Universidade do Minho
Escola de Engenharia
Departamento de Informática

David José Teixeira de Sousa

Learning User Well-being and Comfort through Smart Devices

January 2021

Universidade do Minho
Escola de Engenharia
Departamento de Informática

David José Teixeira de Sousa

Learning User Well-being and Comfort through Smart Devices

Master dissertation
Integrated Master in Informatics Engineering

Dissertation supervised by
César Analide
Fábio Silva

January 2021

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

ACKNOWLEDGEMENTS

I would like to start by thanking the advisor and co-advisor for this dissertation. Professor César Analide and Professor Fábio Silva were crucial for the writing of a good dissertation, as well as for the realization of a good project. All the advice provided and all the availability shown allowed greater security during the entire process.

I would also like to thank my family, especially my parents and my brother for always being present in good and bad times and for having a friendly word when times were more controversial. All their support was essential not only during the writing of the dissertation but also during all the years that led to this being possible.

Finally, I would like to thank all my friends, for the moments not only of leisure but also of mutual help through these years. Thank you for all the support provided. To all of them and to all those who in one way or another crossed my path, I am grateful. Without you, none of this would be possible.

David Sousa

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.
I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

ABSTRACT

The growth of concepts such as Intelligent Environments and Internet of things allows us to understand the habits of users and consequently act to improve people's daily lives.

Through information gathering, it is thus possible to gather patterns about different kinds of human behavior and consequently build a learning model with predictive capabilities. In addition, there are increasing concerns from large companies about the influence, positive or negative, that aspects such as comfort and well-being have on the behavior and health of the population.

In fact, as human beings, we are greatly influenced by the environment in which we are inserted. There are therefore conditions in a place that give us certain levels of comfort that will eventually interfere with our well-being. However, it is difficult to identify which of these factors are relevant and how they intervene in our daily lives. Also, the habits we adopt as a result of the routines we follow can contribute to improving or worsening any of these indicators

With the help of the various types of sensors present, for example, in the smart devices (smartphones, smartwatches, wristbands), it is increasingly possible to collect information on these factors, easily and comprehensively. In this sense, firstly the main objective of this dissertation is thus to collect data on factors that may influence the user in order to create a user profile. These factors can be inferred through its interests, the visited locations, and its main activities. This objective involves a large-scale analysis, where there are no geographical restrictions. Furthermore, the study will be independent of the type of space (open or closed) that is explored. In that way, the perspective that will be used is from the user. Then there is an exploration of the data so that some intelligence can be inferred, and in this sense, build a mobile application capable of providing smart notifications based on user needs.

Keywords: Ambient Intelligence, Comfort, Deep Learning, Machine Learning, Smart Devices, Well-being.

RESUMO

O crescimento de conceitos como os da Inteligência Ambiente e Internet das Coisas, permitem perceber os hábitos dos utilizadores e, desta forma, atuar com o intuito de melhorar o dia a dia das pessoas.

Através da recolha de informação é assim possível padronizar diferentes aspetos dos comportamentos do ser humano e consequentemente construir um modelo de aprendizagem com capacidades preditivas. Para além disso, há cada vez mais preocupações por parte de grandes empresas, na influência, positiva ou negativa, que aspetos como o conforto e o bem-estar possuem no comportamento e na saúde da população.

De facto, como seres humanos somos bastante influenciados pelo ambiente em que nos encontramos inseridos. Existem, portanto, condições num local que nos proporcionam certos níveis de conforto que, eventualmente, irão interferir com o nosso bem-estar. No entanto, há a dificuldade em identificar quais desses fatores são relevantes e de que forma intervêm no nosso quotidiano. Para além disso, os hábitos que adotamos resultantes das rotinas que seguimos podem contribuir para melhorar ou piorar algum destes indicadores.

Assim, com a ajuda dos vários tipos de sensores presentes, por exemplo, nos vários dispositivos inteligentes (telemóveis, relógios, pulseiras) é cada vez mais ampla a possibilidade de recolher informação de forma fácil e abrangente. Nesse sentido, primeiramente o principal objetivo desta dissertação é assim recolher dados sobre fatores que podem influenciar o utilizador, de forma a criar um perfil de utilização. Esses fatores podem ser inferidos através dos seus interesses, dos locais visitados, e das suas principais atividades. Deste modo, este objetivo passa por uma análise em grande escala, onde não existem restrições geográficas, sendo que o estudo efetuado será independente do tipo de espaço (aberto ou fechado) que é explorado. Desta forma, a perspetiva abordada será sempre a do utilizador. Seguidamente, haverá uma exploração dos dados para que seja possível gerar alguma inteligência, e assim criar uma aplicação capaz de fornecer notificações inteligentes assentes nas necessidades do utilizador.

Palavras-Chaves: Aprendizagem Máquina, Aprendizagem Profunda, Bem-estar, Conforto, Dispositivos Inteligentes, Inteligência Ambiente.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AmI - Ambient Intelligence
ANSI - American National Standards Institute
ASHRAE - American Society of Heating, Refrigerating and Air-Conditioning Engineers
API - Application Programming Interface
APK - Android Application Pack
AQI - Air Quality Index
GMM - Guassian Mixture Model
ICT - Information and Communication Technologies
IDE - Integrated Development Environment
INE - National Institute of Statistics
KNIME - Konstanz Information Miner
LSTM - Long Short Term Memory
RNN - Recurrent Neural Networks
RSSI - Received Signal Strength Indicator
SMOTE - Synthetic Minority Oversampling Technique
SSID - Service Set Identifier
SVM - Support Vector Machine
WEKA - Waikato Enviroment for Knowledge Analysis

1

INTRODUCTION

According to the Oxford dictionary “habit” is defined as “a thing that you do often and almost without thinking, especially something that is hard to stop doing”. Now by the quick interpretation of this concept, we can infer that there are patterns that we are running throughout our lives. Recognizing these standards and acting accordingly can give people a better quality of life.

In this sense, concepts such as smart devices, intelligent environments, are crucial in building such systems. These bring systems that are very useful as they are able to ”feel” the environment around them. The perception of the context in which a person is inserted enables them to capture crucial data, which will then be used to construct a model that learns the user’s comfort and well-being.

Well-being and comfort have had a major impact on society. However, it is quite difficult to understand what these terms are and how they may be affected. In fact, one of the main questions that come up when we talk about monitoring comfort and well-being is: What to measure? The environment condition alone is correlated to well-being and comfort status (Silva and Analide, 2015). The heterogeneity of the environment and of the population makes difficult to perceive what may be affecting us in a given place. However, these days it is increasingly easy to explore these terms. Effectively more and more people have access to smart devices. These devices make the use of sensors easy and inexpensive since, for example, when we buy a smartphone it comes with a huge variety of sensors. This will allow some globalization, making possible to gather more heterogeneous information and give the possibility to answer this and others questions.

1.1 MOTIVATION

The use of technology has been growing exponentially over the years. According to *Gartner* by 2020, there will be over 20 billion devices connected to the Internet (Hung, 2017). The mass use of these devices allows us to create a new horizon of possibilities. It is in the exploration of these possibilities that the so-called term of artificial intelligence emerges. In fact, we are increasingly faced with the creation of ”smart” applications. On the one hand, these applications create solutions for a variety of new problems, on the other come facilitate existing solutions.

Furthermore, well-being and comfort are very complex terms and realizing what influences them can be quite difficult to accomplish. This is because we are greatly influenced by the environment around us, and the fact that it is dynamic implies a constant change in our feelings. In this sense, well-being conditions are increasingly in vogue and an increasing concern for people's health. This is because "higher levels of well-being are associated with decreased risk of disease, illness, and injury, better immune functioning, speedier recovery, and increased longevity" (CDC, 2018). Consequently comfort are also a vogue concern. Universities, hospitals, institutions spend most of their time promoting increased comfort and consequently reducing discomfort. We have the example of several studies that try to understand the impact that comfort has on employees performance, and how it can be used to make their performance as high as possible (Haynes, 2008). However, these companies face difficulties as they do not know how to perceive the degree of comfort so they usually end up relying on a single factor (for example, temperature). In short, the growing concern in this area, accompanied by technological development makes the work quite viable.

1.2 OBJECTIVES

The main objective of this dissertation is to evaluate, monitor and predict comfort and well-being in various types of environments.

However, I can divide this objective into four sub-objectives. Since the theme of the project implies the existence of data, a first phase is required for the collection and processing of data. Therefore, it is important to study how the data should be collected and which technologies should be used to store this data.

Secondly, build a model that can predict the degree of well-being and comfort associated with different aspects. At this stage, it is crucial to take into account model performance and optimization.

Subsequently, the third sub-objective will be to send notifications to the end-user, which will serve as a kind of advice on their comfort and well-being. Also in this sense, as the forth sub-objective, the goal is to give the possibility of viewing comfort and well-being on a map, in aggregate mode (for several users) through the use of various techniques. Throughout these phases, I will create an application in which all these assumptions are met.

