between. If an impossible transition is detected, the activity with the next highest log probability will be taken. As a result, the FSM will return the activity classifications. To detect social signals (SS), we will then filter for the stationary activities only using the classification results from the FSM, as the social signals classification models can only classify social signals while the subject is stationary. The raw stationary RESpeck data will then be input into the 1-D CNN

SS 2-class model, 1-D CNN SS 3-class model, and 1-D CNN SS 5-class model to detect coughs, talking, and hyperventilation, which Celina Dong developed in [33]. The detection of coughs will be verified using the log probabilities of the three SS models. If the log probability of a cough is above 0.9 for all three models, then we can determine that the subject is indeed coughing. To get the cumulative number of steps at each timestamp, the raw RESpeck data can be input directly into the Step Count model, developed by Shuai in [82].

Next, the output activity classifications, social signals classification, and step count are aggregated for each subject, re-sampled into minutes, and combined with the original raw RESpeck data and Airspeck data containing activity level, breathing rate, and PM exposure values. Therefore, for each minute of the monitoring period, we will have the activity level, activity type, breathing rate, step count, whether a cough was detected, whether hyperventilation was detected, and PM (PM1, PM2.5, and PM10) exposure levels. This aggregated data can then be visualized and further analyzed.

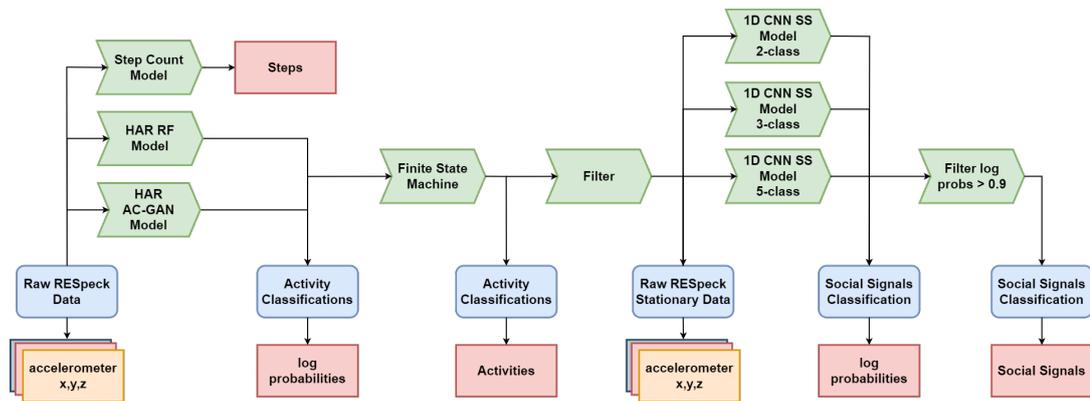


Figure 3.1: Physical activity and social signals classification: Project methodology overview

3.2 AC-GAN HAR Model

The AC-GAN HAR model is a deep learning model developed in [21] which implements an Auxiliary Classifier Generative Adversarial Network method to classify a total of eleven activities, including ascending stairs, descending stairs, lying on their back, lying on their left side, lying on their right side, lying on stomach, sitting/standing, cycling, misc. movement, running, and walking from RESpeck data.

3.2.1 Dataset and Data Pre-processing

The model was trained on a labelled RESpeck dataset collected from 14 healthy volunteers in [21]. Participants were asked to perform 19 everyday activities and 12 social signals for a minimum of 30 seconds each. The activities included sitting normally, sitting bent backwards, sitting bent forwards, standing, lying, lying on back, lying on stomach, lying on left side, lying on right side, walking at normal pace, shuffling (slow-walking), ascending stairs, descending stairs, cycling, running, swinging back and forth in place, standing up from sitting down, sitting down from standing up, getting up from lying down, and lying down from sitting. Social signals included coughing, talking, eating, drinking, singing, laughing, breathing normally, breathing deeply, sighing, sobbing, yawning, hiccuping, and hyperventilating. The distribution of classes [21] can be seen in the Appendix in Figure A.3.

Activities selected for classification were ascending stairs, descending stairs, lying on back, lying on left side, lying on right side, lying on stomach, sitting/standing, cycling, misc. movement, running, and walking. A median filter algorithm was applied to identify noisy signals. Then, the sliding window technique was utilized to segment the data into windows of size 48 timestamps equal to approximately 4 seconds with an overlapping rate of 50%.

3.2.2 Auxiliary Classifier Generative Adversarial

Stylianou [21] implemented an architecture for HAR and social signal classification called the Auxiliary Classifier Generative Adversarial Network (AC-GAN) that was inspired by the Semi-Supervised Generative Adversarial (S-GAN) Network developed in [48]. It extends S-GAN by providing the generator with the label of the activity as input. The architecture of the AC-GAN model can be seen in Figure A.5 [21] and Figure A.4, respectively.

The discriminator follows the same framework as S-GAN. On the other hand, the generator concatenates the feature extraction designed for the S-GAN along with a Fully-Connected (Dense) layer to convert the input label into a feature map. A diagram [21] of this can be seen in Figure A.5. In each iteration, the discriminator is provided with 50% real samples, and 50% generated samples that are chosen randomly. Afterward, the discriminator has some layers frozen and trained using randomly generated points to improve the generator model. The model is trained on 25,000 mini-batches with batch sizes of 128 samples using the Adam optimizer. The class labels are randomly picked

from a uniform distribution of classes.

3.2.3 AC-GAN Performance and Implementation on Unlabelled Datasets

In Stylianos's study, the AC-GAN was shown to perform well in classifying different activities in labelled RESpeck datasets, achieving an accuracy of 0.91 with 10-fold cross-validation. Here, to investigate its performance on the real-world RESpeck datasets, we applied the AC-GAN model to the eleven RESpeck datasets described in Section 2.3.

Previously, we have seen AC-GAN's ability to accurately classify human activities on labelled data. Here, the model was applied to the eleven real-world mentioned datasets in Section 2.3, and the classification results were visualized. The activity type classification (in minutes) for each subject in the Apcaps, PEEPS, Dublin, and QIP datasets are shown in Figures A.6(a)-(d), respectively. From these activity plots, we observe an unrealistically large number of cycling classifications for almost every subject. For example, in Apcaps, we see many subjects spend up to 800 (13 hours) minutes cycling over a span of 24 hours, which is impossible. The same can be observed for PEEPS and Dublin as well. Subjects in the Dublin dataset are also seen to be running for extremely long periods. For example, the subject 'DBIA02', spent close to 1000 minutes (16 hours) engaging in running and cycling in the span of 24 hours, which is highly unlikely. Additionally, several subjects in the QIP dataset are shown to be cycling as well. This is impossible as subjects from the QIP dataset are post-operative patients and are recovering in the hospital. Consequently, these are likely to be misclassifications. In fact, there are very few (close to none) classifications for 'walking' for many of these subjects, which is likely incorrect as subjects from all datasets (except for QIP), are expected to have walked each day. From this, we can conclude that the AC-GAN is unreliable and does not perform well on real-world datasets.

3.3 HAR with Ensemble Classifiers and Dimensionality Reduction Methods

Due to the misclassifications from AC-GAN, we decided to implement other methods to perform HAR instead. The efficiency and accuracy of ensemble machine learning classifiers for HAR such as XGBoost, LightGBM, and Random Forest (RF) have been mentioned in many past studies [20][90]. In [21], Stylianos implemented the RF classifier on the same dataset mentioned in Section 3.2.1 but without prior feature

extraction as a baseline method. Despite this, the RF classifier (with a window size of 48) attained an average accuracy score of 0.85. For reference, the AC-GAN achieved an accuracy score of 0.91 for the same window size [21]. Authors from several previous studies [7][90][20] have shown that with the correct features, ensemble machine learning classifiers are capable of performing HAR accurately. Though multiple studies have suggested deep learning models such as RNN [59] and LSTM [38], training and predicting times are significantly longer for these models. Still, ensemble methods are able to achieve comparable results with the right features [24][45]. Consequently, as this study requires the implementation of the HAR model on different RESpeck datasets, we opted to explore ensemble classifiers instead.

3.3.1 Dataset, Data segmentation, Feature Extraction, and Data Scaling

The dataset used for training and testing consisted of the same labelled 12.5 Hz RESpeck data collected from volunteers described in Section 3.2.1 with the addition of 25 Hz RESpeck data from [22], which was downsampled to 12.5 Hz. This was necessary as the real-world datasets were recorded at 12.5Hz; therefore, using the same frequency will allow the model to generalize between training and the real dataset. Previously, activities such as slow-walking / shuffling and swinging in place were excluded for AC-GAN. Here, the dataset includes two additional classes: 'Shuffling' and 'Not Worn.' The 'Not Worn' class was added to detect when the RESpeck device was turned on and collected data but not worn on the subject. Additional activities, such as swinging or rotating in place, and transitional activities, such as getting up from a chair or the movement of sitting down, were added to the 'Misc. Movement' class to help accommodate a broader range of activities.

The frames were divided into 4-second segments consisting of 50 time-stamped sensor values without overlapping. It has been established in prior studies that the presence and absence of overlapping windows do not impact the model performance for HAR [28]. One frame will correspond to a single activity. Each dataset file is iterated individually to prevent mismatched frames between files/persons. In the end, 37735 frames were created using this data segmentation method. In summary, the activities (and distribution of) included for classification were ascending stairs (3854 frames), descending stairs (3462 frames), lying on back (318 frames), lying on right side (1438 frames), lying on left side (1437 frames), lying on stomach (1313 frames),

sitting/standing (7388 frames), cycling (4239 frames), misc. movement (837 frames), running (3671 frames), shuffling (769 frames), walking (8920 frames), and not worn (89 frames).

We then extracted 14 features from each frame's x, y, and z acceleration. The extracted features include mean, median, mode, standard deviation, max, min, range, skew, kurtosis, 10th percentile, 25th percentile, 50th percentile, 75th percentile, and 90th percentile. These features were chosen as they have been proven to work well and achieve high accuracy scores for activity classification in many HAR datasets [90][20]. This is because, in this case, we are extracting features that we know are relevant and beneficial to the classification performance. Standard Scaler is used to standardize the features in the dataset to unit variance to control the impact of each feature on the model's weights and the model itself.

3.3.2 Ensemble Classifiers and Dimensionality Reduction Methods

Ensemble machine learning classifiers, including Random Forest, XGBoost, and LightGBM, were implemented on the labelled dataset. It can be affected by the curse of dimensionality or weak generalization due to the high number of features [85]. Dimensionality Reduction is one method to solve the problem and improve machine learning models' performance. Three dimensionality reduction methods were used in this section: Principal Component Analysis (PCA) [4], Incremental Principal Component Analysis (IPCA) [98], and Sparse Principal Component Analysis (SPCA) [108]. In our tests, we trained and evaluated our model, with and without dimensionality reduction, to be thorough in our investigation.

3.3.3 Performance of Models on Labelled Data

To evaluate and compare the performance between the different combinations of classifiers (RF, XGBoost, and LightGBM) and dimensionality reduction methods (PCA, SPCA, and KPCA), we apply Stratified K-Fold Cross-validation (K=10). The accuracy, F1, precision, and recall scores for each classifier and dimensionality reduction method is shown in Table A.1. The top-three highest accuracy scores are shown in bold. From the results, the Random Forest classifier, Random Forest + SPCA, and LightGBM + SPCA are the best-performing classifiers. With the aforementioned features extracted, RF achieved the highest accuracy score of 0.92.

3.3.4 RF Classifier Implementation on Real-world Datasets

Given RF's high performance on the labelled data, it was then applied to the eleven real-world datasets described in Section 2.3. The visualized classification results with RF for Dublin, Apcaps, PEEPS, and QIP datasets can be seen in Figures A.7, 4.21, 4.22, and 4.23, respectively. Compared to the previous visualized activity classification results with AC-GAN, the number of cycling classifications has been greatly reduced. Subjects from Dublin, Apcaps, and PEEPS are now mostly engaged in dynamic activities (ascending stairs, descending stairs, walking, cycling, and shuffling) on average between 100 and 300 minutes in 24 hours, which is much more realistic. As validation, the 'cycling' classifications for QIP subjects have disappeared as they are post-operative patients in the hospital and, therefore, incapable of this activity.

3.4 1-D-CNNs Social Signals Classification Models

Celina Dong's social signals classification methodology [33] consists of three 1D Convolutional Neural Network (CNN) models: a 2-class model ('coughing' and 'breathing'), a 3-class model ('coughing', 'breathing'), and a 5-class model ('coughing', 'breathing', 'talking', 'hyperventilation', and 'other').

3.4.1 Dataset, Data Pre-processing and Feature Extraction

The dataset used was the same dataset described in Section 3.2.1 with the addition of 25 Hz RESpeck data collected from 12 participants in [33] who were asked to perform 12 social signals across eight stationary activities for a minimum of 30s each. However, for the data to be usable, it was down-sampled to 12.5 Hz to match the frequency of the previous RESpeck dataset and real-world datasets. The following filter methods were used: smoothing, gradient, and double gradient.

3.4.2 Model Description



Figure 3.2: Social signals classification model structure

The model features a simple two-layer CNN with an L2 regularizer. The model uses a 4-second window size with an overlapping rate of 50%. The model was trained with

a learning rate of $1e-3$ with the Adam optimizer without decay. 128 filters and kernel size 5 was used. The dropout rate, which proved to perform best, was 0.3[33]. The architecture of the 1-D CNN social signals classification model [33] is shown in Figure 3.2. In the 5-class model, it was observed in [33] that the model struggled to classify the 'other' class due to having similar signal patterns as other classes. Consequently, the Extreme Value Theory, an out-of-distribution detection algorithm, was implemented as part of the CNN model.

3.4.3 Model Performance on Labelled and Real-world Data

In the case of labelled data, the 2-class model, 3-class model, and 5-class model achieved accuracy scores of 0.8279, 0.5424, and 0.4131, respectively. The models classify coughs well but struggle with the 'hyperventilation' and 'other' classes. Reasons for this will be discussed further in the 'Limitations' section. As for their performance on real-world RESpeck data, we can see an example of the visualization of social signals classifications in Figure A.8. The figure shows combined plots for social signal classification of the subject DBIA01 from the Dublin dataset. In the Dublin dataset, subjects are each exposed to second-hand smoking for periods ranging between 30 minutes to two hours over the 24 hours of data collection. In the plot, that duration is shaded in green. For this specific subject, we see that during the SHS period, there is a rise in PM_{2.5} levels, and at the same time, several coughing and hyperventilation episodes (circled in red) are detected by the model.

3.5 Hierarchical Step Counting Machine Learning Model

The step counting model implemented in [82] consists of two main steps: (1) classifying the activity being performed to identify whether the subject is walking and (2) applying the step counting algorithm according to the type of walking (normal walking, running, shuffling, ascending stairs, and descending stairs). In the first step (classification of walking), the step counting model features a hierarchical three-layer machine learning model, where a Gated Recurrent Unit (GRU) network is implemented at each layer. In the second step, a discrete wavelet transform-based step counting algorithm is utilized for each type of walking. Figure 3.3 illustrates the step-counting pipeline.