1.3 METHODOLOGY

Initially, the work will go through a problem definition and identification phase. Subsequently, in order to give greater credibility to the whole dissertation, scientific articles that validate all that, will be mentioned and will be considered throughout this dissertation. In this sense, the working method initially involves the collection of some articles and solutions already in force, similar to the proposed theme. Thus, by combining various approaches it will be possible to obtain a more reliable end result and thus to gain a better understanding of the themes that the dissertation involves. Finally, the working method involves the validation of the developed solution so that it can be verified that the initial problem has been solved. In short, the methodology followed is called action-research since there is a collaboration between the researcher and the member of an organization (in this case the University of Minho) to solve a problem. It is thus a cyclical process that first involves a focus selection and an identification of research questions. The next step is to collect and analysis of data. Finally, before the project be presented to the scientific community it is necessary a conclusion phase where the following project is validated. In short, "action research is not about hypothesis testing and producing empirically generalized result" (Young et al., 2010), but is about analysis of the results of the experiments that are performed.

1.4 PLANNING

According to the available time, it is possible to divide the dissertation elaboration into six essential tasks.

In this sense, the first comprises a study on the proposed theme through the writing of its state of art. Still in this phase, it is intended to study some existing tools and understand the challenges the whole project faces. The second task is intended to be elaborated at the same time as the first, since it presupposes the elaboration of a pre-dissertation report to be delivered on 15 January 2020. Subsequently, implementation will begin. As it was a long and very practical phase, a three-month interval was established for the construction of what was requested (Task 3). At the end of this task, it is necessary to study the results and test the proposed solution, which is task number 4. In a final phase, it is intended to make scientific dissemination of the study carried out thus providing contributions to the entire community. Finally, the last task comprises the whole process of writing the dissertation. It is intended that this task occur simultaneously with others.

Although there was a slight deviation from the dates initially set due to the global pandemic, all the defined objectives were met and achieved.

This whole process, modeled initially, can be consulted in a more practical way, through a *Gantt* diagram, in Fig. 1.

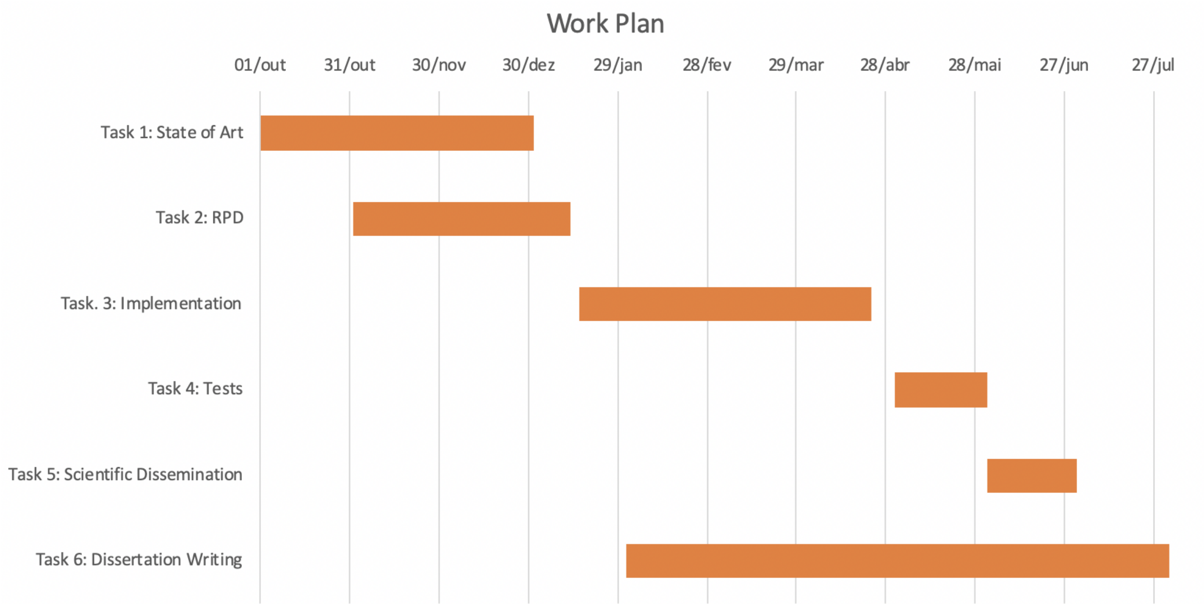


Figure 1: Work Plan

1.5 DISSERTATION STRUCTURE

This dissertation is divided into six fundamental parts. The first refers to the theme definition. This part answers questions like what are the objectives and the motivations for the whole study. Still in this phase is approached the kind of methodology that will be used and the planning for finished what is proposed.

The second refers to the state of the art, that is, to the definition of all concepts considered central to the understanding of the whole study.

The third part shows of similar works in this area and how they are equal or different from the proposal dissertation theme.

The forth will be used to allude to a general and specific architecture that can be followed to answer the main objectives of the dissertation theme. It will be presented the built application and some of the functionalities. Furthermore, it will be also described some of the main technologies used.

The fifth part will present a case study with some experiments made using the developed application. In this part, the work focuses mainly on the construction of machine learning models and the results obtained through the application of different models. The six-chapter, comprises the conclusion of the entire dissertation, concerning future work. It is also presented some relevant work that alludes to the scientific contribution made.

Finally, the document ends with the bibliography and with the appendices relative to the developed dissertation.

2

STATE OF THE ART

2.1 WELL-BEING AND COMFORT

As one of the focus of this dissertation is comfort and well-being, it is important to define these concepts in order to clarify their meaning. In this sense, on the next sub-chapters, is presented a little about these and how they are distinguished from each other.

2.1.1 *Definition of Well-being*

The term well-being is quite difficult to define and its refinement has been the subject of many studies over the years. It is a term that is often used by the community, but not always with a correct connotation. In fact, because of the associated confusion, we do not always know the meaning of "being well". In this sense, several authors seek to define this concept. On the one hand, according to Elena Alatarsteva and Galina Barysheva, we can subdivide this term into two strands: one subjective and one objective. "This distinction makes it possible to differentiate between well-being as an inner condition of an individual and well-being according to a quality of life criterion as a result of and as a member of development in society." (Alatartseva and Barysheva, 2015) The authors go further and highlight all the factors that may be included in each one of these new terms. On the one hand, objective well-being is characterized by wage levels, residence conditions, educational opportunities, social quality, the environment, security and civil rights. On the other hand, subjective well-being is conceptualized only as an internal state of an individual.

In addition, by consulting various articles published by researchers from different branches, it is possible to arrive at the definition of this concept more specifically, that is, by categorizing it into different classes:

- **Community/Societal well-being:** The ability to participate in a community, having engagement and involvement in a area where you live (Rath and Harter, 2010).
- **Economic/Financial well-being:** There are some studies that defend that have a good financial life can reduce stress and increase security. In fact, a study published on

"Happiness and Economics" (Frey, 2002) defend that there is a link between well-being and economics.

- **Psychological well-being:** When well-being is related to mental aspects is commonly be referred as psychological well-being. Some evidences suggest that psychological well-being is associated with diseases and mortality risk (Trudel-Fitzgerald et al., 2019).
- **Physical well-being:** This type of well-being emerges associated to physical health, improve, for example, with exercise habits (Rath and Harter, 2010).
- **Social well-being:** Some authors, see the relationships with others as a essential type of well-being (Davis, 2019). This is because, maintaining a support network can help to overcome loneliness (Rath and Harter, 2010). In this case, some studies tries to establish a relationship between the Big Five domains (that is, the five factors that model our personality: openness to experience, conscientiousness, extraversion, neuroticism and agreeableness) and social well-being (Joshanloo et al., 2012).
- **Work well-being:** It is related with the ability to interests pursue, doing what is liked. Some studies, focus in investigate the relationship between job satisfaction and work engagement as a dimension of work-related well-being (Rothmann, 2008).

To achieve well-being, we need to make sure that all of these types are being satisfied. Although there are several ways to categorize well-being, in general, they all end up touching the same points.

While several studies address the difficulties that the definition of well-being is (Dodge et al., 2012), others seek to define this concept more practically. In search of a good definition, without ambiguity, some studies by some authors were presented. However, it is admitted that there are other studies with different views from those presented here.

Alongside these definitions are several works that aim to measure the well-being level of a population. In the case of well-being, habits are seen as the key. For better or for worse, habits influences health, well-being, and quality of life (Stoewen, 2017). Take physical well-being as an example. Physical exercise habits can help improve this type of well-being. So, if a person change their habits for the better, they can change their life for better. It is also important to emphasize that the places we visit are part of these habits. Several studies denote their importance and how effectively they can influence us. A 2017 study (Pearce and Curtis, 2017) showed that much that needs to be done to improve well-being goes also beyond our control. The places where we live and where we spent our time can influence our health. Being mental well-being one of the main concerns of this study is refereed that places can be damaging to our mental health if, for example, they are characterised by high levels of crime, pollution, poor physical conditions and a lack of secure and rewarding employment opportunities.

On a global way, well-being is really used and differ in their level in each country. In fact, societies with higher well-being are the ones who are more economically developed and with low levels of corruption. Also, cultural factors can play a critical row in well-being. Because of the subjective nature of well-being, it is usually measured thought self-reports (CDC, 2018). In Portugal, *INE*(National Institute of Statistics) published an article in 2018 assessing the welfare index of the Portuguese population (INE, 2017). In this study, well-being is explored as a whole and was created more than 79 indicators grouped by two perspectives: material living conditions and quality of life.

In short, well-being is a complex concept but efforts were being made with the objective to understand what's needed and how and where we can work to improve our lives in a complex world.

2.1.2 *Definition of Comfort*

Finding and maintaining comfort is common to all people. Many people quickly identify loss of comfort and work diligently to restore it.
In our days, the term comfort is often seen related to the marketing of products like chairs, cars, clothing, hand tools, and a lot of others (Vink and Hallbeck, 2011). Evidence from the literature suggests that comfort, as a concept, has historical and more recent relevance for nursing. There are several authors, only in this context, which has its own definition. A few examples of that are Orlando (1961), where comfort is identified only as a pivotal to nurse's role, Roper et al. (1980), and Paterson & Zderad (1988) where comfort is defined as three states (ease, relief, and transcendence) (Tutton and Seers, 2003). It is, therefore, a concept that needs further exploration to increase knowledge in this area because it is difficult to reach a consensus on its definition. In order to establish the definition of comfort, this dissertation will use the meaning present in the Oxford dictionary. According to this, comfort is a broader concept that can be defined from a physical and a psychological perspective. In the first case, it is seen as a "state of physical ease and freedom from pain or constraint". In the second it is defined as: "The easing or alleviation of a person's feelings of grief or distress".

Despite all these definitions, they all note the existence of factors that influence comfort. Some papers show that different activities during measurements can influence comfort, concluding that characteristics of the environment and the context, can change how people feel (Vink and Hallbeck, 2011). For example, let consider that we have two chairs exactly equal to each other and we have two people on each chair do different work: computer work and telephoning work. Despite all the conditions be the same, the chair may be more comfortable for computer work, so will have positive effects on the first person and negative effects on the second one. This allows us to realize that comfort is a complex problem that varies from person to person. In fact, little aspects such as the mood (temporal state of the human mind) of a person can have a positive or negative impact on comfort (Silva and Analide, 2018). Alongside these other examples can be made. Temperature is always used as a crucial factor in various papers. In fact, the temperature can have a big influence on how we feel in a short time. In addition, other authors consider features such as air quality and humidity as factors that play a critical role (Zhang et al., 2011). These can and will be explored later (Chapter 5).

2.1.3 *Distinction between Well-Being and Comfort*

Although comfort is often associated with a synonym for well-being, a clear difference can be demonstrated between both concepts. By the previous definitions, it is noticeable that comfort is most used to classify the atmosphere that surrounds the human being. This implies that comfort is associated with momentary aspects. However, well-being arises characterized by an exhaustive variety of groups, which makes it associated with a long-term context. By way of example, a person may find himself comfortable but unhappy (and vice versa). In this sense, well-being is referred to in many studies as being progressively improved, while comfort is something that can be improved on time. As we can see, the mental health organization in the UK said that "it is important to realize that well-being is a much broader concept than moment-to-moment happiness" (Foundation, 2015). As we saw in the previous definitions, well-being is seen as the satisfaction of a set of terms. So, if we focus, for example, on mental well-being, reference can be made to the Warwick-Edinburgh scale. As we can see from the Fig. 2, we realize that the questions are based on events that occurred over two weeks.

The Warwick-Edinburgh Mental Well-being Scale (WEMWBS)

Below are some statements about feelings and thoughts.
Please tick the box that best describes your experience of each over the last 2 weeks

STATEMENTS	None of the time	Rarely	Some of the time	Often	All of the time
I've been feeling optimistic about the future	1	2	3	4	5
I've been feeling useful	1	2	3	4	5
I've been feeling relaxed	1	2	3	4	5
I've been feeling interested in other people	1	2	3	4	5
I've had energy to spare	1	2	3	4	5
I've been dealing with problems well	1	2	3	4	5
I've been thinking clearly	1	2	3	4	5
I've been feeling good about myself	1	2	3	4	5
I've been feeling close to other people	1	2	3	4	5
I've been feeling confident	1	2	3	4	5
I've been able to make up my own mind about things	1	2	3	4	5
I've been feeling loved	1	2	3	4	5
I've been interested in new things	1	2	3	4	5
I've been feeling cheerful	1	2	3	4	5

Figure 2: Warwick Edinburgh Mental Well-Being Scale (School, 2019)

Although comfort and well-being are distinct terms, this does not imply that they exist separately. Indeed, both terms arise intrinsically linked. Comfort at a certain level contributes to well-being. We can take the example of thermal comfort. A long term bad thermal comfort can "have a large effect on the mental and physical health of the occupants of any building" (Maclean, 2017). This is because the occupants will become distracted by the adverse temperature and their morale, health (physical well-being, for example) and productivity will be affected. However, it is important to note that well-being is only influenced if these conditions stay for too long.

2.2 AMBIENT INTELLIGENCE

Interaction with computers (not necessarily personal computers and laptops) these days is increasing and increasingly difficult to avoid. With the advent of microprocessors and smartphones, it is normal for the user to have no way to escape contact with such systems. This widespread availability of resources sparked the realization of Ambient Intelligence (Cook et al., 2009).

In a more theoretical way, Ambient Intelligence (AmI), formerly known as ubiquitous computing, consists of a "discipline that brings intelligence to our everyday environments and makes those environments sensitive to us." (Cook et al., 2009) This is a term that arises associated with the conditions of well-being and comfort, since the main objective of this concept is to build environments that allow improving the efficiency of systems, thereby increasing people's quality of life. User experiences with computers over recent decades have created an interesting context where expectations of these systems are growing and people's fear of using them has decreased (Park et al., 2012).

This context allows us to create usage profiles by having the system identify and act according to the needs of each one. However, for this to be possible, systems that provide environmental intelligence have to go through three phases (Fig. 3): sensing, reasoning, and acting (Aztiria et al., 2010).

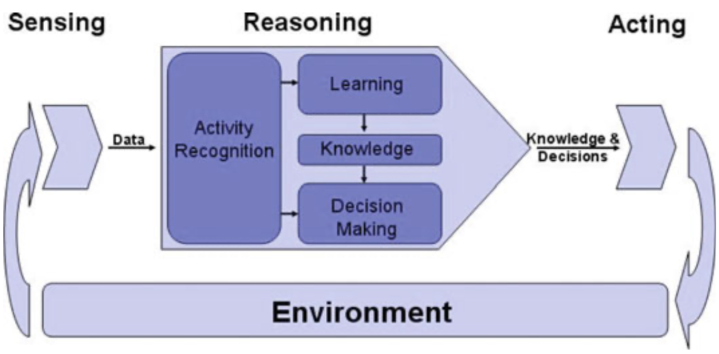


Figure 3: Phases in environmental intelligent system (Aztiria et al., 2010)

2.2.1 *Sensing*

Environment Intelligence relies on information gathering. In this sense, the use of sensors is a crucial aspect for the sensing process: "Ambient intelligence research builds upon advances in sensors and sensor networks, pervasive computing, and artificial intelligence" (Heger, 2014). Moreover, being one of the focus of this dissertation the sensing, it is important to know this concept better. There are two types of sensors: Physical Sensors and Virtual sensors. Physical sensors are the traditional sensors that we use to observe a physical phenomenon. On the other side, a virtual sensor is something used as an API.

The monitoring process can be divided into two main areas: monitoring of users and their activities and monitoring of the environment. Both aspects are extremely important. While the first allows a better understand of the habits of users, the second allows knowing the environmental phenomena that surround them (temperature, humidity, light level, etc.), providing another context to data that is collected. Of course, once the information has been obtained, the system architecture may follow a centralized model in which the sensors send the data to a central server for processing or a distributed model where each sensor has the processing power. The process of filtering, disambiguation, and interpreting data before it is used is called middleware. Thus, data collected from different sources can be combined, through a sensor fusion process, and integrated to produce more refined and complete information (Heger, 2014).

2.2.2 *Reasoning*

Once the data has been collected and properly processed, some kind of "intelligence" needs to be applied. To make algorithms responsive, adaptive, and beneficial to users, several types of reasoning must take place. These include user modeling, activity prediction and recognition, decision making, and spatial-temporal reasoning (Cook et al., 2009).

Modeling

One of the main features of a responsive algorithm is the ability to model user behavior. User modeling approaches are based on the data that is used to build the model, the type of model that is built and the nature of the model-building algorithm.

Activity prediction and recognition

A reasoning algorithm is useful if "have the ability to predict and recognize activities that occur in AmI environments" (Cook et al., 2009). Depending on the type of sensors data that is captured, different approaches can be used. Predicting, for example, resident locations, and even resident actions, allows the AmI system to anticipate the resident's needs and assist with performing the action.

Decision Making

Decision making is the process of making choices by identifying a decision, gathering information, and assessing alternative resolutions (Dartmouth, 2018). Decision-making can be seen as a step-by-step process. This type of process can help to make more deliberate and thoughtful decisions. For this motive, this approach helps in identifying the most satisfying solution possible. For example, the AmI system may respond to a sensed health need by calling a medical specialist and sending health vitals using any available device (step 1). If there is no response from the specialist, the AmI system would phone the nearest hospital and request ambulance assistance (step 2).

Spatial Temporal Reasoning

"For a system to make sensible decisions it has to be aware of where the users are and have been during some period of time" (Cook et al., 2009). This, together with other information, can be used for inferring the context where the user is inserted. After this, the system could perform the most adequate responses. For example: Whenever someone turns on the cooker and leaves it unattended for more than 10 units of time, then the system has to take action. If the user leaves the house it is important to turn off the cooker to prevent a possible fire.