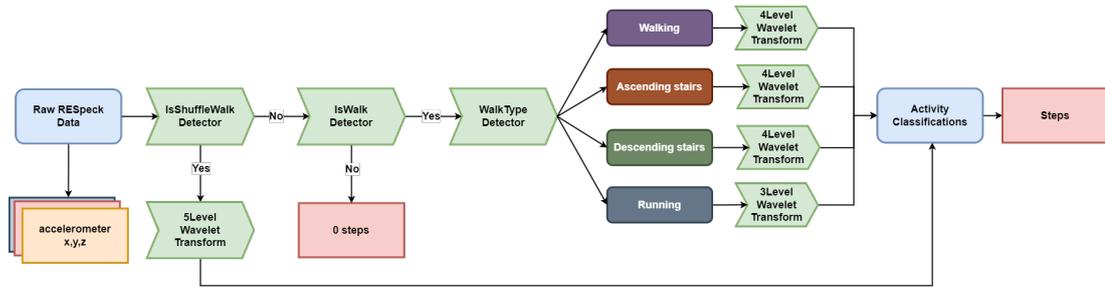


Figure 3.3: Step counting model pipeline

3.5.1 Model Structure

The classification of shuffle walking is performed in the first layer (IsShuffle) of the hierarchical machine learning model. Since the trends from accelerometer data for shuffle walking are rather different compared to other types of walking, it is classified first. Fluctuations in accelerometer values for shuffle walking are very minimal. After the classification of IsShuffle has been done, non-shuffle walking data is fed into the next layer/classifier. In the second layer (IsWalk), the classifier distinguishes whether the input data is walking (normal walking, running, ascending stairs, and descending stairs) or non-walking. This layer implements a Softmax optimizer, which is more suited for binary classification, and a window size of 64. Walking data is then fed into the third layer (FourTypeWalk) and classified into one of the following four types: ascending stairs, descending stairs, walking normally, and running / jogging. The step counting algorithm is then applied to data following the classification result.

3.5.2 Frequency and Domain Step Counting Algorithm

Walking is a cyclic movement that involves eight stages per gait, with one particular stage having peak acceleration. The main concept behind the step counting algorithm is to select the frequencies when the acceleration peaked for the different types of walks. Fast Fourier transform (FFT) determines the frequency range for the different types of walks. Therefore, signals with frequencies outside the range can be classified as noise and filtered out using wavelet decomposition and recombination. After applying the wavelet transform, the peaks from different wavelet transform levels are obtained following the frequency ranges determined with FFT. As the raw RESpeck data has already been smoothed out, we only need to detect the local maximum to find the peak. Lastly, peak detection is used to count the number of steps.

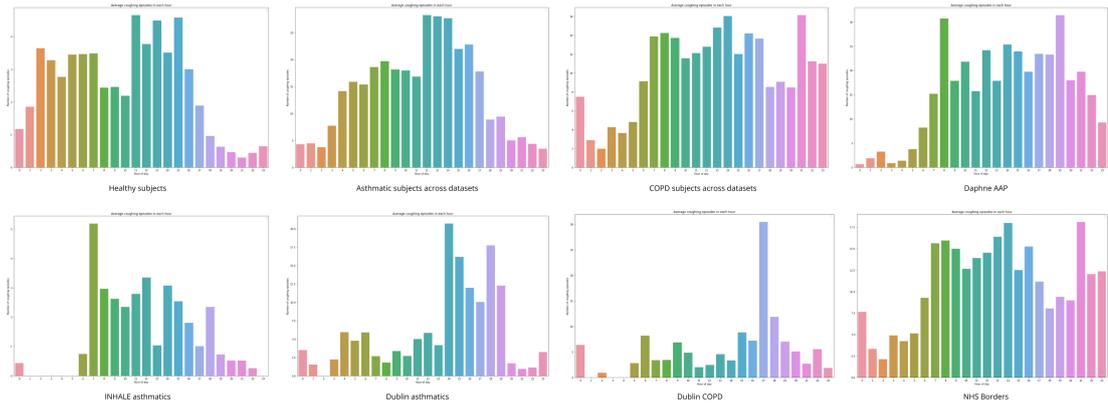


Figure 4.2: Number of coughing episodes by hour of day across datasets

seen that the Apcaps dataset is the most active dataset, with an average of 200 minutes of dynamic activity in a day. Dynamic activities include walking, running, shuffling, ascending stairs, descending stairs, and cycling. Apcaps is expected to be the most active dataset as the subjects were healthy and living in rural areas and would therefore be engaging more often in dynamic activities such as walking or cycling. On the other hand, while PEEPS subjects are also healthy subjects with no known afflictions, they were mostly UN employees who were living in more urban areas and commuted to work, thus resulting in a lower number of active minutes per day as compared to Apcaps. It has been stated in prior literature that people living in rural areas are often more physically active than those living in urban areas[44][49]. In contrast, QIP has the lowest number of daily active minutes (approximately 10 minutes). This is to be expected as the QIP dataset is collected on post-operative subjects who are hospitalized and lying down most of the time. Daphne AAP and PHILAP datasets are collected on both asthmatic adolescents. Subjects from this dataset are expected to be more physically active as they belong to a younger age group [73]. A study that investigated physical activity levels in 1591 adolescents (consisting of both asthmatics and non-asthmatics) found that 55% of the cohort were physically active and that the asthmatic adolescents were more active than their non-asthmatic peers [19]. Despite the datasets being both collected on young asthmatics in India, PHILAP has a much lower average activity level and a number of active minutes in a day. However, this is likely due to the subjects from PHILAP not wearing the Respeck device for extensive durations of time, which will be discussed in more detail later in the following sections.

From the breathing rate graph in 4.1(a), the datasets with higher breathing rates compared to others include Dublin Asthma, DAPHNE MCC, DAPHNE AAP, Apollo-C,

PHILAP, APCAPS, and PEEPS. We can attribute the higher breathing rate values from Dublin Asthma, Daphne AAP, Apollo-C and PHILAP to their condition and exposure to higher levels of PM_{2.5}. Daphne MCC's high breathing rate average, likely to be a combined consequence of the pregnancy and exposure to pollution, will be discussed in Section 3.6. At first glance, Apcaps' and PEEPS' high breathing rate averages might have been unexpected as the datasets are both collected on healthy subjects with no morbidity. Though we have no collected Airspeck data in the Apcaps dataset, if we look at PEEPS average exposure level as a cohort, we can see that PEEPS (along with Daphne AAP and Daphne MCC) have abnormally high exposure to air pollution compared to other datasets. This could be because they are living in urban areas in New Delhi where there are a lot of vehicular and industrial emissions[56], which therefore, may cause a decline in respiratory function and an elevated breathing rate [42] [19] [89]. We can compare these numbers to the breathing rate averages from the INHALE dataset for both asthmatic and healthy subjects. Though INHALE asthmatic subjects have a slightly lower breathing rate than INHALE healthy subjects, it is still lower than PEEPS despite PEEPS being healthy subjects. This can be explained by INHALE's extremely low PM_{2.5} exposure values, which can be seen in 4.1(b).

In Figure 4.1(c) and Figure 4.1(d), the average number of coughing and hyperventilation episodes are shown, respectively. From this, we can see that asthmatic subjects (from Daphne AAP, INHALE Asthma, Dublin Asthma) generally have a higher number of coughing or hyperventilation episodes than subjects with COPD. It can be observed that the NHS Borders and Dublin COPD datasets share similar values for breathing rate, coughing episodes, and hyperventilation episodes. Most asthmatics will experience regular dry coughs; however, for COPD patients, it will depend on the severity of the condition [102]. COPD is typically divided into four stages: mild, moderate, severe, and very severe [70][84]. In earlier stages of COPD, coughs will typically occur after physical exertion, then eventually while at rest in later and more severe stages [102]. Between datasets collected on asthmatics, Daphne AAP and PHILAP (collected on young asthmatics) have a higher number of coughing episodes as opposed to Dublin Asthma and INHALE asthma (collected on adult asthmatics). We can speculate that the reason Daphne AAP and PHILAP have a higher number of coughing episodes is because they are generally more active. Many asthmatic patients have reported that physical activity or exercise often triggers their asthma[58]. From this figure, we can see that INHALE healthy subjects have the lowest number of coughing episodes. This is likely because they are healthy and exposed to very low levels of PM_{2.5} on average, as

seen in Figure 4.1(b). However, strangely, healthy INHALE subjects have high average values for hyperventilation episodes. These are likely misclassifications and will be discussed more in the 'Model Performance Evaluation' section. The Apcaps dataset has the lowest average number of coughing and hyperventilation episodes per minute. This is likely due to the fact that they are healthy and active subjects living in rural areas where lower emission and air pollution is expected[107]. In comparison, PEEPS has a much higher number of average coughing and hyperventilation episodes per minute despite being healthy subjects and less active. This can be attributable to the fact that they are exposed to much higher levels of air pollution (as shown in 4.1(b)).

Figure 4.2 compares the number of coughing episodes in each hour of the day across different datasets. Firstly, it is evident that the healthy subjects have much lower coughing episodes than asthmatic and COPD subjects. It has been found that asthma and COPD symptoms will often worsen during nighttime or in the early hours of the morning [88][55][80][71][5]. This is reflected in the NHS Borders, Daphne AAP, and INHALE asthmatics datasets where there are a higher number of coughs during nighttime (Daphne AAP and NHS Borders) or early morning (INHALE asthmatics, NHS Borders, and Daphne AAP). However, Dublin asthmatics and Dublin COPD subjects experienced more coughs in the evening when they were brought to be exposed to second-hand smoking. Additionally, it can be noticed that subjects from the NHS Borders dataset experience more persistent coughing throughout the day as compared to Dublin COPD subjects, despite both datasets being collected on individuals with COPD. This is likely because subjects from NHS Borders are older patients as COPD develops gradually over many years, and most individuals will not have any notable symptoms until they reach their late 40's or 50's[77].

4.2 INHALE Analysis

From the hourly averages of breathing rate, the number of active minutes, and exposure levels in Figure 4.3(a), we can see that subjects are primarily active during the period between 6 am to 8 pm. We can also see raised levels of exposure to PM_{2.5} from 6 pm to 9 pm, which is the same time the breathing rate increases. This trend can be seen between 7 am to 9 am as well. This could be the result of cooking at those times of the day. Although there is no apparent difference between the breathing rate levels of INHALE asthmatic and healthy subjects, INHALE healthy subjects are generally more active than asthmatics subjects, as seen in Figure 4.3(b).

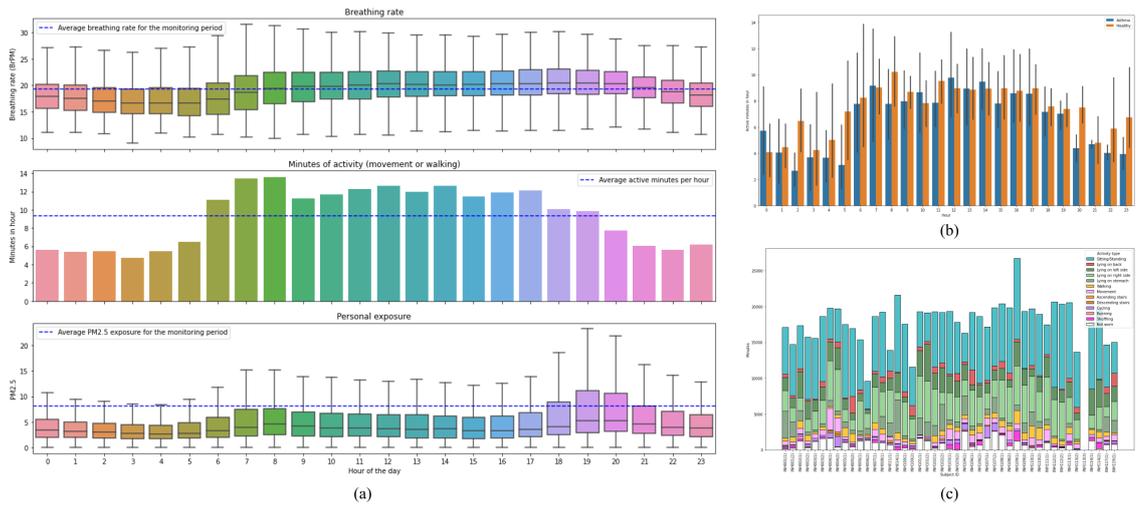


Figure 4.3: INHALE plots for breathing rate, activity, and exposure levels (a) Hourly averages for BR, active minutes and exposure (b) Active minutes in an hour split by cohort (c) Activity type classification

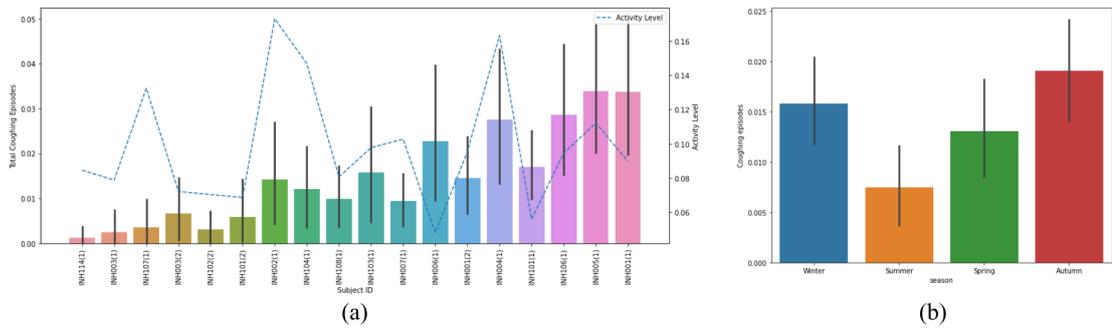


Figure 4.4: Average coughing episodes (a) Average coughing episodes per minute and activity level (b) Average coughing episodes by season

The average number of coughing episodes per minute, along with the average activity level for each subject, is depicted in Figure 4.4(a). Asthmatics are seen to experience a more significant number of coughing episodes. Especially in some asthmatic subjects, it can be noticed that the overall activity level is relatively low, but the number of coughing episodes is high. This means that the coughing is likely not triggered by activity but probably from asthma. However, this can also show that asthmatics are more sensitive to activity than healthy subjects are, which has been shown in several previous studies [26]. For some healthy subjects, we see that they have much higher activity levels but a very low number of average coughing episodes. We also see more coughing episodes in winter and autumn, as seen in Figure 4.4.