2.2.3 *Acting*

The most natural way of acting is using actuators integrated into standard devices. Another mechanism is through robots. In order this to work, is necessary to define human-centric computer interfaces that are natural and provides context-awareness. With a natural interface is possible to use ambient intelligence closer to human communication (like talk, for example). Context awareness, on another side, is the key to build AmI because if the system can exploit the user's activity then it can act according (Cook et al., 2009).

2.3 SMART CITIES

There are many definitions of smart cities. The term is often used along with alternative adjectives, that emerge by replacing "smart" with "intelligent" or "digital". The label "smart cities" is a fuzzy concept and is not always used properly. There is neither a single template of framing a smart city, nor a one-size-fits-all definition of it (Albino et al., 2015).

However, in this context a Smart City can be viewed as an "urban innovation and transformation that aims to harness physical infrastructures, Information and Communication Technologies (ICT), knowledge resources, and social infrastructures for economic regeneration, social cohesion, better city administration, and infrastructure management". (Edward Curry and Sheth, 2016).

In fact, the emerging ICT paradigms such as data-intensive computing, Internet of Things, Cloud Computers and others, are essential to the realization of the vision of Smart Cities. Actually, real-world Smart Cities are being enabled by a combination of these paradigms using a mixture of architectures (centralized, decentralized, and a combination of both) and infrastructures (Edward Curry and Sheth, 2016).

One thing that smart cities bring is the concept of services. "Cities are structures of services and these services are things through which people interact within the city systems, together with other people" (Kuchta, 2014) . The most interesting advantage of this services is the opportunity to access to suitable data. In this way, stakeholders can access wide online services, with portals for basic information, citizen services, business, and tourism, all based on a common infrastructure. As the main purpose of this dissertation is to capture information, this aspect can be really useful. Smart cities are deploying different online services in different areas. These services cover a wide variety of sectors, such as transportation (intelligent road networks, connected cars, and public transport), public utilities (smart electricity, water, and gas distribution), education, health and social care, public safety. The concept of smart cities is increasingly in vogue so that different applications are emerging. These can range from fields since as disaster control to intelligent building constructions. This promotes the concept of the connected city since these services have applications in many areas like smart grid, smart home, security, building automation, remote health and wellness monitoring, location-aware applications, mobile payments, and other machine-to-machine (M2M) applications (Kuchta, 2014).

2.4 MACHINE LEARNING

One of the best-known areas of artificial intelligence is machine learning which is known for "provides systems the ability to automatically learn and improve from experience without being explicitly programmed" (Team, 2017). With the definition of machine learning comes another, **Deep Learning**. Deep Learning is a subset of machine learning. It is considered the next evolution of machine learning algorithms. The designation "deep" comes because of the many layers that usually this type of technique has. The main difference between these two areas is that deep learning can automatically discover the features to be used while in machine learning the features have to be provided manually (Garbade, 2018a).

In this work, these concepts can be useful to make predictions. When we speak of the term "predictions", we speak of the outputs that are achieved from the training that a given algorithm had on a given dataset. In many cases, the term does not refer to future issues. It can be used, for example, to predict if there has been fraud in a given transaction that has already happened. Predictions are very important as they allow highly accurate guessing, thus helping in various aspects. However, predictions are the last step of an extensive process. Before the application of this concept is necessary to collect and prepare the data. In this sense, despite the existence of various approaches the one that could be followed could be seen in the next figure.

Figure 4: Machine learning process (van Rijmenam, 2019)

The name of each phase is self-explanatory (Fig. 4). First is necessary to proceed to some data collection. After that, is necessary to prepare the data making it suitable for the training process. In fact, in the real world, the data often is dirty and can not be directly used. This means that raw data alone is not very useful.

Data cleaning, discretization, integration, and transformation are some examples of steps that can be applied in this phase. After this initial phase, is necessary to choose a model to train. The chosen model will be evaluated and will be the target of parameter tuning which means that some of the parameters of the model will be chosen in order to improve the performance. Finally, the best model is used to make the predictions wanted.

The concept of good guessing is dependent on how the models fit the problem and how good that model has learned. In fact, all of these terms are bounded by the concept of "learning". There are different types of learning: supervised learning, unsupervised learning, and reinforcement learning.

There are several types of algorithms that can be associate with these types of learning. Thus, below will be defined and will be given some examples of useful traditional machine learning algorithms. However, some examples of deep learning algorithms will also be given in the section. Although the type of learning used with these algorithms is supervised learning, such separation is due to the fact that the learning classification, depends on the way that they are used.

2.4.1 *Supervised Learning*

Is given to the machine what each data block means. This is, the model is trained on a labeled dataset (Marsland, 2015). A label dataset is one that has both parameters: input and output parameters. In this case, learning is based on regression and classification techniques. If the output variable is a category, such as "yes" or "no", it is a classification problem. If the output variable is a real value, such as "weight" it is a regression problem. In this dissertation, may be useful, for example, a category technique because one of the main purposes is to identify well-being and comfort levels which can be represented by classes.

Naïve Bayes: Classification Algorithm

It is an algorithm based on Bayes theorem in which each value is treated independently. It allows, through probabilities, to predict a class or category based on a certain set of features. In simple terms, this classifier assumes that the presence of a feature in a class is independent of the presence of any other feature. For example, a piece of fruit can be considered an apple if it is red, round and if have 7 cm in diameter. Even though these characteristics depend on each other, Naïve Bayes algorithm will consider each of these properties to contribute independently to the fact that the piece of fruit is an apple (Ray, 2017). This problem is widely used in the classification of text and in problems that involve many classes. Although it is a simple algorithm, it usually presents good results mainly in very large datasets.

Support Vector Machine: Classification Algorithm

In this type of algorithm, the data is plotted in an n-dimensional space (where n represents the number of features). The value of each characteristic, in the case of SVM, will be represented as a specific coordinate in a referential (support vector). Subsequently, in general, the objective is to find a hyperplane (borders) that marks the difference between two classes, usually represented by the furthest distance from the nearest support-vector. This type of algorithm is very used principally for cases where very high predictive power is required (Ray, 2017).

Decision Tree: Classification Algorithm

As its name implies, it is a tree-like structure in which each branch is used to illustrate a possible result of a decision. With a decision tree, it is possible to make decisions with a given set of input. The most common procedures in building a decision tree are induction and pruning. In the first one, the tree is built and in the second one, the tree is simplified by removing several complexities (Team, 2019). This type of algorithm is useful in data exploration.

From this algorithm comes another one called **Random Forest**. Like its name implies, consists of a large number of individual decision trees that operate as an ensemble. Each tree in the random forest represents a class prediction. In a simple way, there is a voting phase, where the trees are voting randomly. The class with the most votes is chosen by the algorithm, becoming in that way the model's prediction (Wakefield, 2019).

Nearest Neighbours: Regression and Classification Algorithm

It is an algorithm that can be used in both types of problems, classification, and regression. However, it is often used in the industry for classification problems. K-Nearest Neighbors estimates the probability that a data point belongs to a particular group. In this way, this classifier is able to learn patterns from the data. In general, the algorithm looks at a set of points that surround a single data point, in order to determine which of the groups that point belongs to. Depending on the number of nearby points, this decision is made (Ray, 2017).

2.4.2 *Unsupervised Learning*

Unlike what happens in supervised learning, in this case, it is the machine that has to learn by itself concepts that it has never seen before. It is used when the information in training is not classified. For this reason, the process is longer and consequently less popular. The goal of unsupervised learning is to model the underlying structure or distribution in the data in order to learn more about the data. There are two types of unsupervised problems: clustering and association. Clustering is when is necessary to discover inherent groups in data, while the association is when is necessary to use rules to describe large portions of data. In this dissertation, on one hand, we may have a clustering problem, because we may want, for example, to group people by their well-being. On the other hand, we may have an association problem if the goal would be to identify that the people that have X well-being, tend to go to Y place.

K Means: Clustering Algorithm

The main objective of this algorithm is to find iteratively, among k groups, the group to which a given dataset belongs. K-means algorithm identifies k number of centroids (known as the location that represent the center of a cluster), and then allocates every data point to the nearest cluster while keeping the centroids as small as possible (Garbade, 2018b). This algorithm starts by chosen random centroids and then in an iterative way tries to optimize the position of that centroids.

Mean Shift Clustering

Mean Shift Clustering is a hierarchical clustering algorithm. The main objective is similar to K-Means since it tries to group the data in clusters without being trained on label data. However, the particular difference between this algorithm and the K-Means is that in this we do not need to know the number of categories, which makes the Mean Shift computationally heavy ($O(n^2)$) (Maklin, 2018).

2.4.3 *Reinforcement Learning*

It is a learning process, where there is an interaction with the environment and through the production of certain actions in that, a given algorithm will be compensated or penalized (Team, 2017). This is experience-based learning where the machine looks for the right approach. This can be useful for this dissertation if the algorithm is rewarded or penalized every time it gives a notification, in the right place, about the user's comfort or well-being. The most known algorithms of Reinforcement Learning are Q-Learning and Sarsa.

2.4.4 *Deep Learning Algorithms*

In the case of the algorithms more specifically associated with the concept of deep learning, their classification in the types of learning defined above is more complex. Effectively, this type of algorithms can be seen as supervised or unsupervised depending on how they are used. A term that arises with deep learning algorithms is the term of artificial neural networks. These types of algorithms were drawn to simulate the way the human brain analyzes and processes information (Frankenfield, 2018). There are many types of neural networks that can be used depending on what is necessary.

Recurrent Neural Networks

In the case of this dissertation can be really useful to explore the concept of Recurrent Neural Networks (RNN). RNN is a generalization of a feed-forward neural network, however, unlike this type of networks, RNNs has an internal memory that they can use to process sequence of inputs. The most advantage of this is that with RNN is possible to "model sequence of data so that each sample can be assumed to be dependent on previous ones" (Mittal, 2019). One of the most popular versions of RNN is Long Short Term Memory (LSTM) Network which is used to process and predict time series given time lags of unknown duration (Mittal, 2019). It trains the model by using back-propagation. In fact, comfort and well-being can be predicted more easily taking into consideration what happened in a previous time. In this way is important to deepen this algorithm. In an LSTM network, three gates are present (Mittal, 2019):

- **Input gate:** The main objective is to discover witch value that comes from the input that will be used to modify the memory. The *sigmoid* function through the attribution of the values 0 or 1 decide which values to let and the *tanh* decides the level of importance of the values by given weighting. Thee weight range between 1 and -1.
- **Forget gate:** Looks to the previous state (h_{t-1}) and the content input (x_t) and for each number in the cell state outputs a number. 0 means "omit" this and 1 means "keep

this". In short, the main function of the forget gate is to discover what details should be discarded from the block.

- **Output gate:** The memory and the input of the block are used to decide the output. Once again, *sigmoid* function decide which values to let by given the values 0 and 1. *Tanh* function gives weights to the values which are passed deciding their level of importance ranging from -1 to 1 and multiplied with the output of *sigmoid*.

Fig. 5 appears as an example of what has been described.

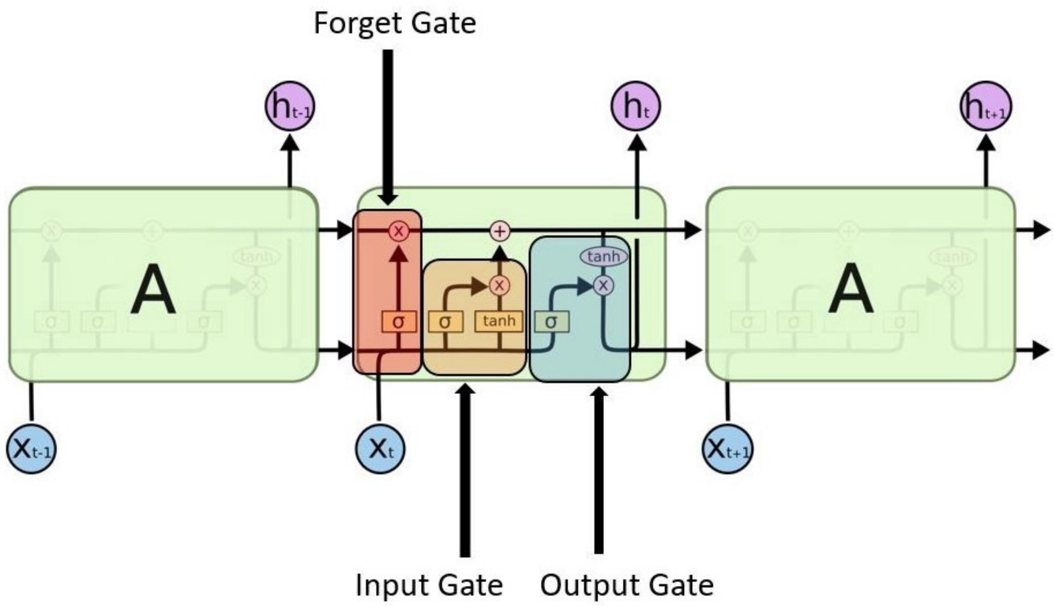


Figure 5: LSTM gates (Mittal, 2019)

Convolutional Neural Networks

It is a type of feed-forward neural network composed by several layers (multi-layer), witch is commonly used with visual data, such as image classification. While the first convolution layer captures low-level features, the next layers (witch can be convolutional, normalization, polling and fully connect layers) extract higher-level features which makes this algorithm efficient on image recognition. Furthermore, the fewer training parameters makes it easy to train. However, this algorithm have more application than image classification such as recommendation engines and video recognition (Simplilearn, 2019).

Gaussian Mixture Model

Gaussian Mixture Model is algorithm that is used to represent a normally distributed sub-population (in a cluster) within an overall population. Since GMM does not require the data to which a sub-population belongs, it can learn the sub-populations automatically. This algorithm is used for the soft-clustering method. For example, let's consider we have the height of men and women. However, we do not know which one is a man and which one is a woman. The only thing that we know is that the mean height of males in male distribution is 5′8″ and for females, it is 5′4″. We can say that distributions follow the sum of two scaled and two shifted normal distributions. This assumption can be made with the help of the Gaussian Mixture Model (Team, 2019).

Generative Adversarial Networks

With the increase of deepfakes problem, this type of algorithm can be quite current. In fact, Generative Adversarial Networks is used to generate new, synthetic instances of data that can pass for real data (Nicholson, 2019). This algorithm has the term "adversaria" because of the fact this architecture comprises two sub-models (generator and discriminator), pitting against each other. This type of network can discover patterns in input data. For this motive, this algorithm is usually considered unsupervised. It is used principally in cybersecurity and health diagnosis (Simplilearn, 2019).

2.4.5 *Machine Learning in practice*

The information above allows verifying that several types of algorithms can be applied to solve a problem that requires the application of concepts such as Machine Learning. To apply these concepts there are some useful tools.

One example of a popular tool is the Waikato Environment for Knowledge Analysis (WEKA). WEKA is an open-source machine learning software. Is recognized as a landmark system in data mining and machine learning. WEKA provides an extensive collection of machine learning algorithms and furthermore it allows data pre-processing making it possible for users to quickly try out and compare different machine learning methods on new data sets (Hall et al., 2009).

Another example that can be made is Konstanz Information Miner(KNIME). KNIME is a modular environment, which enables simple integration of new algorithms and tools as well as data manipulation or visualization. In fact, KNIME enables easy visual assembly and interactive execution of a data pipeline (Berthold et al., 2009).

Furthermore, there are specialized libraries and frameworks for the treatment of these problems. One of the most popular is *TensorFlow*. It was released in November 2015 by Google. This library makes it easy to build and train deep learning models (especially when combined with high-end APIs, like *Keras*) and allows the implementation of simple and flexibles architectures for neural networks. It can be used in Python and Javascript. Another example that can be maid is *PyTorch*. *PyTorch* was developed by Facebook for Python. It is easier to learn than TensorFlow, however, it has fewer resources.

In addition, there are some libraries that are really useful as they have mechanisms for data analysis and mining, such as *scikit-learn* (Python library that focuses on the implementation of Machine Learning algorithms) or *Deepnet* (R library, focus on deep learning algorithms).

2.5 Composite monitoring indicators

For accurate and realistic measurement to occur, it is necessary to define some terms in order to know the dimensions that compose them. In this sense, there are two classes of indicators. Simple indicators, where data collection is done through a direct reading of values and composite indicators where data extraction is more complex and requires some types of assumptions. So, here are some types of these indicators and what are the precautions to be taken so that they can be correctly inferred.

2.5.1 *Thermal Comfort*

Generally, associated with the concept of environment arises the term thermal comfort. Thermal comfort is defined as a mental condition in which satisfaction with the surrounding thermal environment is expressed. It is quite common for the thermal comfort indicator to be the air temperature since most people can easily relate to it. However, air temperature alone is not a reliable metric and there are other factors that allow the concept of thermal comfort to be improved. These factors are six and are mainly divided into two groups: environmental and personal. Environmental factors include air temperature, radiant temperature, air velocity, and humidity. Already with regard to personal factors, there is heatstroke caused by clothing and metabolic heat (Health and Executive, 2019). For these reasons, thermal comfort is a sensation that depends on one's personal opinion. A thermally comfortable environment for one person may be uncomfortable for another. A realistic goal in this type of study will be to create a thermal environment that provides comfort to the greatest number of people.

PMV/PDD method

The PMV model (Predicted Average Vote) and consequently the PPD model (Predicted Percentage of Dissatisfied) was developed in 1970 by Fanger with the aim of defining thermal comfort. PMV is an index "that aims to predict the mean value of votes of a group of occupants on a seven-point thermal sensation scale". This scale is symmetrical relative to 0 and has values of 1 to 3 which can be positive, corresponding to sensations of heat or negative, meant for cold sensations (Appendix B). The calculation of an individual's PMV is made according to the combination of the six parameters defined before (1). PMV has recommended limits between -0,5 and +0,5 (ANSI/ASHRAE Standard).

$$PMV = 0.303\, e^{-0.036M} + 0.028 \times \left\{ \begin{array}{c} (W-M) - 3.05e - 3[5733 - 6.99(M-W) - p_a] \\ -0.42[(M-W) - 58.15] - 1.7e - 5M(5867 - p_a) - 1.4e - 4M(34 - t_a) \\ -3.96e - 8.f_{cl}[(t_{cl}+273)^4 - (\overline{t_r}+273)] - f_{cl}.h_c(t_{cl} - t_a) \end{array} \right\} \quad (1)$$

With:

- W - Mechanical power, in Watt per square meter (W/m^2);
- M - Metabolic rate in watts per square meter (W/m^2);
- p_a - Partial pressure of water vapor in Pascal (Pa);
- t_a - Air temperature in Celsius degrees (ºC);
- f_{cl} - Surface area of clothing;
- t_{cl} - Surface temperature of clothing in degrees Celsius (ºC).
- $\overline{t_r}$ - Average radiant temperature in degrees Celsius (ºC);
- h_c - Convective heat transfer coefficient, in watts per square metre kelvin $(W/(m^2 * K))$

The PDD emerges because is also necessary to get a more holistic idea of the level of satisfaction of the occupants in space and how thermal comfort consequently can be achieved. Because of that, Fanger developed another equation, which is related to PMV, to predict the percentage of dissatisfied. The PPD can range from 5% to 100%, depending on the calculated PMV. However, the recommending factor vary depending on where the occupant is located in the building (Guenther, 2019).