Figure 4.5 shows a comparison of the breathing rate at each activity of an asthmatic

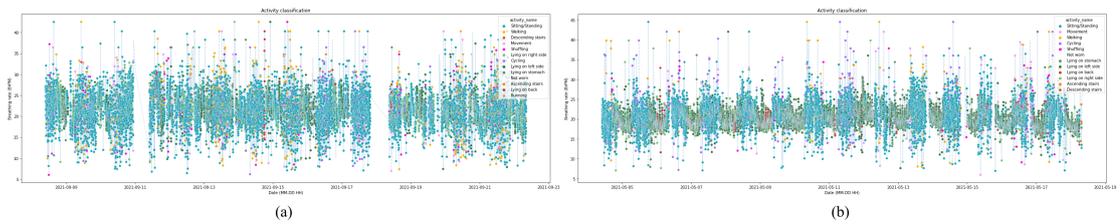


Figure 4.5: Comparison of breathing rate and activity classification between an asthmatic and healthy subject (a) INH005 (asthmatic) (b) INH102 (healthy)

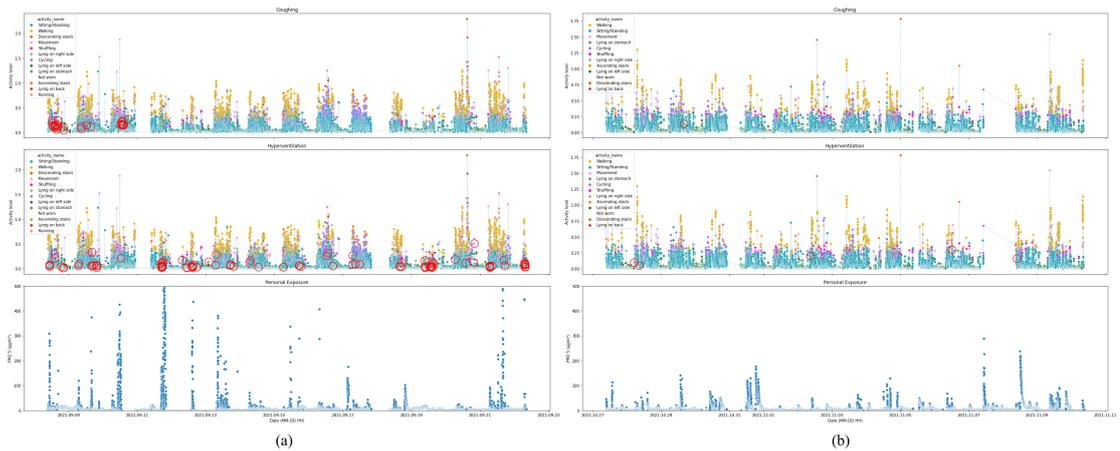


Figure 4.6: INHALE single subject combined plots for coughing, hyperventilation and exposure levels (a) INH005 (asthmatic) (b) INH102 (healthy)

subject (INH005 Figure 4.5(a)) and a healthy subject (INH102 4.5(b)). From this, it can be seen that in INH102 (healthy), the change in breathing rate is very visible from activities such as sitting/standing to lying down. We see a clear drop in breathing rate when the subject lies down or a large increase in the breathing rate when the subject performs a dynamic activity. However, if we look at INH005's (asthmatic) breathing rate trend, we do not see such a visible change in breathing rate as we did with INH005. For instance, we do not see a clear drop in breathing rate when the subject goes from sitting/standing to lying down or an obvious increase when the subject goes from a static activity to dynamic activity. INH005's breathing rate is consistently high throughout (even when at rest), and this trend can be commonly seen with other asthmatic subjects as well.

Figure 4.6 shows the comparison of the effect of air pollution on asthmatic subject (INH005), seen in Figure 4.6(a) as opposed to a healthy subject (INH102)4.6(b). As a cohort, subjects in this dataset are not exposed to high levels of air pollution. Therefore, it can be difficult to see the difference in response to air pollution between healthy and

asthmatic subjects. Therefore, two subjects were chosen- one healthy and one asthmatic, from the dataset who were exposed to higher levels of air pollution compared to other subjects. From the graphs, we see that each time there was a spike in PM2.5 values, INH005 will start to cough and sometimes hyperventilate, while on the other hand, INH114 does not seem to be affected by the high exposure values and only coughed once during the entire duration.

4.3 Apollo-C Analysis

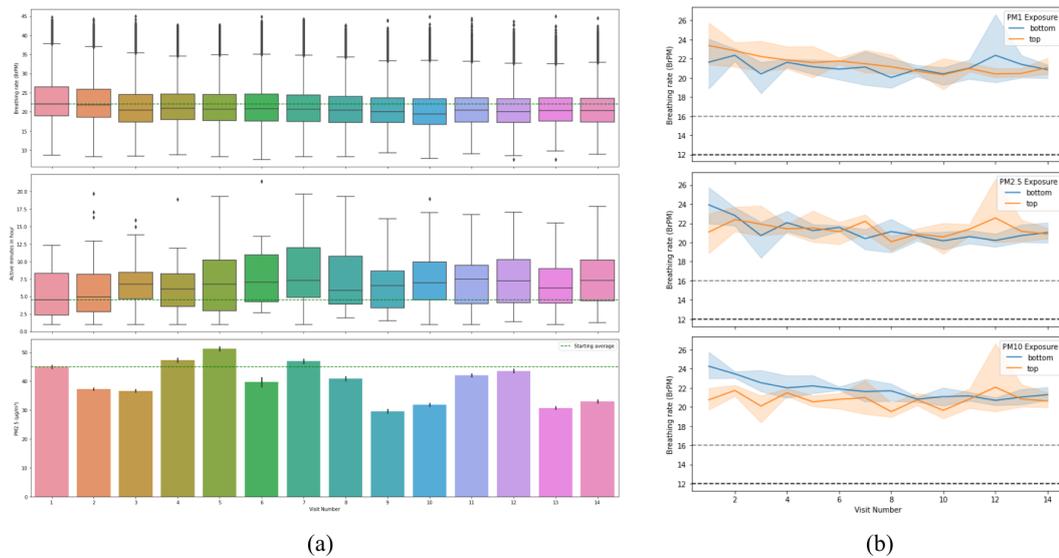


Figure 4.7: Apollo-C plots for breathing rate, number of active minutes, and PM2.5 exposure levels across visits (a) Averages for Br, number of active minutes in an hour, and PM1.5 exposure levels across visits (b) Effect of air pollution on breathing rate across visits

For the Apollo-C dataset, we want to investigate the recovery of lung function on breathing rate, physical activity levels, and the number of coughing and hyperventilation episodes as the patient’s recovery progresses (over 14 weeks). A lower breathing rate and increased activity level are associated with better recovery. Long COVID refers to persisting symptoms from COVID after the infection has already gone. These symptoms could last from weeks to months, depending on the person. Some symptoms of long COVID include shortness of breath, coughing, chest pain, and a sore throat [51] [75]. In one study conducted on post-Covid patients, it was stated that two-fifths of patients with long COVID reported a ”worsened quality of life” [60]. Most people with long COVID

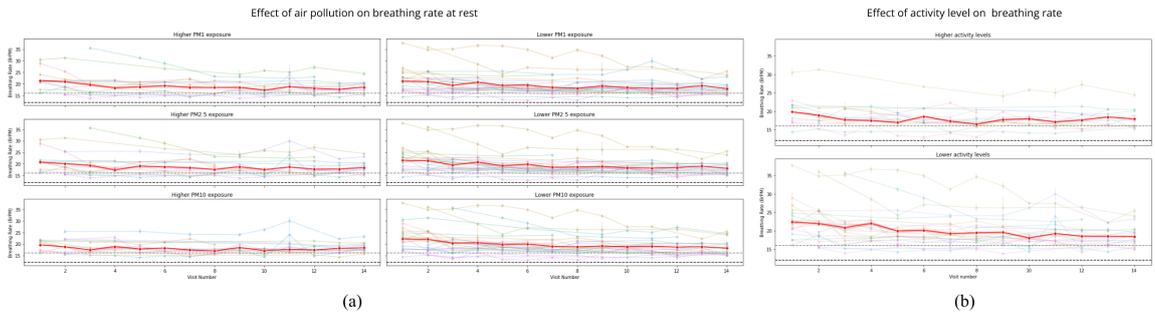


Figure 4.8: Apollo-C breathing rate trends (a) Effect of air pollution on breathing rate at rest (b) Effect of activity level on breathing rate

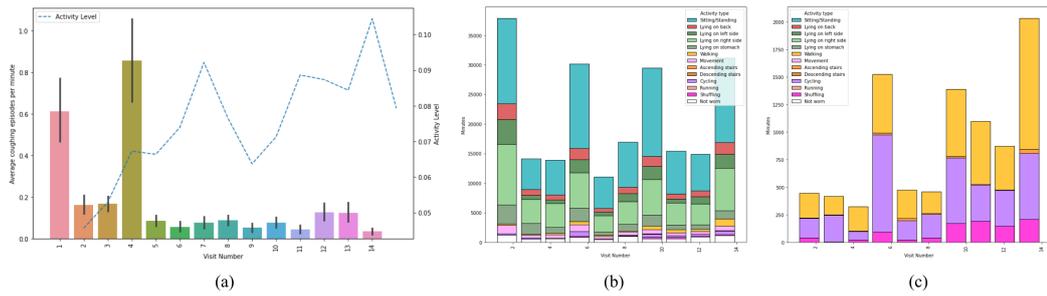


Figure 4.9: Apollo-C number of coughs and activity type distribution across visits (a) Average number of coughing episodes per minute across visits (b) Activity type distribution across visits in minutes (c) Dynamic activity distribution across visits in minutes

can make a recovery within 12 weeks [66]. Figure 4.7(a) shows the trends for breathing rate, the average number of active minutes in an hour, and exposure levels across visits. From the breathing rate plot, it is evident that the cohort’s average breathing rate slowly decreases with each visit. By visit 14, the average breathing rate is approximately 20 breaths per minute, while the breathing rate at the starting average is around 23 breaths per minute. Note that the regular breathing rate for the average adult is roughly 12 to 16 breaths per minute [3]; therefore, the cohort’s average breathing rates throughout the visits is much higher than that of an average person. This is likely because one of the symptoms of long COVID includes difficulty breathing and fatigue which can lead to a higher breathing rate and lower physical activity levels [75]. Similar trends can be seen for the average number of active minutes in an hour as well, where the cohort becomes more active as their recovery progresses. Figure 4.7(b) shows the impact of air pollution on the recovery of long COVID. Here, the cohort is divided into two groups: the top exposure and the bottom exposure group. The top exposure group consists of subjects with the top ten highest average exposure levels to PM2.5, PM1, and PM10. It

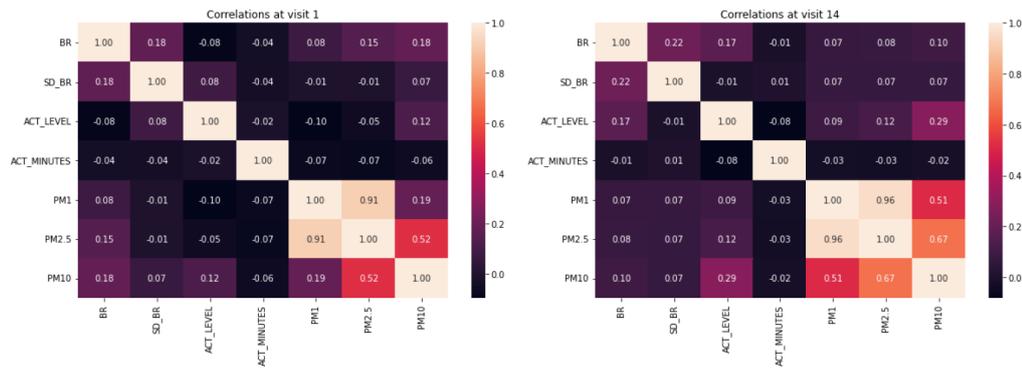


Figure 4.10: Apollo-C correlation heatmaps at visit 1 vs visit 14

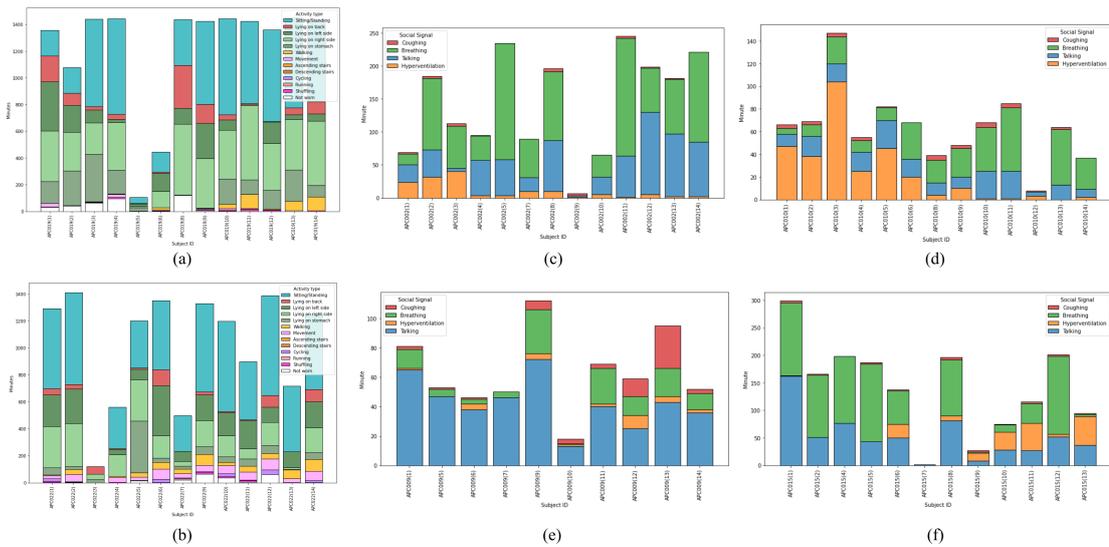


Figure 4.11: Apollo-C single subject visualizations (activity type classification and social signals distribution)

can be seen that the breathing rate of subjects in the top exposure group fluctuates more throughout the visits and does not show a visible decline like the breathing trend of the subjects in the bottom exposure group.

Similar results can be seen in Figure 4.8(a), which shows the effect of air pollution on the breathing rate at rest. Here, the breathing rate at rest is the average breathing rate taken from 2 am to 4 am, where we assume that most subjects are sleeping. Similar to previous plots, we see that the breathing rate of the subjects in the top exposure group stagnates and does decline as time progresses and even increases at times. On the other hand, subjects in the low exposure group show a visible decrease in the breathing rate at rest. This corresponds with results from previous studies showing that air pollution worsens COVID symptoms [101] [6]. Figure 4.8(b) shows the impact of activity level

on the breathing rate. From this graph, we are able to see that subjects with lower activity levels (in the bottom activity level group), on average, have a better recovery as their breathing rate decreases visibly.

To see further if the subjects recovered, a comparison of the correlation heatmap at visit 1 and visit 14 is generated, which can be seen in Figure 4.10. In visit 1, we see a positive correlation of 0.12 between breathing rate and activity level. In visit 14, this value increases to 0.29. This is because, in visit 1, the breathing rate is consistently high regardless of whether the activity level is high or low. This is also demonstrated in the previous breathing rate trend plots (when the subject is at rest). However, by visit 14, as their recovery progresses, there is a higher positive correlation which means that their breathing rate at rest is lower and increases normally with higher activity level. The positive correlation between PM variables and breathing rate decreased in visit 14. This would mean that exposure to air pollution now has a smaller impact on the breathing rate of the patients as their recovery progresses. Initially, most subjects tend to have higher breathing rates and are less active as one of the symptoms of long COVID is fatigue and shortness of breath and will often struggle with engaging in physical or dynamic activities [47][100]. This is also reflected in the plots from Figure 4.9. In terms of coughing episodes and activity level (Figure 4.9(a)), in the earlier visits, subjects experienced a larger number of coughing episodes on average even though having very low activity levels. However, as time passes, the average activity level increases, and the number of coughing episodes decreases.

Figure 4.11 presents some examples of subjects whose condition have improved over time (seen in Figures 4.11(a)-(b)) and subjects whose condition have worsened (seen in Figures 4.11(e)-(f)). For subjects whose condition have improved, we can see that as their recovery progresses, they become more active (by walking more or engaging in more dynamic activities), which can be seen in Figure 4.9(c), or have less coughing and hyperventilation episodes as compared to previous visits. However, for subjects whose condition did not improve, we see more coughing or hyperventilation episodes in their later visits than in their first visits or a consistent number of coughing/hyperventilation episodes throughout.