2.5.2 *Light*

Another factor to consider when talking about comfort is light. The fact that the human being is in an environment with low light conditions can harm him and have some consequences for his health, especially visual health. In this sense, there are some tabulated values that allow having a reference to what is considered normal.

These values are tabulated as *lux* (Appendix A) and represent the lighting unit. Corresponds to the incidence of 1 lumen on a surface of 1 square meter. (ToolBox, 2004)

2.5.3 *Noise*

Sound is measured in decibels (dB). On one side, noises over 70 dB for a long period of time can cause hearing loss. On another side, noise above 120 dB may cause immediate loss of this capability. The more exposed to loud sounds the greater the risk of hearing loss. However, the louder something appears to a person, it is not the same as telling the true intensity of that sound. Sound intensity is the amount of sound energy in a given space. It is a logarithmic scale so the noise is not directly proportional to its intensity. This means that a 20 dB sound is ten times louder than a 10 dB sound. Also, two sounds of the same intensity may not mean that they are at the same height. The pitch of the sound has to do with perception and varies from place to place (for example, in a quiet place we can hear a sound that is not perceived when we are on the street). The risk of hearing problems increases with sound intensity rather than sound high. However, it should be noted that the effect of listening to a low-intensity sound for long periods is the same as the effect of a high-intensity sound for short periods of time. (Chuck Kardous and Lotz, 2016)

2.5.4 *Air Quality*

Air quality is important because it influences our health. In this sense, one of the major ways to measure air quality is to measure how polluted the air is. The air generally is essentially compound by two gases: nitrogen and oxygen. Nonetheless, there are several other gases and particles: ground-level ozone, carbon monoxide, sulfur dioxide, nitrogen dioxide, and airborne particles, or aerosols.

To measure this there is an index called Air Quality Index or AQI. This works as a "thermometer that runs from 0 to 500 degrees. However, instead of showing changes in the temperature, the AQI is a way of showing changes in the amount of pollution in the air" (SciJinks, 2019). There are some tabulated values for AQI levels (Fig. 6). An AQI under 50 means that the air quality is good. At a very high level can mean significant risks for human health.

However, in spite of air pollution play a critical role in air quality, other factors can affect this indicator, principally when talking about indoor environments (AFSCME, 2019). Some examples of these factors are:

- **Chemical exposures:** Chemicals in the air, that result from some materials or activities, can affect the air quality;
- **Infectious agents:** Sometimes, health can have critical importance in air quality. In hospitals, for example, biological agents, spread by ill people, can affect the air;

- **Physical agents:** The right amount of heat, humidity and air movement are important for comfort and can have an impact on the air quality.

In short, the term *air quality* itself, is not only about pollution but about a combination of factors that influence the air.

Air Quality Index Levels of Health Concern	Numerical Value	Meaning
Good	0 to 50	Air quality is considered satisfactory, and air pollution poses little or no risk.
Moderate	51 to 100	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.
Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is not likely to be affected.
Unhealthy	151 to 200	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects.
Very Unhealthy	201 to 300	Health alert: everyone may experience more serious health effects.
Hazardous	301 to 500	Health warnings of emergency conditions. The entire population is more likely to be affected.

Figure 6: AQI values (AirNow, 2019)

2.5.5 *Ultraviolet Index*

According to the World Health Organization, the UVI is a measure of the level of Ultraviolet radiation. The higher the UVI, the greater the damage provoked to skin and eyes (Fig. 7), and, consequently, less time is necessary for harm to occur. This value can vary throughout the day and depend on the geographical location and weather daylight it can have different implications.

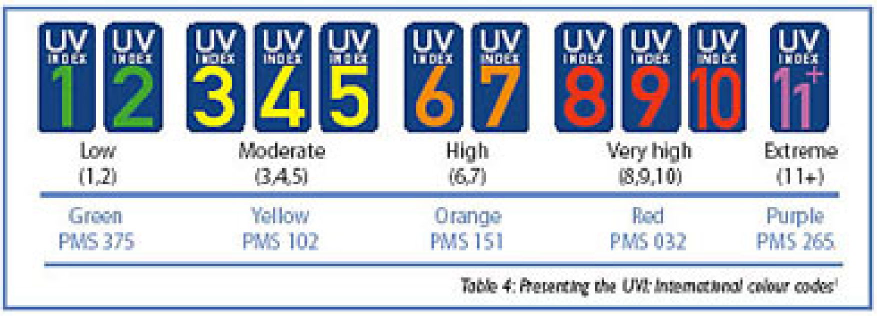


Figure 7: UVI values (WHO, 2020)

2.6 CHALLENGES

Some of the challenges present in this project are:

- **Privacy:** Since the main purpose is to collect as much information as possible about a given user, there may be some concerns about the privacy of that information.
- **Battery:** Making a data collection app means have to do background activities so that captured information be more as possible. This can lead to a high battery drain, something, not all users are willing to.
- **Time:** Gathering information takes time to do and large numbers of users. Actually, being a dissertation inserted in an academic context makes difficult to reach a large number of users for the expected development time.
- **Device Variability:** When collecting data, it is always necessary to assess their reliability. In this case, the large variability of devices is a disadvantage as it may mean that not all have the same sensors. This may result in a lack of data making it inconsistent.
- **Services:** During the set time some of the services may stop working, making the data that will be used infeasible. Furthermore, the use of some of the services through some *APIs* is not free so the large-scale use implies high costs.

However, when developing the application, these challenges will always be taken into account in order to minimize the problems. For example, concerning privacy, some steps can be taken to ensure anonymity. The main measure will be not to save any data that identifies the user. In the case of battery drain, every way to optimize the use of the battery will be taken into account while developing an application. An example of how this can be achieved is by reducing the number of accesses to the system. Furthermore, services with the least possible limitations will be chosen and, in a way, an attempt will be made to optimize the number of access to them.

In short, during the development of an application, possible solutions and optimizations will always be considered to minimize the consequences of these challenges to the maximum.

3

RELATED WORK

Generally speaking, there are several solutions we use every day that are related to the concept of data collection and intelligent suggestions. Take for example the case of virtual assistants such as Siri, Cortana, and others. This type of system collects as much information as possible about the user and the environment around them, and from there suggests and performs actions that simplify people's lives. In this sense, the following are some projects that aim to improve people's lives and that in one way or another meet what is proposed in this dissertation.

3.1 BEWELL

It is an application designed to make people care about certain concepts of their well-being, concepts that are often overlooked. This is the case, for example, lack of sleep, hygiene or high social isolation. The authors believe that this situation is caused by an absence of adequate tools for effective self-management of overall well-being and health.

The main idea of this application is be capable of monitoring multiple dimensions of human behavior, encompassing physical, mental and social dimensions of well-being.

In terms of architecture (Fig. 8) the developers write in the article that: "*BeWell* application and system support consists of a software suite running on Android smartphones and cloud infrastructure. The software components installed on the phones include: i) Sensing Daemon, which is responsible for sensing, classification, data processing (e.g., privacy preserving audio processing) and uploading of sensor data; ii) Mobile *BeWell* Portal, which displays the user's behavioral patterns and the well-being score associated with these behaviors; and iii) Mobile Ambient Well-being Display, which provides an always-on visualization of the user's well-being scores" (Lane et al., 2011) .

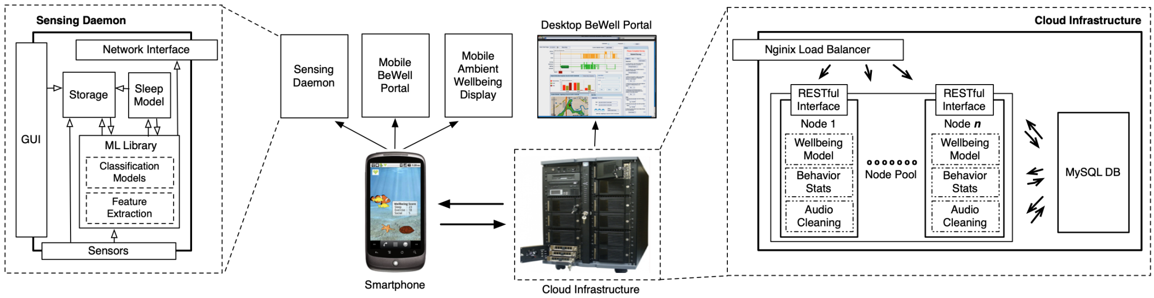


Figure 8: BeWell Architecture (Lane et al., 2011)

3.2 THERMOOSTAT

"Heating, Ventilation, and Air Conditioning (HVAC) typically account for more than 40% of energy use in institutional buildings (DOE 2010)" (Sanguinetti et al., 2016). In spite of this, thermal comfort conditions in campus buildings are frequently poor. The main purpose of the *ThermOOstat* project was to build a participatory thermal sensing app to capture the feedback of the University of California occupants.