4.4 Dublin Analysis

In the Dublin dataset, we wish to see the effect of second hand smoking (SHS) on asthmatic and COPD subjects. Additional personal subject data such as their smoking

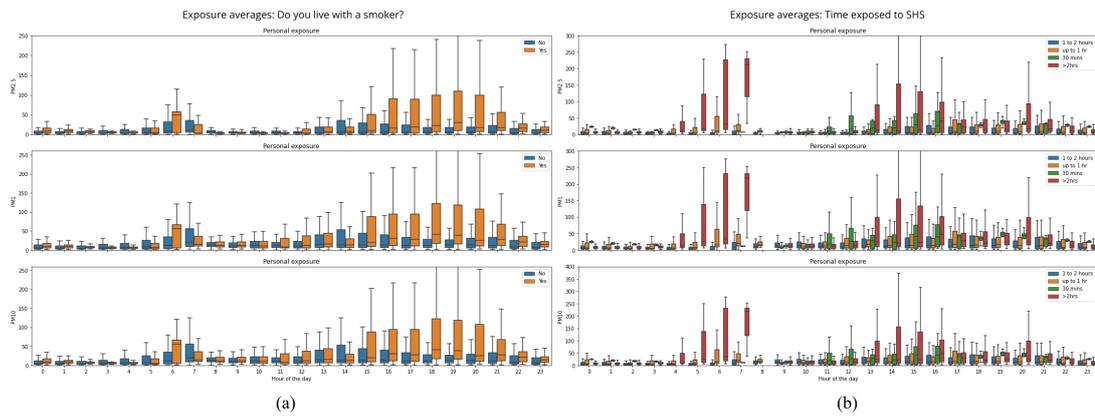


Figure 4.12: Dublin average exposure levels by hour of the day (a) split by "Do you live with a smoker?" (b) split by "time exposed to SHS"

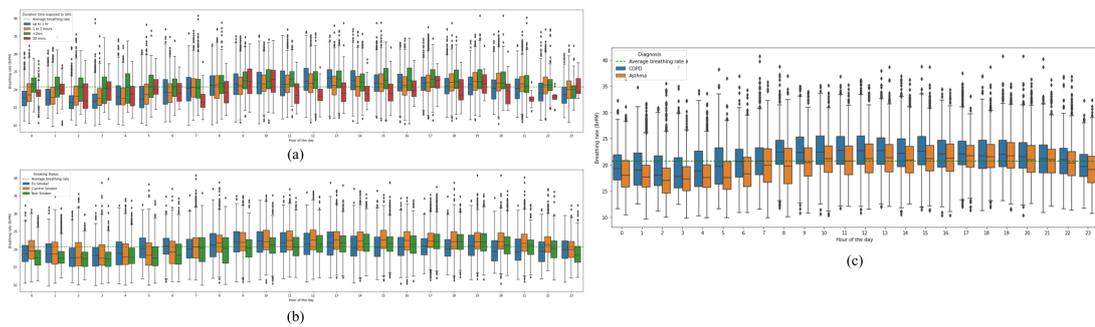


Figure 4.13: Dublin average Br by hour of day (a) split by duration of time exposed to SHS (b) split by smoking status (c) split by diagnosis (asthma or COPD)

status, whether they live with a smoker, and the amount of time exposed to SHS was also recorded. Figure 4.12(a) shows the comparison of exposure to PM_{2.5} levels between subjects that live and do not live with a smoker, and Figure 4.12(b) comparison of exposure levels depending on how much time they were exposed to SHS. From these graphs, as expected, it is evident that subjects who live with a smoker experience much larger levels of PM_{2.5} as compared to subjects that do not, especially during the daytime. Figure 4.13 shows plots for breathing rate averages by hour of day split by duration of time of exposed to SHS (Figure 4.13(a)), smoking status (4.13(b)), and diagnosis of whether they are asthmatic or have COPD (4.13(c)). From this, it can be seen that COPD subjects generally have higher breathing rates than asthmatics subjects. On the other hand, asthmatics subjects experience more coughing episodes as compared to COPD patients. Subjects exposed to over two hours of SHS during the day have the highest breathing rates. As for smoking status, subjects with the highest breathing rates are current smokers; however, there is not much of a noticeable difference in breathing

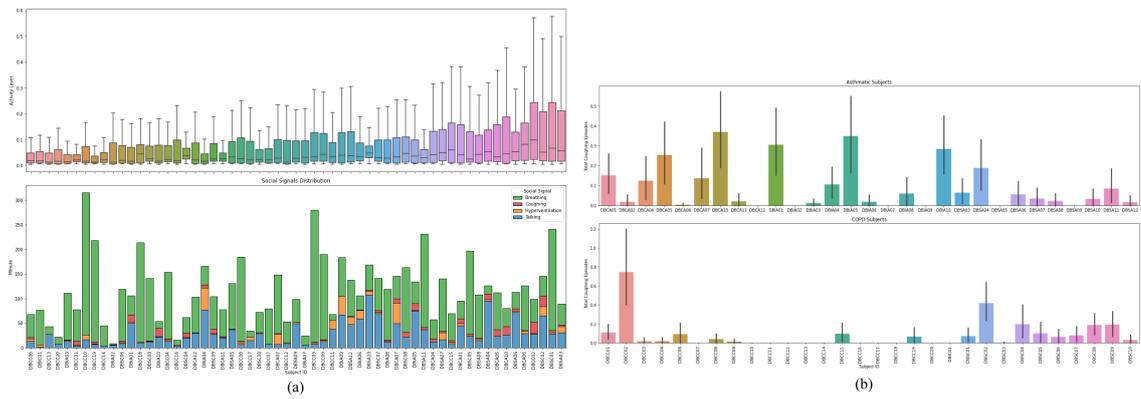


Figure 4.14: Dublin Social signals plots (a) Social signals distribution plot ordered by activity level (b) Average number of coughing episodes per minute between asthmatic subjects and COPD subjects

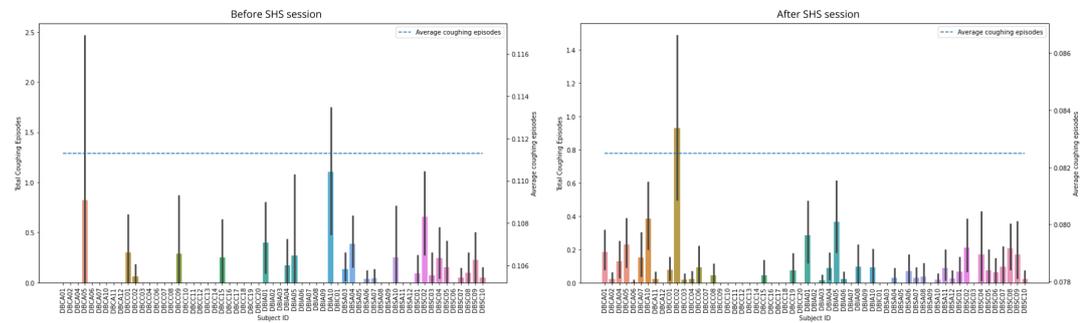


Figure 4.15: Dublin average number of coughing episodes per minute before and after exposure to SHS

rates between ex-smokers and non-smokers.

The distribution of social signals ordered by activity level in ascending order is shown in Figure 4.14(a). From this, we can see that subjects with higher activity levels tend to cough more. Figure 4.14(b) shows the comparison of the number of coughing episodes between asthmatic and COPD subjects. From this, it can be seen that asthmatic subjects experience more coughing episodes compared to COPD subjects. In order to see the impact of SHS on the subjects, Figure 4.15 shows two plots comparing the average number of coughing episodes for each subject before and after the SHS session took place. From the figure, we can see that many subjects did not cough before the SHS session, but coughed after they were exposed to SHS. Before the SHS session, 24 out of 60 subjects coughed at least once. After the SHS session, 39 out of 40 people coughed at least once. This shows that SHS does negatively impact asthmatic and COPD subjects.

4.5 Daphne MCC Analysis

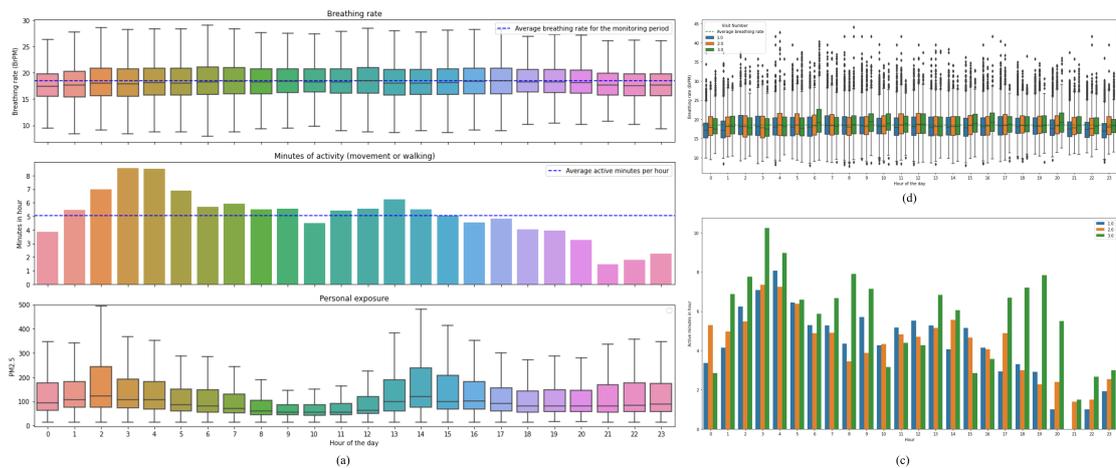


Figure 4.16: Daphne-MCC hourly activity, breathing rate and exposure plots (a) Hourly averages for breathing rate, minutes of activity, and exposure (b) comparison of average breathing rates split by trimester (c) comparison of number of active minutes split by trimester

From the results in Figure 4.16, the average breathing rate throughout the monitoring period is around 19 breaths per minute. In another study, it is reported that pregnant women have an average respiratory rate of 15 to 21 breaths per minute [29], which aligns well with the results shown in the hourly averages breathing rate plot in Figure 4.16(a). In 4.16(b), we notice that the breathing rate of subjects is highest in Visit 3, followed by Visit 1, then Visit 2. Multiple sources have suggested that pregnant women can experience shortness of breath due to increased levels of progesterone[2]. As the pregnancy progresses and the baby in the stomach grows larger, the organs are often squeezed, and the diaphragm is compressed. This may cause the pregnant woman's lungs to not have enough room to expand for a full breath. As a result, this could cause pregnant women to take more breaths [2][93]. Despite being healthy subjects with no respiratory afflictions, we do not see much difference between the breathing rate at rest (during night time) and the breathing rate during the day. If we compare this to the breathing rate trends for APCAPS, we will see a visible decrease in breathing rate from daytime to nighttime. Again, this could be due to their pregnancy causing difficulty breathing, which resulted in a higher breathing rate even when at rest [65][40]. In Figure 4.16(b), it can be observed that subjects in visit 3 are generally more active when compared with the previous visits (followed by visit 2 and then visit 1). In several

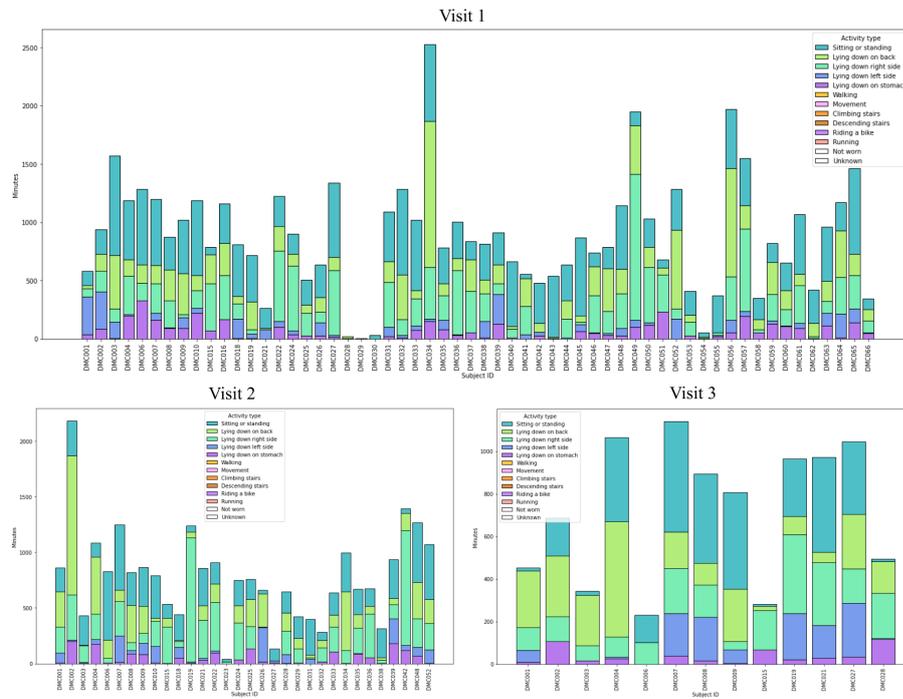


Figure 4.17: Daphne-MCC sleeping positions across visits

studies [10] [37][72][61][54], it is shown that women actually become more active in their third trimester due to the ‘nesting’ instinct. This occurs during late pregnancy, and women become more active and carry out different chores to prepare the home for the expected baby. Meanwhile, some studies show that women become less active as their pregnancy progresses[97][79]. However, we do not know if this is the case for the subjects in this dataset as there is no RESpeck data from these subjects before they were pregnant; therefore, we do not know what their normal breathing rate was. Additionally, compared to other datasets, subjects in Daphne MCC are quite active at night as well. In several studies, including questionnaire data, many pregnant women have reported discomfort while sleeping, inability to sleep, and a disrupted or disturbed sleeping schedule [104][43].

Figure 4.17 shows plots of the sleeping positions of pregnant women. In all visits 1, 2, and 3, it can be seen from the graphs that women opt to either lie on their backs or their right side. Though in visit 1, we can still see a large duration of time that women spent lying on their stomachs. However, as the pregnancy progresses, women lie on their stomachs less and less. This is to be expected because as the pregnancy progresses, the stomach will grow. Though it is generally safe for pregnant women to lie on their stomachs, it can be quite uncomfortable and cause neck and back pain

[76]. It is generally recommended that women sleep on their left. Though either side is considered safe, the left is usually recommended by experts. In the first and second trimesters, most women lie on their right side more than they do on their left. However, in the third trimester, more pregnant women opt to lie on their left side compared to the first trimester. The subjects also lie on their backs less in the third trimester. This could be due to the fact that as pregnancy progresses, lying on their backs can also become quite uncomfortable. In one study, it is reported that many women in their third trimester spend a significant amount of time lying on their backs and that lying on their backs is associated with more respiratory events[35]. It is also recommended that pregnant women do not lie on their backs as it could lead to stillbirth or other birth-related complications [27][95]; however, some studies disagree with this and state that pregnant women lying on their back does not lead to stillbirth [1][83]. Regardless of the suggestions given by experts on sleeping positions while pregnant, there are still many conflicting studies when it comes to the actual sleeping position adopted by pregnant women [78][96][68].

4.6 NHS Borders Analysis

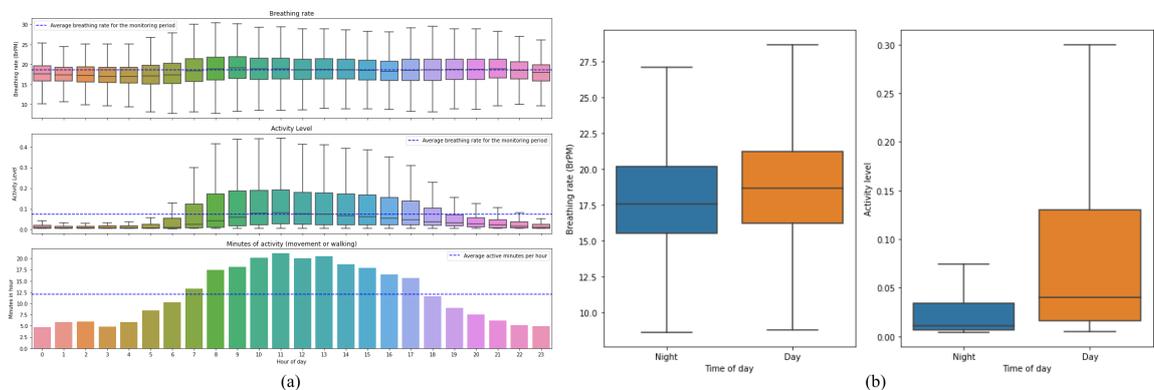


Figure 4.18: NHS averages for breathing rate, activity level and active minutes (a) Hourly averages for Br, activity level, and active minutes in an hour (b) NHS day VS night average activity level and breathing rate

The NHS Borders dataset consists of elderly COPD subjects belonging to three categories: exacerbation recovery (for subjects who have had an exacerbation recently), anxious COPD, and virtual PR. Subjects were invited to wear the Respeck sensor at all times for monitoring purposes and also perform pulmonary rehabilitation exercises

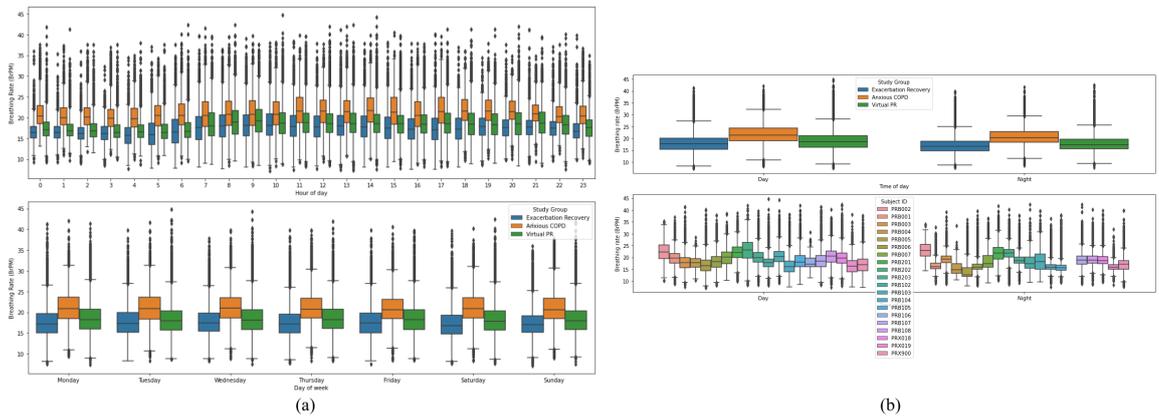


Figure 4.19: NHS Brs comparison between cohorts plots

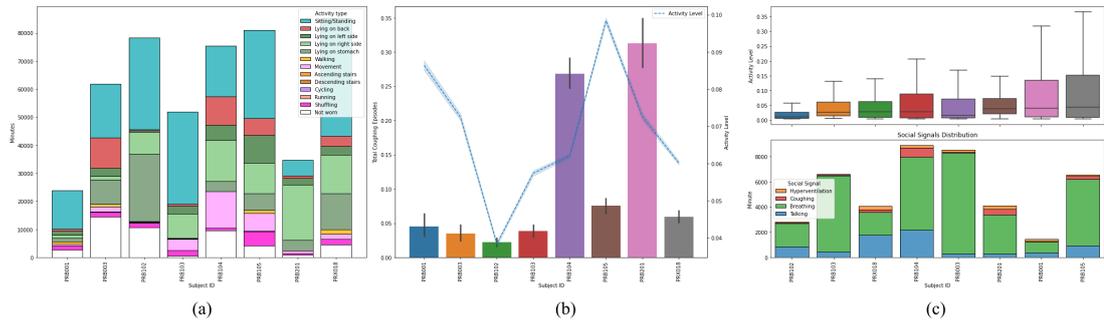


Figure 4.20: NHS HAR and social signals results plots (a) activity type distribution (b) average number of coughing episodes per minute with activity level (c) social signals distribution ordered by activity level

regularly, preferably once a week. Figure 4.18(a) shows the hourly averages for breathing rate, activity level, and the number of active minutes in an hour. From this, it can be seen that the subjects are mainly active during the period between 7 am and 6 pm. Compared to other datasets such as INHALE, PEEPS, and Apcaps, whose hourly averages plots can be found in Appendix *fill in later*, NHS Border subjects are active for a shorter period of time. This is to be expected as they are older subjects and are likely to spend more time sitting or lying down. Figure 4.18(b) compares the breathing rate and activity level at night and during the daytime. We can see a drastic difference in the activity level during nighttime and the daytime. Subjects have very low activity levels at night. However, their breathing rate is quite high- almost as high as their breathing rate during the daytime when they are much more active. This is expected as they are older COPD patients and are likely to be in their later stages of COPD, where symptoms occur even when at rest [102]. From the breathing rate comparison plots between cohorts (exacerbation recovery, anxious COPD, and virtual PR) in Figure 4.19,

it can be seen that COPD patients in the anxious COPD group have a much higher average breathing rate as compared to patients in the exacerbation recovery and virtual PR groups. Findings from prior literature have stated that COPD patients with anxiety will often experience shortness of breath [99][25][16] and increased respiratory rate [87]. Additionally, from the results of the activity classification in Figure 4.20(a), many subjects are seen to shuffle walk. This is expected as shuffle walking is often seen in older people. Additionally, a study that investigated walking abnormalities in COPD subjects found a strong correlation between COPD and the presence of abnormal gait (especially shuffle walking)[103].

4.7 Apcaps, PEEPS, QIP, and Leon Analysis

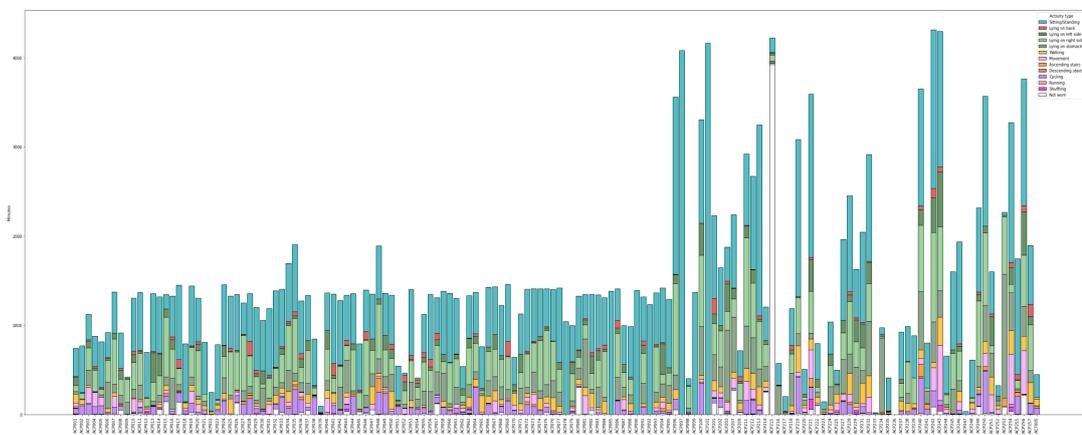


Figure 4.21: Activity type classification distribution across subjects for Apcaps dataset

Figure 4.21 - 4.24 shows the activity type classifications for Apcaps, PEEPS, QIP, and LEON, respectively. In the Apcaps dataset, it can be seen that the subjects spend more time engaging in more dynamic or vigorous physical activities such as running, cycling, ascending stairs, descending stairs, and walking for longer periods of time. Although in the case of PEEPS dataset, the subjects spend more time sitting or standing, and less time walking given that they are mainly office workers.

QIP subjects had spent the majority of their time lying down or sitting and standing with some short durations of miscellaneous movement. However, they rarely ever walked, which is likely to be accurate classifications as QIP subjects are post-operative patients in a hospital ward, and are expected to be lying in bed during their recovery. Some of these patients had back surgery, while others had more simple procedures.

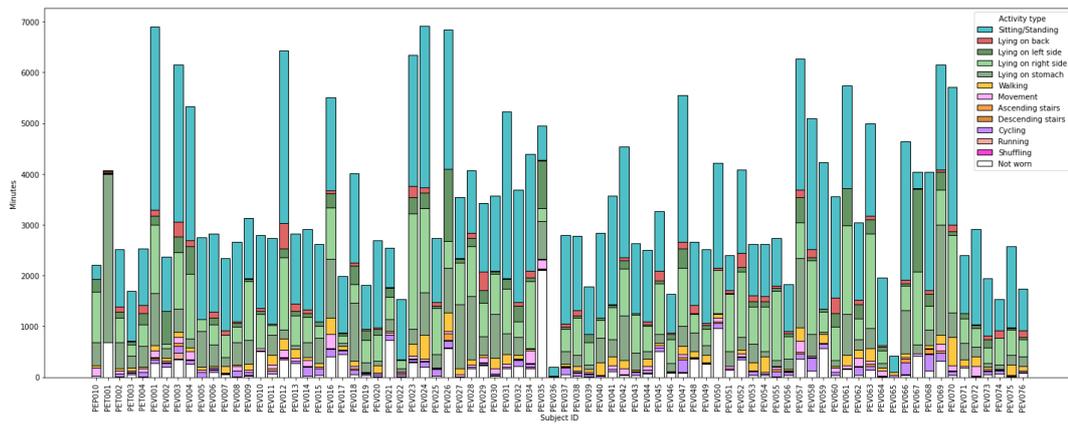


Figure 4.22: Activity type classification distribution across subjects for PEEPS dataset in minutes

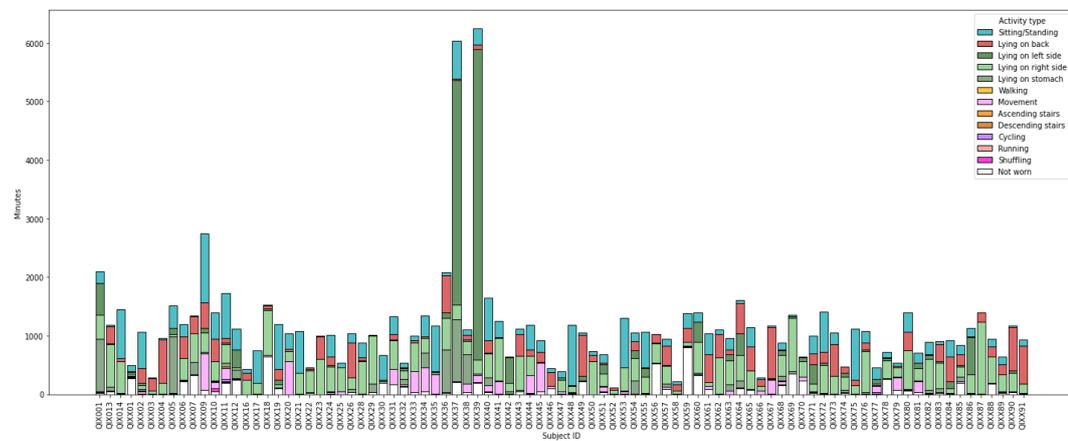


Figure 4.23: Activity type classification distribution across subjects for QIP dataset in minutes

From the QIP HAR classification results in Figure 4.23, it can be noticed that some subjects do not lie on their back at all, which can be explained by the nature of their surgery.

The Leon dataset, as illustrated In Figure 4.24, details the breakdown of activity types for each hour. It is clear that the three volunteers (MXX101, MXX102, MXX103) either walked (yellow) or cycled (purple) the same three routes every weekday for one month. This is clearly reflected in MXX102, and MXX103, who were either cycling or walking between 6 am to 11 am. However, in the case of MXX101, the activity classification results show that they were mostly walking between 2 pm to 4 pm. For the remaining hours for all subjects, the results showed that the subject was mainly 'lying on their stomach.' This is likely due to the RESpeck being stored, still collecting

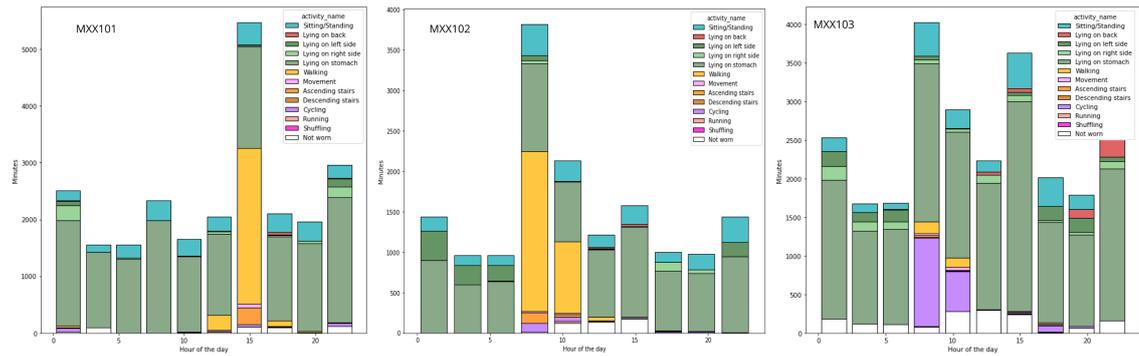


Figure 4.24: Activity type classification distribution by hour for Leon dataset in minutes

data, but not worn.

4.8 Evaluation and Discussion

4.8.1 Performance of models on real-world datasets

Overall, the RandomForest HAR and the step-counter model do classify different activity types accurately. For instance, Figure 4.25 shows a detailed single subject plot of subject INH102's step data obtained from the hierarchical step counter model along with the classified activities and activity level during when that specific activity was performed. From this plot, one can see the line representing cumulative steps increasing every time during 'walking' classifications and stagnating when the activity is 'lying down' or 'sitting/standing.' This implies that both the HAR model and step counter are accurate. Additionally, we can also compare the activity level recorded by the RESpeck device to the classification result. We can see that during physical activities such as 'walking,' 'cycling,' or 'misc. movement', the activity level or intensity is much higher than during 'sitting/standing' or 'lying down' events.

As for the 1-D CNN social signals classification models, we see that most coughs are detected accurately, as it has to be verified by the three models' results. Additionally, the number of coughs detected in each subject/dataset corresponds to the known characteristics of the subject/dataset. Examples of this can be seen in Figure 4.6 and Figure A.8, where one can observe subjects' coughing in response to high exposure to air pollution. However, it is still difficult to verify the 'hyperventilation' classifications, and some of them are likely to be misclassifications, as the healthy INHALE subjects

exhibit high levels of unlikely hyperventilation episodes. The possible cause of this will be discussed in more detail in the next section.