In fact, the general strategy is to solicit thermal comfort "votes" (e.g., hot, cold) from building occupants via a web or mobile app. With the results obtained, the application was still able to collect if people were willing to sacrifice some degrees of temperature for the environment. In short, participation in *TherMOOstat* promotes higher satisfaction with indoor temperatures on a specific campus contributing also to energy saving (Sanguinetti et al., 2016).

3.3 TRACKING THE TRANSIT RIDER EXPERIENCE

The objective of this work is by a smartphone application to understand what are the triggers that are causing negative perceptions of personal comfort and well-being on a trip. The study involves 60 students who travel on a different public transit bus with the smartphone survey app installed. The app uses GPS tracking along with voice and text capture to enabled a contextual understanding. The aim is to improve "the transit service amenities to provide a more comfortable experience, build rider loyalty, and make transit a more desirable mode choice" (Rid, 2015).

3.4 PROSIT

Researchers of Dalhousie University have developed a mobile application with the aim of detecting some mental conditions, such as anxiety and depression. This is possible through an initial data collection using the user's smartphone and the capacity of those devices, for example, using their sensors. This data includes information related to the practice of physical exercise, history of messages and calls, musical tastes, among others. Users are also asked to record a 90-second audio clip describing the most exciting part of their week and to self-report their feelings on a five-point scale.

This app is currently being tested and has about 300 participants who have been testing the features (Smith, 2020).

3.5 OPENDATA APPLICATIONS

OpenData emerges to designate the idea that data is freely available to anyone who wants to use it. This concept and the works related to it are in one way or another linked to the main objective of this dissertation. In fact, one of the objectives is to create a kind of geographic map to consult the comfort and well-being indicators. In this sense, first, there is a collection of data, individually, which is done through smartphones and then there may be the availability of this data for consultation by the community. In this way, it is possible to gather the information collected from a high set of sensors and later provide a system where the conditions of comfort and well-being can be easily accessed and updated in real time.

It is on this premise that several projects arise. One of them, which can be given as an example, is the work "Mobile Data Collection With Wired and Wireless Sensors" (Chaudhri et al., 2012). Its main objective is "to provide a high level framework that allows for customization and flexibility of applications that interface with external sensors, and thus support a variety of information services that rely on sensor data". Subsequently, the article introduces four types of applications in which, through the framework presented, data can be accessed, also in a simple way and in some cases in real time. These applications include sensors as important as medical sensors. It is in this context and thanks to data interoperability, that more and more concepts are being applied in this area.

3.6 COMPARATION WITH THE PROPOSAL

In the case of the *BeWell* application it only makes use of some device sensors whereby only a few pre-established well-being metrics are evaluated. If we look at the *ThermOOstact* application we realize that it only cares about the comfort of the university campus. The fourth example is the most similar to the dissertation's purpose. In fact, the application

Prosit can be seen as an inspiration for what is proposed. However, as we can see these applications are really specific as they only act in a specific problem.

The purpose of this dissertation is more comprehensive because a lot of different data will be collected and according to that data will be possible to explore a wide range of concepts. This brings a great advantage because it allows the creation of a broader application. On the one hand, at a personal level, it is possible to create an application that fits the context in which a person is inserted thus providing intelligent notifications regarding that context. The purpose application is about learning how the data captured by smart devices can affect people. The theme of this dissertation focuses on sensor data and discovering, by inference, how that can affect people, positively or negatively. Furthermore, a lot of studies, like the *ThermOOstact*, focus on closed spaces since is more difficult to study open spaces because of heterogeneity. The proposed dissertation intends to study any type of space regardless of the geographical barriers that may exist. On the other hand, it is also possible to exploit an enterprise level as each employer can use the application to understand what the ideal characteristics of a particular location are and use it to their advantage.

4 ARCHITECTURE

According to the system architecture, it is possible, in addition to fundamental components such as the database, to divide the architecture into three strands. A first strand oriented only to data collection so that the data can be stored and processed later. In this strand is important to explore what can be captured by the different smart devices. The objective is to collect data relevant to the classification of comfort and well-being. A second strand more related to the building of predictive models. It is in this strand that some processing will be applied to the captured data, and it will be decided what is important to use or what is less relevant, depending on the target. A lot of different programming languages can be used in this strand. Finally, a third strand that needs the last two strands. This strand is responsible to provide a smart component to the system. Furthermore, it is important in this phase give to the user the opportunity to access their comfort and well-being. This architecture (Fig. 9) only serves to demonstrate how a solution to the problem can be built. The following section shows a more specific answer to the initial problem.

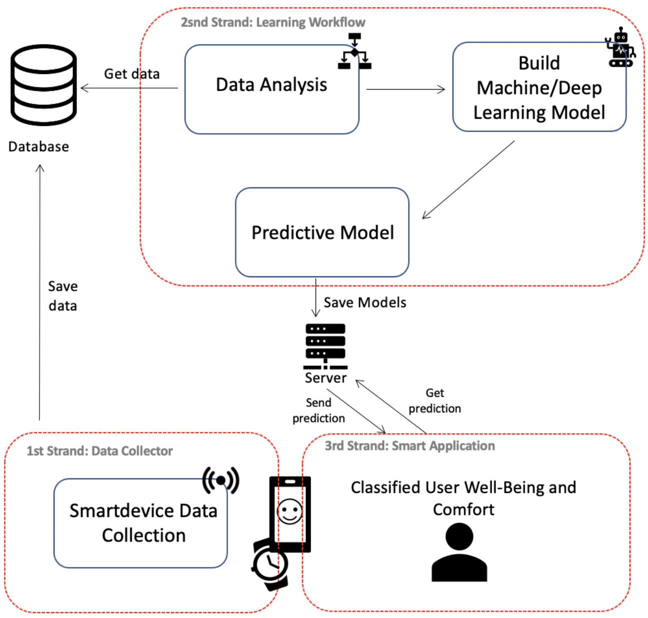


Figure 9: General architecture

4.1 PROTOTYPE SYSTEM

The collected data will be stored in a database, in this case, it was being used the *firebase*. With the data captured some models of machine learning were built. For this purpose was used Python programming language because of the extensive libraries that this language has. In the next sub-chapters will be described, with detail, every component of this architecture. The approach presented intends to follow what is outlined in Fig. 10. However, other alternatives for the various components could have been used, as this architecture is only part of a study.

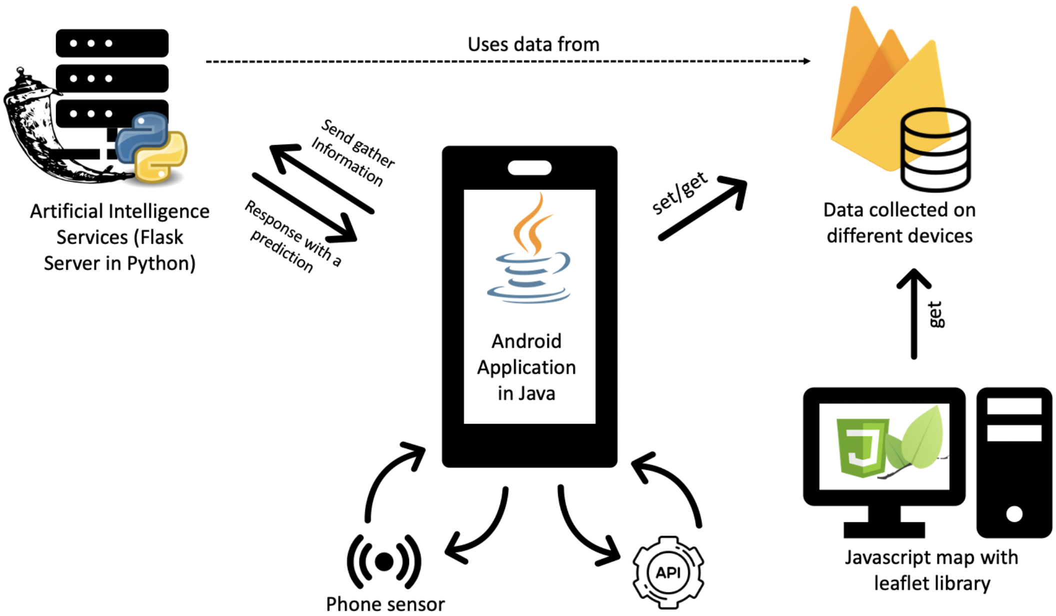


Figure 10: Case study implementation

4.1.1 *Android System*

For the whole application logic, it was used Java, where it was possible to access android sensors and establish a bridge between the client and the database. The Java language was chosen because of the large number of libraries it has and because I have more experience with it. The use of the Android Studio IDE allowed the frontend to also be developed more easily. Android was the operating system chosen since it is the most used in Portugal[1](at the time this work was done). Furthermore, being an open-source operating system has fewer restrictions in accessing some of the functionalities of a smartphone (like sensors, for example).