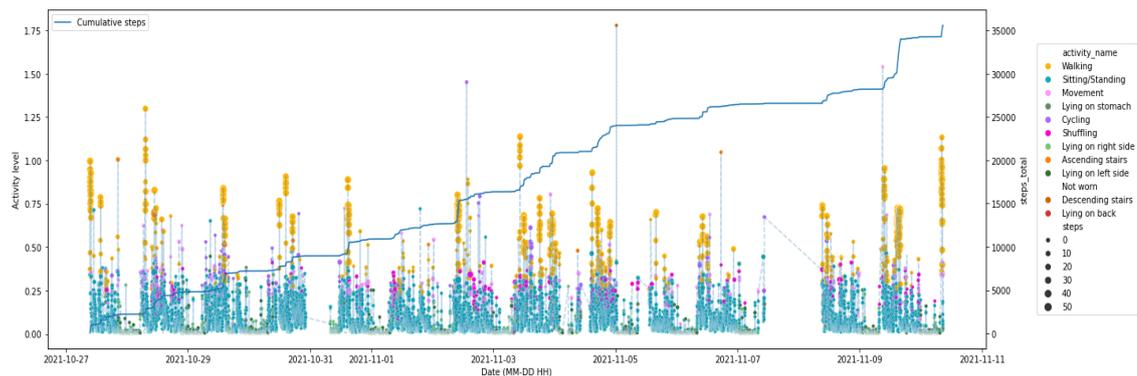


Figure 4.25: INH102 aggregated activity and step data

4.8.2 Limitations of the study

There are four main limitations of this study. Firstly, the social signals classification models used to detect coughing and hyperventilation episodes can only be applied to stationary activity. This means that some coughs that occur while the subject is in the middle of performing a dynamic activity may go unnoticed. Secondly, the classifications for hyperventilation episodes may not be accurate when applied to real-world datasets. This is because the social signals classification model was trained on labeled data collected from volunteers. Volunteers were asked to 'fake hyperventilate,' which could be misrepresentations of what actual hyperventilation episodes look like. Thirdly, the original HAR classification result timestamps were in nanoseconds and were then re-sampled into 1-minute data. The most commonly occurred activity type was applied for the re-sampled minute. In general, this should not be a problem, although the time spent for each activity type now becomes less detailed as some were performed for very short durations (mostly transition activities such as 'getting up from lying down') of time could be missed. Finally, since the study relies on data collected from real-world subjects, it is difficult to control when the subjects wear the RESpeck device. They might not be wearing the Respeck device at all, which leads to large gaps in the dataset, causing the collected data to be unusable. This is reflected most clearly in the PHILAP dataset, where several subjects have worn the RESpeck for only a few hours (shown in Figure A.9) when they were invited to wear the device for 48 hours. However, this is a natural consequence of analyzing real-world datasets.

Chapter 5

Conclusion and Future Works

5.1 Summary

In this project, three different existing machine learning methods (AC-GAN HAR model, 1-D CNN social signals classification models, and the step counter model) were applied to eleven real-world RESpeck datasets (INHALE, Apollo-C, Dublin, DAPHNE MCC, DAPHNE AAP, PHILAP, NHS Borders, APCAPS, PEEPS, QIP, and Leon) consisting of subjects with different levels of activity and types of morbidity. Our results have shown the performance of each model. The step counter performed well in detecting the number of steps. The social signals detected coughs well but struggled with classifying hyperventilation, most likely due to the 'hyperventilation' class collected by fake hyperventilating. Lastly, the AC-GAN failed to generalize on real-world datasets resulting in unrealistic numbers of 'cycling' and 'running' predictions. To address the issues of AC-GAN's misclassification and slow predicting times, an alternative HAR machine learning method that uses a Random Forest classifier and manually extracted features was applied to the datasets, greatly reducing misclassifications.

The results of applying the models for classification of physical activity and social signals to the Respeck and Airspeck datasets were visualised, analysed, interpreted, and compared at a meta-level. The accuracy of the classifications on the unlabelled datasets has been validated thanks to the variety in the datasets in terms of the subjects (health, morbidity of different types), location (urban, rural), and occupation (office workers, students). We have found that overall, healthy subjects are generally more active than asthmatic subjects, long COVID patients, and COPD subjects. Other factors, such as age and residence area, also affect a subject's activity level. Asthmatic subjects tend to cough most, followed by long COVID patients and then by COPD patients. We have

also noticed that in almost all datasets, high exposure to PM particles can increase the number of coughing episodes in a subject or negatively impact the subject's condition. Some main conclusions made following our results are stated as follows.

In the INHALE dataset, our results have shown that asthmatic subjects have irregular and fluctuating breathing rate patterns compared to healthy subjects, cough more, and are less active. In Apollo-C, we see that generally, most long COVID patients gradually recover and their symptoms become milder. We also find that high exposure to PM can hinder their recovery. In the Dublin dataset, results showed that exposure to SHS negatively impacts subjects and can trigger asthma and COPD symptoms as subjects are seen to cough more after exposure to SHS. In the NHS Borders dataset, we see a minimal difference in the breathing rate at rest and the breathing rate while active despite the activity level being significantly lower while at rest. We also observed many shuffle-walking episodes, which could be attributed to both old age and side-effect of COPD. For Daphne-MCC, our results reflected the phenomenon known as 'nesting,' where pregnant women often become more active during their third trimester.

5.2 Future Works

There are many directions for future work for this project. Firstly, additional deep learning architectures combined with manual feature extraction could be explored. Secondly, the labelled dataset used for training and testing could be further improved for a more balanced distribution between classes, and the collected hyperventilation data should consist of actual hyperventilation episodes instead of simulated ones. The machine learning models can be further developed to classify a larger variety of activities and social signals. Furthermore, there exists additional recorded data such as the patient's lung function, COPD Assessment Test scores, standardised early warning scoring (SEWS) charts, information of doctor's visits, and symptoms diaries that have yet to be incorporated into the visualizations and analysis of the datasets. As many of these datasets are derived from ongoing projects, there is fresh data from new subjects that can augment the datasets and improve classifications. Lastly, as these datasets are time-series data, forecasting models could be implemented to predict asthma attacks or COPD exacerbations based on previous data combined with information from activity levels, breathing rate, and social signals leading up to the respiratory event.

Bibliography

- [1] Science Update: Sleeping position during early and mid pregnancy does not affect risk of complications, NIH-funded study suggests.
- [2] Shortness of Breath In Pregnancy. *Harvard Health Publishing*.
- [3] Vital Signs (Body Temperature, Pulse Rate, Respiration Rate, Blood Pressure), June 2022.
- [4] Hervé Abdi and Lynne J. Williams. Principal component analysis: Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4):433–459, July 2010.
- [5] A. Agusti, J. Hedner, J. M. Marin, F. Barbé, M. Cazzola, and S. Rennard. Night-time symptoms: a forgotten dimension of COPD. *European Respiratory Review*, 20(121):183–194, September 2011.
- [6] Nurshad Ali and Farjana Islam. The Effects of Air Pollution on COVID-19 Infection and Mortality—A Review on Recent Evidence. *Frontiers in Public Health*, 8, 2020.
- [7] Turke Althobaiti, Stamos Katsigiannis, and Naeem Ramzan. Triaxial Accelerometer-Based Falls and Activities of Daily Life Detection Using Machine Learning. *Sensors*, 20(13):3777, January 2020.
- [8] Justice Amoh and Kofi Odame. DeepCough: A deep convolutional neural network in a wearable cough detection system. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, October 2015.
- [9] Justice Amoh and Kofi Odame. Deepcough: A deep convolutional neural network in a wearable cough detection system. In *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, 2015.

- [10] Marla V. Anderson and M. D. Rutherford. Evidence of a nesting psychology during human pregnancy. *Evolution and Human Behavior*, 34(6):390–397, November 2013.
- [11] Ong Chin Ann and Lau Bee Theng. Human activity recognition: A review. In *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, pages 389–393, November 2014.
- [12] D. K. Arvind, C. A. Bates, D. J. Fischer, and J. Mann. A sensor data collection environment for clinical trials investigating health effects of airborne pollution. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pages 88–91, March 2018.
- [13] D. K. Arvind, Janek Mann, Andrew Bates, and Konstantin Kotsev. The airspeek family of static and mobile wireless air quality monitors. In *2016 Euromicro Conference on Digital System Design (DSD)*, pages 207–214, 2016.
- [14] Gérard Biau and Erwan Scornet. A random forest guided tour. *TEST*, 25(2):197–227, June 2016.
- [15] Agata Brajdic and Robert Harle. Walk detection and step counting on unconstrained smartphones. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, UbiComp '13, pages 225–234, New York, NY, USA, September 2013. Association for Computing Machinery.
- [16] Gretchen A. Brenes. Anxiety and Chronic Obstructive Pulmonary Disease: Prevalence, Impact, and Treatment. *Psychosomatic Medicine*, 65(6):963–970, November 2003.
- [17] Raymond Brueckner and Björn Schuler. Social signal classification using deep blstm recurrent neural networks. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4823–4827, May 2014. ISSN: 2379-190X.
- [18] Erhan Bulbul, Aydin Cetin, and Ibrahim Alper Dogru. Human activity recognition using smartphones. In *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pages 1–6, 2018.
- [19] Christopher Carlsten, Sundeep Salvi, Gary W. K. Wong, and Kian Fan Chung. Personal strategies to minimise effects of air pollution on respiratory health:

- advice for providers, patients and the public. *European Respiratory Journal*, 55(6), June 2020.
- [20] Passara Chanchotisation and Chanvichet Vong. Implementation of Machine Learning and Deep Learning Algorithms with Dimensionality Reduction Methods for Internet of Things Gait Analysis and Monitoring Systems. *Sensors and Materials*, 33(12):4245, December 2021.
- [21] Stylianos Charalampous. Human activity and social signal classification using auxiliary classifier generative adversarial network and transfer learning. MInf Project (Part 1) Report, School of Informatics, The University of Edinburgh, 2021.
- [22] Stylianos Charalampous. Transformer-based human activity recognition calibration system. MInf Project (Part 2) Report, School of Informatics, The University of Edinburgh, 2022.
- [23] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794, San Francisco California USA, August 2016. ACM.
- [24] Nurul Amin Choudhury, Soumen Moulik, and Diptendu Sinha Roy. Physique-Based Human Activity Recognition Using Ensemble Learning and Smartphone Sensors. *IEEE Sensors Journal*, 21(15):16852–16860, August 2021.
- [25] J. A Cleland, A. J Lee, and S. Hall. Associations of depression and anxiety with gender, age, health-related quality of life and symptoms in primary care COPD patients. *Family Practice*, 24(3):217–223, May 2007.
- [26] Marco Aurélio de Valois Correia Junior, Emília Chagas Costa, Laienne Carla Barbosa de Barros, Andressa Araújo Soares, Emanuel Sávio Cavalcanti Sarinho, José Angelo Rizzo, and Silvia Wanick Sarinho. PHYSICAL ACTIVITY LEVEL IN ASTHMATIC ADOLESCENTS: CROSS-SECTIONAL POPULATION-BASED STUDY. *Revista Paulista De Pediatria: Orgao Oficial Da Sociedade De Pediatria De Sao Paulo*, 37(2):188–193, June 2019.
- [27] Robin S. Cronin, Minglan Li, John M. D. Thompson, Adrienne Gordon, Camille H. Raynes-Greenow, Alexander E. P. Heazell, Tomasina Stacey, Vicki M.

- Culling, Victoria Bowring, Ngaire H. Anderson, Louise M. O'Brien, Edwin A. Mitchell, Lisa M. Askie, and Lesley M. E. McCowan. An Individual Participant Data Meta-analysis of Maternal Going-to-Sleep Position, Interactions with Fetal Vulnerability, and the Risk of Late Stillbirth. *eClinicalMedicine*, 10:49–57, April 2019.
- [28] Akbar Dehghani, Omid Sarbishei, Tristan Glatard, and Emad Shihab. A Quantitative Comparison of Overlapping and Non-Overlapping Sliding Windows for Human Activity Recognition Using Inertial Sensors. *Sensors (Basel, Switzerland)*, 19(22):5026, November 2019.
- [29] A. T. Dennis and L. Hardy. Defining a Reference Range for Vital Signs in Healthy Term Pregnant Women Undergoing Caesarean Section. *Anaesthesia and Intensive Care*, 44(6):752–757, November 2016.
- [30] Christine Dewi and Rung-Ching Chen. Human activity recognition based on evolution of features selection and random forest. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 2496–2501, 2019.
- [31] Rahul Dey and Fathi M. Salem. Gate-variants of Gated Recurrent Unit (GRU) neural networks. In *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pages 1597–1600, August 2017. ISSN: 1558-3899.
- [32] Celina Dong. A comparative study in the classification of social signals using the wearable RESpeck Sensor. MInf Project (Part 1) Report, School of Informatics, The University of Edinburgh, 2021.
- [33] Celina Dong. Social signal classification using the wearable RESpeck Sensor. MInf Project (Part 2) Report, School of Informatics, The University of Edinburgh, 2022.
- [34] Gordon B. Drummond, Darius Fischer, Margaret Lees, Andrew Bates, Janek Mann, and D. K. Arvind. Classifying signals from a wearable accelerometer device to measure respiratory rate. *ERJ Open Research*, 7(2), April 2021.
- [35] Galit Levi Duniets, Orna Sever, Ari DeRowe, and Riva Tauman. Sleep position and breathing in late pregnancy and perinatal outcomes. *Journal of clinical*

- sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, 16(6):955–959, June 2020.
- [36] Andreas Ejupi and Carlo Menon. Detection of Talking in Respiratory Signals: A Feasibility Study Using Machine Learning and Wearable Textile-Based Sensors. *Sensors*, 18(8):2474, August 2018.
- [37] Flavia Faccio, Eleonora Mascheroni, Chiara Ionio, Gabriella Pravettoni, Fedro Alessandro Peccatori, Camilla Pisoni, Chiara Cassani, Sara Zambelli, Anna Zilioli, Giuseppe Nastasi, Nicola Giuntini, and Lucia Bonassi. Motherhood during or after breast cancer diagnosis: A qualitative study. *European Journal of Cancer Care*, 29(2), March 2020.
- [38] Wilton W. T. Fok, Louis C. W. Chan, and Carol Chen. Artificial Intelligence for Sport Actions and Performance Analysis using Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM). In *Proceedings of the 2018 4th International Conference on Robotics and Artificial Intelligence, ICRAI 2018*, pages 40–44, New York, NY, USA, November 2018. Association for Computing Machinery.
- [39] Teodora Georgescu. Classification of coughs using the wearable respeck monitor. UG Dissertation, School of Informatics, The University of Edinburgh, 2019.
- [40] Sorel Goland, Sharon Perelman, Nardin Asalih, Sara Shimoni, Osnat Walfish, Mordechai Hallak, Zion Hagay, Jacob George, Avraham Shotan, and David S. Blondheim. Shortness of Breath During Pregnancy: Could a Cardiac Factor Be Involved?: SOB in pregnancy. *Clinical Cardiology*, 38(10):598–603, October 2015.
- [41] Pragun Goyal, Vinay J. Ribeiro, Huzur Saran, and Anshul Kumar. Strap-down pedestrian dead-reckoning system. In *2011 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–7, 2011.
- [42] Elena Grigorieva and Artem Lukyanets. Combined Effect of Hot Weather and Outdoor Air Pollution on Respiratory Health: Literature Review. *Atmosphere*, 12(6):790, June 2021.
- [43] C Hedman, T Pohjasvaara, U Tolonen, A. S Suhonen-Malm, and V. V Myllylä. Effects of pregnancy on mothers’ sleep. *Sleep Medicine*, 3(1):37–42, January 2002.

- [44] Tan Chee Hian, Zainal Fikiri Mahmud, and Tham Yin Choong. Physical Fitness Level between Urban and Rural Students-Case Study. *Procedia - Social and Behavioral Sciences*, 90:847–852, October 2013.
- [45] Narit Hnoohom, Sakorn Mekruksavanich, and Anuchit Jitpattanakul. Human Activity Recognition Using Triaxial Acceleration Data from Smartphone and Ensemble Learning. In *2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 408–412, December 2017.
- [46] Chengli Hou. A study on imu-based human activity recognition using deep learning and traditional machine learning. In *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, pages 225–234, 2020.
- [47] Helen Humphreys, Laura Kilby, Nik Kudiersky, and Robert Copeland. Long COVID and the role of physical activity: a qualitative study. *BMJ Open*, 11(3):e047632, March 2021.
- [48] Martin-Philipp Irsch. Semi-supervised human activity recognition with the wearable respeck sensor using gans MSc Thesis Report, School of Informatics, The University of Edinburgh, 2021.
- [49] Violet Jayamani, Vijayaprasad Gopichandran, Premila Lee, Greeda Alexander, Solomon Christopher, and Jasmin Helan Prasad. Diet and Physical Activity Among Women in Urban and Rural Areas in South India: A Community Based Comparative Survey. *Journal of Family Medicine and Primary Care*, 2(4):334–338, December 2013.
- [50] Himanshu Joshi, Ananya Verma, and Amrita Mishra. Classification of social signals using deep lstm-based recurrent neural networks. In *2020 International Conference on Signal Processing and Communications (SPCOM)*, pages 1–5, 2020.
- [51] Yu Ri Kang, Ji-Yoon Oh, Ji-Hyang Lee, Peter M. Small, Kian Fan Chung, and Woo-Jung Song. Long-COVID severe refractory cough: discussion of a case with 6-week longitudinal cough characterization. *Asia Pacific Allergy*, 12(2):e19, April 2022.
- [52] Justin J. Kavanagh and Hylton B. Menz. Accelerometry: a technique for quantifying movement patterns during walking. *Gait & Posture*, 28(1):1–15, July 2008.

- [53] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [54] Alexandra Ketterman, Anastasia Makhanova, Tania A. Reynolds, Charleen R. Case, James K. McNulty, Lisa A. Eckel, Larissa Nikonova, Heather A. Flynn, and Jon K. Maner. Prevalence and predictors of “nesting”: Solutions to adaptive challenges faced during pregnancy. *Evolution and Human Behavior*, 43(3):188–196, May 2022.
- [55] Wajahat H. Khan, Vahid Mohsenin, and Carolyn M. D’Ambrosio. Sleep in asthma. *Clinics in Chest Medicine*, 35(3):483–493, September 2014.
- [56] Prashant Kumar, Mukesh Khare, Roy M. Harrison, William J. Bloss, Alastair C. Lewis, Hugh Coe, and Lidia Morawska. New directions: Air pollution challenges for developing megacities like delhi. *Atmospheric Environment*, 122:657–661, 2015.
- [57] Takayuki Kurokawa, Tasuku Miura, Masaru Yamashita, Tomoya Sakai, and Shoichi Matsunaga. Emotion-Cluster Classification of Infant Cries Using Sparse Representation. In *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 1875–1878, Honolulu, HI, USA, November 2018. IEEE.
- [58] Jason E. Lang. The impact of exercise on asthma. *Current Opinion in Allergy and Clinical Immunology*, 19(2):118–125, April 2019.
- [59] Leyuan Liu, Jian He, Keyan Ren, Jonathan Lungu, Yibin Hou, and Ruihai Dong. An Information Gain-Based Model and an Attention-Based RNN for Wearable Human Activity Recognition. *Entropy*, 23(12):1635, December 2021.
- [60] Elisabeth Mahase. Covid-19: What do we know about “long covid”? *BMJ*, 370:m2815, July 2020.
- [61] Lisa Malich. [The Nest as Environment. A Historical Epistemology of the Nesting Instinct in Pregnancy]. *NTM*, 29(1):45–75, March 2021.
- [62] Ihssene Menhour, M’hamed Bilal Abidine, and Belkacem Fergani. A new framework using pca, lda and knn-svm to activity recognition based smartphone’s

- sensors. In *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)*, pages 1–5, 2018.
- [63] Ronald Mutegeki and Dong Seog Han. A cnn-lstm approach to human activity recognition. In *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, pages 362–366, 2020.
- [64] H. Nematallah and S. Rajan. Comparative study of time series-based human activity recognition using convolutional neural networks. In *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1–6, 2020.
- [65] Francesca Neuberger and Catherine Nelson-Piercy. Acute presentation of the pregnant patient. *Clinical Medicine*, 15(4):372–376, August 2015.
- [66] NHS. Long-term effects of coronavirus (long COVID), January 2021.
- [67] Nikita Nikolajev. Multiclass classification of social signals using the respect sensor. UG Dissertation, School of Informatics, The University of Edinburgh, 2020.
- [68] Louise M. O’Brien and Jane Warland. Typical sleep positions in pregnant women. *Early Human Development*, 90(6):315–317, June 2014.
- [69] Keiron O’Shea and Ryan Nash. An Introduction to Convolutional Neural Networks, December 2015. arXiv:1511.08458 [cs].
- [70] Romain A. Pauwels, A. Sonia Buist, Peter M. A. Calverley, Christine R. Jenkins, and Suzanne S. Hurd. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop Summary. *American Journal of Respiratory and Critical Care Medicine*, 163(5):1256–1276, April 2001.
- [71] Chalaywan Pinyochotiwong, Naricha Chirakalwasan, and Nancy Collop. Nocturnal Asthma. *Asian Pacific Journal of Allergy and Immunology*, 39(2):78–88, June 2021.
- [72] Melanie S. Poudevigne and Patrick J. O’Connor. A Review of Physical Activity Patterns in Pregnant Women and Their Relationship to Psychological Health. *Sports Medicine*, 36(1):19–38, January 2006.

- [73] M. Pratt, C. A. Macera, and C. Blanton. Levels of physical activity and inactivity in children and adults in the United States: current evidence and research issues. *Medicine and Science in Sports and Exercise*, 31(11 Suppl):S526–533, November 1999.
- [74] Sreenivasan Ramasamy Ramamurthy and Nirmalya Roy. Recent trends in machine learning for human activity recognition—A survey. *WIREs Data Mining and Knowledge Discovery*, 8(4), July 2018.
- [75] A. V. Raveendran, Rajeev Jayadevan, and S. Sashidharan. Long COVID: An overview. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 15(3):869–875, May 2021.
- [76] Bikash R. Ray and Anjan Trikha. Prone position ventilation in pregnancy: Concerns and evidence. *Journal of Obstetric Anaesthesia and Critical Care*, 8(1):7, January 2018.
- [77] Andrew I. Ritchie and Jadwiga A. Wedzicha. Definition, Causes, Pathogenesis, and Consequences of Chronic Obstructive Pulmonary Disease Exacerbations. *Clinics in Chest Medicine*, 41(3):421–438, September 2020.
- [78] Shania K. Rossiter, Samia Aziz, Alyce N. Wilson, Liz Comrie-Thomson, Tomasina Stacey, Caroline S. E. Homer, and Joshua P. Vogel. Women’s sleep position during pregnancy in low- and middle-income countries: a systematic review. *Reproductive Health*, 18(1):53, March 2021.
- [79] E. K. Rousham, P. E. Clarke, and H. Gross. Significant changes in physical activity among pregnant women in the UK as assessed by accelerometry and self-reported activity. *European Journal of Clinical Nutrition*, 60(3):393–400, March 2006.
- [80] Frank A. J. L. Scheer, Michael F. Hilton, Heather L. Evoniuk, Sally A. Shiels, Atul Malhotra, Rena Sugarbaker, R. Timothy Ayers, Elliot Israel, Anthony F. Massaro, and Steven A. Shea. The endogenous circadian system worsens asthma at night independent of sleep and other daily behavioral or environmental cycles. *Proceedings of the National Academy of Sciences of the United States of America*, 118(37):e2018486118, September 2021.

- [81] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In Wulfram Gerstner, Alain Germond, Martin Hasler, and Jean-Daniel Nicoud, editors, *Artificial Neural Networks — ICANN'97*, Lecture Notes in Computer Science, pages 583–588, Berlin, Heidelberg, 1997. Springer.
- [82] Shuai Shi. An end to end real-time step counting algorithm implementation using the wearable respeck monitor. UG Dissertation, School of Informatics, The University of Edinburgh, 2022.
- [83] Robert M. Silver, Shannon Hunter, Uma M. Reddy, Francesca Facco, Karen J. Gibbins, William A. Grobman, Brian M. Mercer, David M. Haas, Hyagriv N. Simhan, Samuel Parry, Ronald J. Wapner, Judette Louis, Judith M. Chung, Grace Pien, Frank P. Schubert, George R. Saade, Phyllis Zee, Susan Redline, Corette B. Parker, and for the Nulliparous Pregnancy Outcomes Study: Monitoring Mothers-to-Be (NuMoM2b) Study. Prospective Evaluation of Maternal Sleep Position Through 30 Weeks of Gestation and Adverse Pregnancy Outcomes. *Obstetrics & Gynecology*, 134(4):667–676, October 2019.
- [84] Joan B. Soriano, Inmaculada Alfageme, Pere Almagro, Ciro Casanova, Cristobal Esteban, Juan J. Soler-Cataluña, Juan P. de Torres, Pablo Martinez-Camblor, Marc Miravittles, Bartolome R. Celli, and Jose M. Marin. Distribution and Prognostic Validity of the New Global Initiative for Chronic Obstructive Lung Disease Grading Classification. *Chest*, 143(3):694–702, March 2013.
- [85] C. O. S. Sorzano, J. Vargas, and A. Pascual Montano. A survey of dimensionality reduction techniques, March 2014. arXiv:1403.2877 [cs, q-bio, stat].
- [86] Abhishek Tiwari, Salaar Liaqat, Daniyal Liaqat, Moshe Gabel, Eyal de Lara, and Tiago H. Falk. Remote copd severity and exacerbation detection using heart rate and activity data measured from a wearable device. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pages 7450–7454, 2021.
- [87] Athanasios Tselebis, Argyro Pachi, Ioannis Ilias, Epaminondas Kosmas, Dionisios Bratis, Georgios Moussas, and Nikolaos Tzanakis. Strategies to improve anxiety and depression in patients with COPD: a mental health perspective. *Neuropsychiatric Disease and Treatment*, 12:297–328, February 2016.

- [88] Ioanna Tsiligianni, Esther Metting, Thys van der Molen, Niels Chavannes, and Janwillem Kocks. Morning and night symptoms in primary care COPD patients: a cross-sectional and longitudinal study. An UNLOCK study from the IPCRG. *NPJ Primary Care Respiratory Medicine*, 26:16040, July 2016.
- [89] OAR US EPA. Particle Pollution and Respiratory Effects, September 2014.
- [90] Chanvichet Vong, Thitiphoom Theptit, Virinya Watcharakonpipat, Passara Chanchotisatien, and Seksan Laitrakun. Comparison of Feature Selection and Classification for Human Activity and Fall Recognition using Smartphone Sensors. In *2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering*, pages 170–173, March 2021.
- [91] Michalis Vrigkas, Christophoros Nikou, and Ioannis A. Kakadiaris. A Review of Human Activity Recognition Methods. *Frontiers in Robotics and AI*, 2, 2015.
- [92] Jayavrinda Vrindavanam, Raghunandan Srinath, Hari Haran Shankar, and Gaurav Nagesh. Machine learning based covid-19 cough classification models - a comparative analysis. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pages 420–426, 2021.
- [93] Arlene M. Wallace, Diane B. Boyer, Alice Dan, and Karyn Holm. Aerobic exercise, maternal self-esteem, and physical discomforts during pregnancy. *Journal of Nurse-Midwifery*, 31(6):255–262, November 1986.
- [94] Jian-Hua Wang, Jian-Jiun Ding, Yu Chen, and Hsin-Hui Chen. Real time accelerometer-based gait recognition using adaptive windowed wavelet transforms. In *2012 IEEE Asia Pacific Conference on Circuits and Systems*, pages 591–594, 2012.
- [95] Jane Warland. Back to basics: avoiding the supine position in pregnancy. *The Journal of Physiology*, 595(4):1017–1018, February 2017.
- [96] Jane Warland and Jillian Dorrian. Accuracy of self-reported sleep position in late pregnancy. *PloS One*, 9(12):e115760, 2014.
- [97] Patricia E. Watson, P. E. Watson, and Barry W. McDonald, B. W. McDonald. Activity levels in pregnant New Zealand women: relationship with socioeconomic factors,

- well-being, anthropometric measures, and birth outcome. *Applied Physiology, Nutrition, and Metabolism*, July 2007.
- [98] Juyang Weng, Yilu Zhang, and Wey-Shiuan Hwang. Candid covariance-free incremental principal component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):1034–1040, August 2003.
- [99] Thomas George Willgoss and Abebaw Mengistu Yohannes. Anxiety Disorders in Patients With COPD: A Systematic Review. *Respiratory Care*, 58(5):858–866, May 2013.
- [100] Jack Wright, Sarah L. Astill, and Manoj Sivan. The Relationship between Physical Activity and Long COVID: A Cross-Sectional Study. *International Journal of Environmental Research and Public Health*, 19(9):5093, April 2022.
- [101] Xiao Wu, Rachel C Nethery, M Benjamin Sabath, Danielle Braun, and Francesca Dominici. Exposure to air pollution and covid-19 mortality in the united states: A nationwide cross-sectional study. *medRxiv*, 2020.
- [102] Josef Yayan and Kurt Rasche. Asthma and COPD: Similarities and Differences in the Pathophysiology, Diagnosis and Therapy. *Advances in Experimental Medicine and Biology*, 910:31–38, 2016.
- [103] Jennifer M. Yentes, Harlan Sayles, Jane Meza, David M. Mannino, Stephen I. Rennard, and Nicholas Stergiou. Walking abnormalities are associated with COPD: An investigation of the NHANES III dataset. *Respiratory Medicine*, 105(1):80–87, January 2011.
- [104] Edward R. Yeomans and Larry C. Gilstrap. Physiologic changes in pregnancy and their impact on critical care:. *Critical Care Medicine*, 33(Supplement):S256–S258, October 2005.
- [105] Wojciech Zaremba, Ilya Sutskever, and Oriol Vinyals. Recurrent Neural Network Regularization, February 2015. arXiv:1409.2329 [cs].
- [106] Xudong Zhang. Classification of indoor/outdoor environment and modes of transportation based on airspeck/gps data MInf Part1 Thesis Report, School of Informatics, The University of Edinburgh, 2022.

- [107] Niya Zhou, Zhihong Cui, Sanming Yang, Xue Han, Gangcai Chen, Ziyuan Zhou, Chongzhi Zhai, Mingfu Ma, Lianbing Li, Min Cai, Yafei Li, Lin Ao, Weiqun Shu, Jinyi Liu, and Jia Cao. Air pollution and decreased semen quality: A comparative study of Chongqing urban and rural areas. *Environmental Pollution*, 187:145–152, April 2014.
- [108] Hui Zou, Trevor Hastie, and Robert Tibshirani. Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2):265–286, June 2006.

Appendix A

Additional Figures



Figure A.1: RESpeck sensor with ruler



Figure A.2: Airspeck sensor [106]

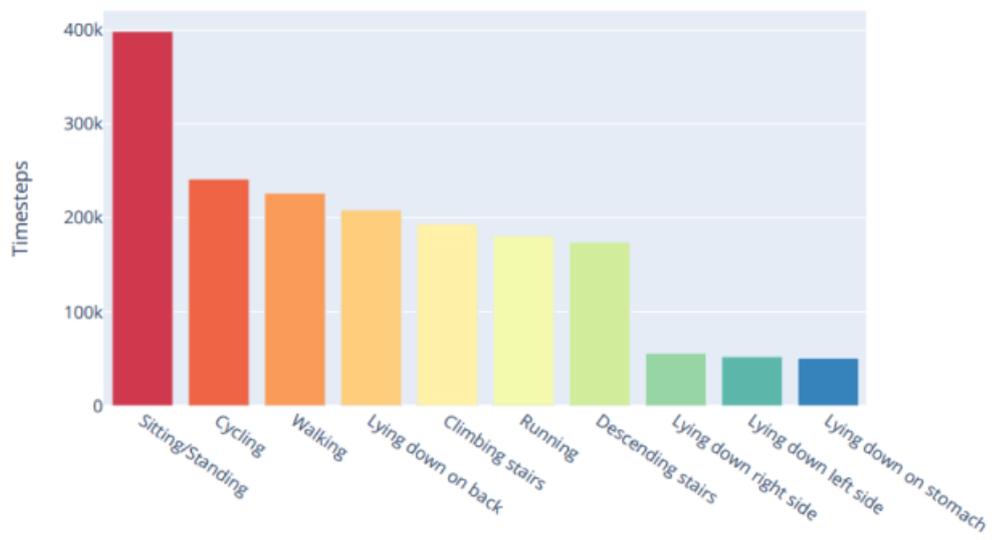


Figure A.3: Distribution of timesteps in each class in HAR labelled dataset [21]

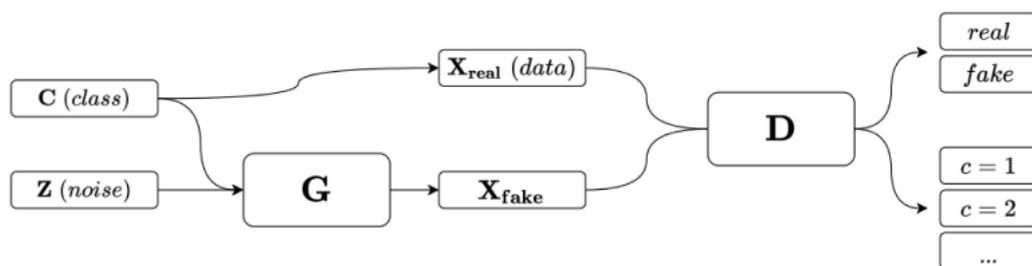


Figure A.4: AC-GAN architecture diagram [21]

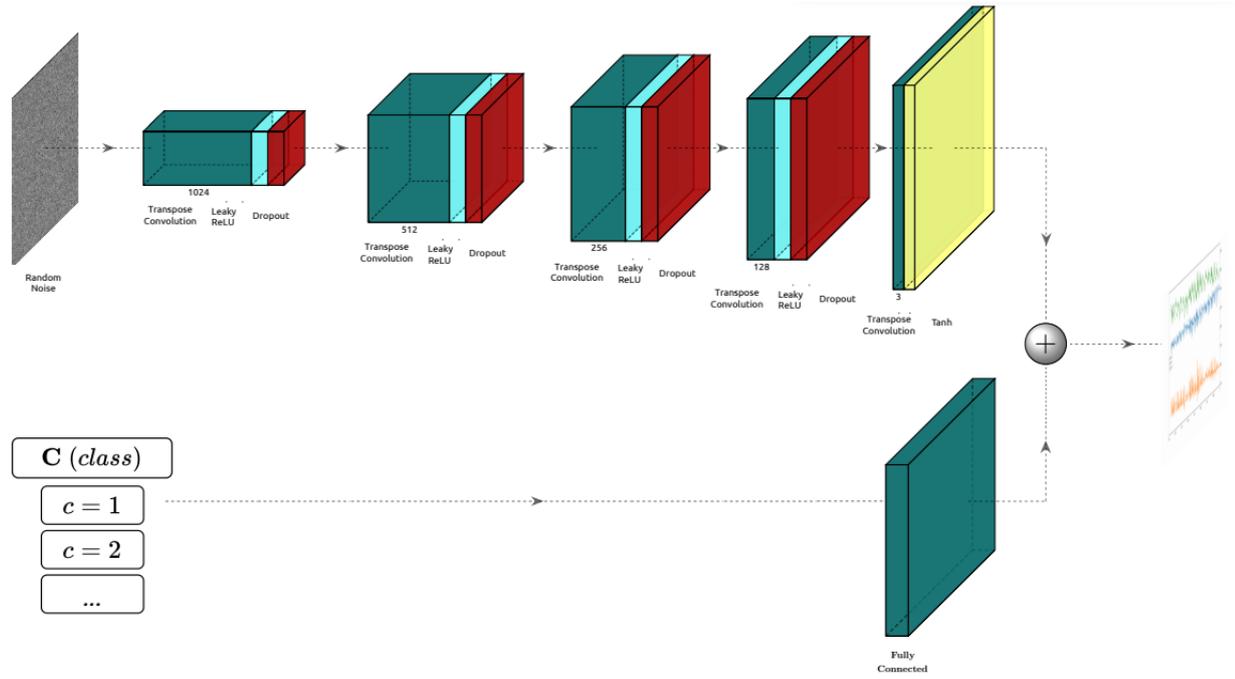


Figure A.5: AC-GAN generator diagram [21]

Classifier	Dimen Reduc.	Acc	F1	Prec	Recall
RF	-	0.9217	0.9212	0.9340	0.9114
XGBoost	-	0.8841	0.8788	0.8941	0.8680
LightGBM	-	0.9123	0.8979	0.9082	0.8900
RF	PCA	0.8764	0.8214	0.8880	0.7918
XGBoost	PCA	0.8611	0.8382	0.8694	0.8208
LightGBM	PCA	0.8965	0.8778	0.8952	0.8653
RF	SPCA	0.9210	0.9211	0.9330	0.9118
XGBoost	SPCA	0.8840	0.8781	0.8921	0.8682
LightGBM	SPCA	0.9149	0.8981	0.9035	0.8985
RF	IPCA	0.8770	0.8210	0.8950	0.7910
XGBoost	IPCA	0.8619	0.8366	0.8710	0.8171
LightGBM	IPCA	0.8968	0.8778	0.8944	0.8663

Table A.1: Accuracy, F1, precision, and recall scores for ensemble and dimensionality reduction methods

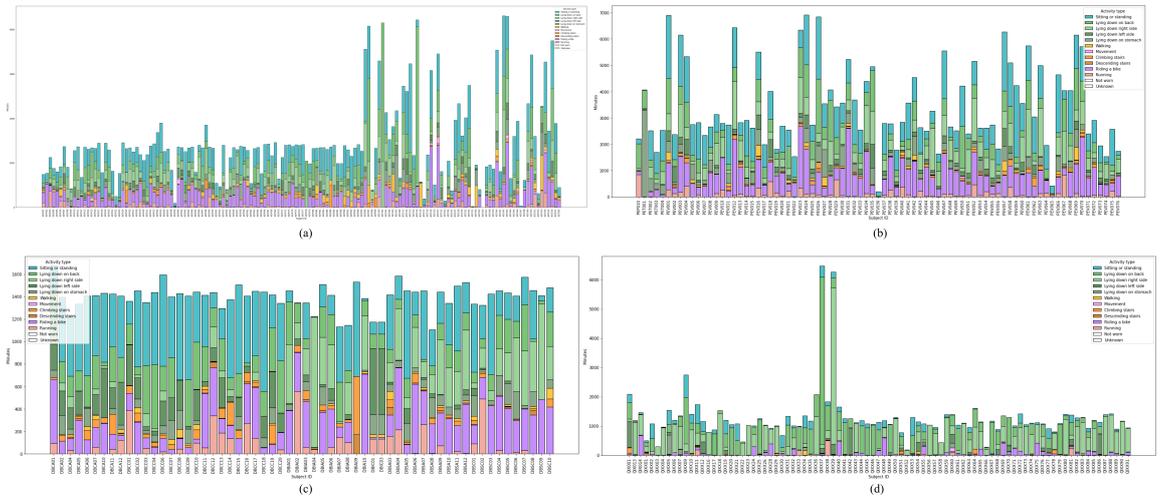


Figure A.6: Activity classification with AC-GAN (a) Apcaps (b) PEEPS (c) Dublin (d) QIP

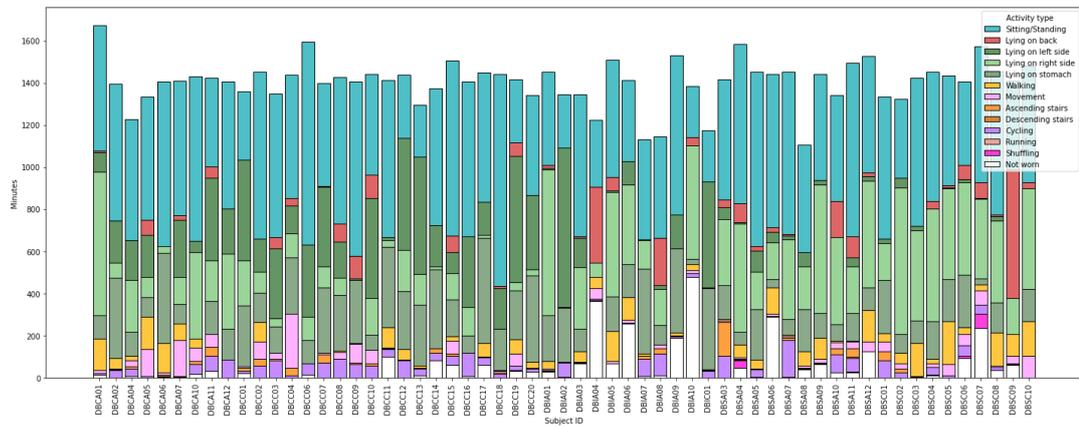


Figure A.7: Dublin dataset activity type classification for each subject in minutes with RF classifier

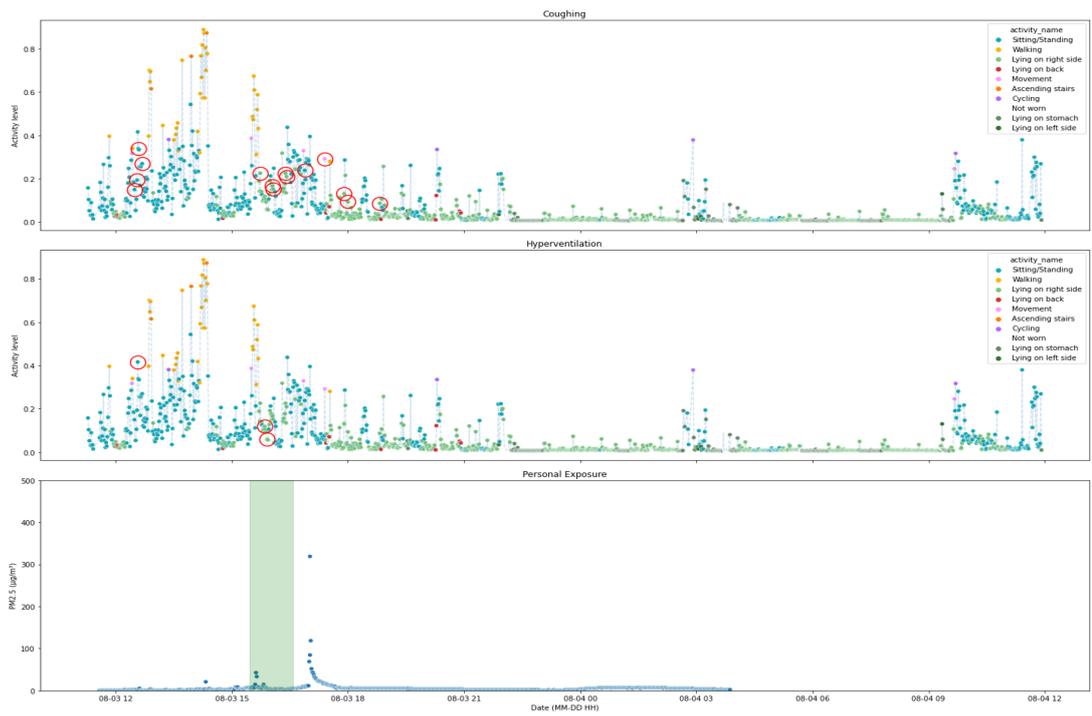


Figure A.8: Dublin dataset DBIA01 single subject visualization: combined social signals plot

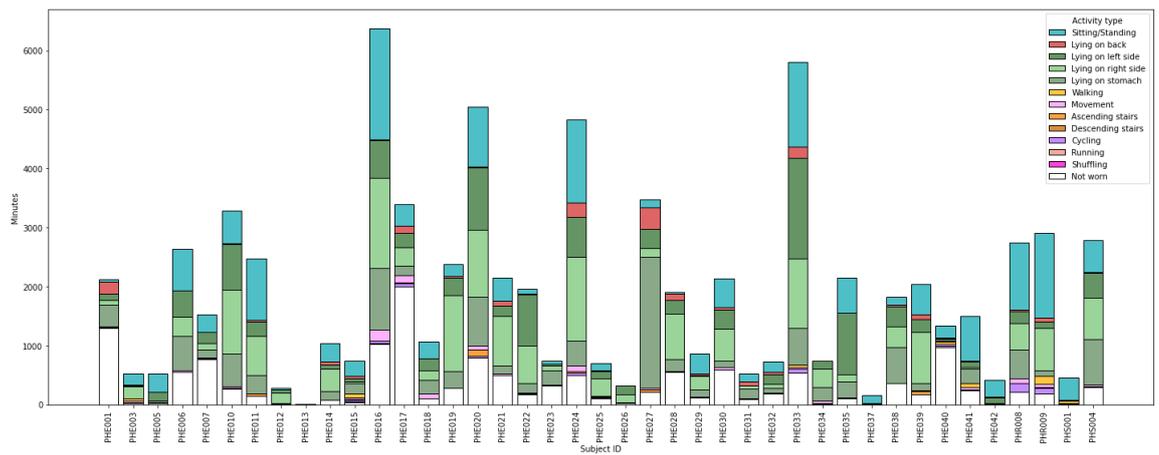


Figure A.9: Philap activity classifications