1 https://gs.statcounter.com/os-market-share/mobile/portugal

In this case, the study was developed only for smartphones since these devices allow many advantages such as the use of embedded sensors and communication structures at zero cost, high mobility, and easy application dissemination.

4.1.2 *Flask Server*

A python server, developed with Flask, will be responsible for receiving the different requests for access to the models that were developed in the previous section. It is in this server that the data received is processed and transformed according to the models that will be used. Flask is a micro-framework, however, according to the documentation, "micro" does not mean that the entire web application has to fit into a single Python file, although this is entirely possible. It also doesn't mean that Flask lacks functionality. "Micro" in micro-framework means that Flask seeks to keep the core simple, but extensible. Despite being "micro", it is ready for use in production, fulfilling a wide range of needs.

4.1.3 *Firebase Database*

Firebase provides a realtime database and backend as a service. The service provides to application developers an API that allows application data to be synchronized across clients and stored on Firebase's cloud. Although there are several types of databases, firebase was selected, mainly due to its good integration with mobile applications, something that is required in this project. Furthermore, the database has some advantages. Some of the more interesting for this dissertation are that it allows a simple and secure way to authenticate different users and real-time synchronization of the data in the application.

4.2 DATA COLLECTOR STRAND

This first strand, as its name implies, is a very simple data collection application. In this application, some crucial decisions have been made. Since the main objective was to capture information, it was first decided what kind of data would be collected and how.

4.2.1 *What data was collected?*

Every time a "new" user starts the application he has to fill out a short questionnaire (Fig. 11). This will serve to gather more general information about the type of user dealing with the application. However, it is not saved any information that identifies him (such as email, for example) and this questionnaire will only appear once in the application usage time (if a ”new” user enters the application two times, only for the first time the form will appear).

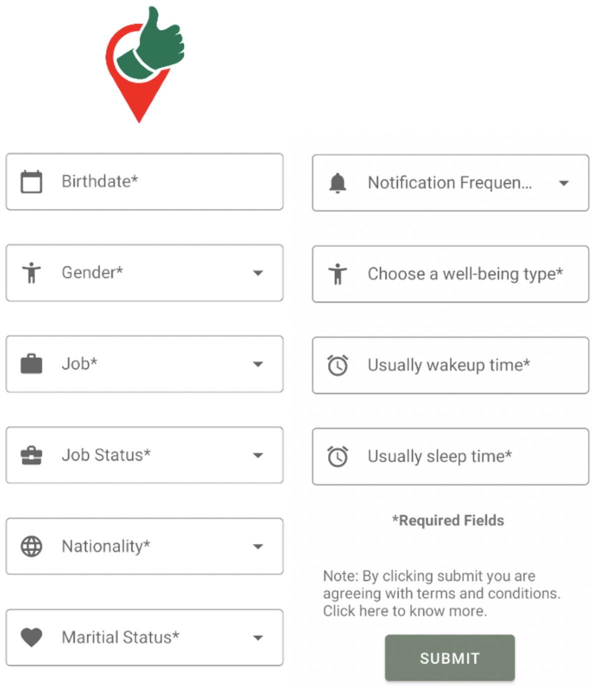


Figure 11: Initial Survey

Since it is an application developed for Android is possible to use the inherent sensors and capture other relevant information of this type of system. The data that were captured are the following one:

- **Environment Type:** What is the type of environment that the user is inserted (outdoor or indoor environment). Useful to understand if the user is in an open or closed space. Users in an enclosed space are not subjugated to some variations in environmental conditions;
- **Steps:** Number of user steps for the day. Useful to understand the user’s activity level;
- **Ambient Temperature:** Measures the ambient temperature in degrees Celsius (°C). A scarce sensor in most equipment, however, aims to measure the environmental temperature that is felt in a given space. As refereed before, temperature is linked to comfort and poor temperature conditions, in the long run, can cause decreased well-being;
- **Light:** Measures the ambient light level (illumination) in lx. Sensor present in most smartphones. It is very useful to perceive the lighting of a certain place. Poor lighting

can be related to conditions of comfort and well-being. ”Bad lighting is associated with a range of ill-health effects, both physical and mental” (Agarwal, 2018);

- **Magnetometer:** Measures the geomagnetic field of the environment for the three physical axes (x, y, z) in μT. The measurement of the magnetic field can have some influences, although not so noticeable, on our well-being and comfort;

- **Gyroscope:** Measures the rotation rate of a device in rad/s around each of the three physical axes (x, y, z). Sensor widely used, and combined with the accelerometer, also serves to deduce some of the activities of users and thus understand habits that may be influencing them;

- **Accelerometer:** Measures the acceleration force in m/s^2 that is applied to a device on the three physical axes (x, y, z), including the force of gravity;

- **Humidity:** Measures relative humidity in percent (%). Humidity as a sensor also exists on a small number of devices. It is a factor that can cause health problems and in this way it can cause decays in our well-being;

- **Location:** Latitude and longitude obtained by the location sensors on the smartphone. Understanding the places users pass through can be important to understand the level of comfort and well-being as mentioned above;

- **WiFi Connection:** See if the device is connected to the internet through WiFi. There are many studies that prove that users who have a WiFi connection feel more comfortable, hence it is an interesting indicator to be collected;

- **WiFi Info:** Some WiFi information like SSID, RSSI and Mac Address of the provider. Information about WiFi can be used to extract some interesting factors such as the name of the location where the user is using the SSID;

- **Battery:** Remaining battery percentage. More and more users have their lives dependent on a smartphone. The number of people using the type of services provided by it has been growing and smartphones have been replacing objects constantly present in our daily lives (such as ATM cards, tickets for public transport, among others). In this sense, low battery levels can cause signs of stress and anxiety affecting the comfort and well-being of users (Ellis, 2019);

- **Calls Number:** Number of answered calls for the day. The number of calls, especially those of long duration, can help to combat loneliness, thus contributing to social well-being;

- **Calls Duration:** Total time spent on calls for the day (in seconds);

- **Number of Events:** Total number of events for the day. The number of events in a day is useful to understand if there is availability for the user to meet other people and thus also improve their social well-being;
- **Activity:** Performed activity at the time by the user. Can be one of the following activities: "Still", "Walking", "Running", "Tilting", "In_Vehicle", "On_Bicycle" or "Unknown". With the activity is also saved the confidence level of the activity indicated. It is an indicator that can be quite useful mainly for physical well-being;

In addition, in conjunction with this it was used APIs so that the information gathered was even wider. This panoply includes the *OpenWeather* API that allows collecting a vast amount of data associated with weather conditions (such as **weather state**, **wind speed**, **ultraviolet radiation index**, **pressure** and the **temperature**) and the *Foursquare* API that allows capturing location-related data (such as **local category**). The category of a location is a very important concept, since for example in places like restaurants users tend to spend more money, affecting financial well-being. Besides, another example that can be given is that places like parks can contribute to good mental and physical well-being (as explained before). Still in this sense, some spaces can provide a greater degree of comfort than others. Therefore, it is a very interesting attribute to be collected.

Additionally, an API was also used to understand air quality (such as **AQI** and the **dominant polluter**). Air quality is often mentioned in studies and ends up affecting our comfort and well-being. The only metric that could not be collected was the sound. It is a metric that varies from device to device so that capturing it would first require calibration.

4.2.2 *How users can indicate their comfort or well-being?*

The second decision to make was how users would be asked about their comfort and well-being. In this sense, to avoid being too intrusive, two situations were taken into account. On the one hand, the questions about comfort are made by context analysis. That is, when it is detected that the user has moved, questions are asked more frequently. This check is made by using a circle formula to confirm if the new location is inside the area defined with the older location. This comparison only happens at specific times, to avoid some problems. On one side, to avoid battery draining since this check can be really consuming. On another, this check cannot be performed every second because if for example, the user is travel, he will constantly trigger comfort questions.

Another situation is that these questions only appear at specific times. In this case, is given to the user the possibility to choose how often the questions occur. There are three options: low, medium or high frequency:

- **Low:** As the name indicates, low means that the user will receive notifications with less frequency. In this specific case, the notifications will only appear every hour and a half, if the user doesn't change his location or each half-hour otherwise;
- **Medium:** If the medium option is selected, every one hour, the user will receive a notification. The location inspection will be reduced to fifteen minutes.
- **High:** This means that they will receive a new notification every thirty minutes, and in case they are in constant location change, the notification will happen every seven minutes. If the users don't mind about the intrusion they can always choose high notifications frequency.

In the case of the well-being question, this will appear only less frequently. This is due to what was defined in the previous section. Indeed, it is assumed that well-being throughout the day will not have as many variations as comfort. However, every time the user stops the data capture, after 5 minutes a well-being notification shows up. In cases that the user doesn't stop the data collection, the notification shows up every two hours.

One another hand, the user is also allowed to answer a question about comfort or well-being voluntarily without waiting for a notification.

To make both process faster and more intuitive was used as a response scheme smiley faces. This idea came about through an adaptation of the Likert scale, widely used in opinion polls. As Fig. 12 demonstrate, the questionnaire for well-being is more complex to be possible to associate the different type of data collected to different types of well-beings (defined in state of art).

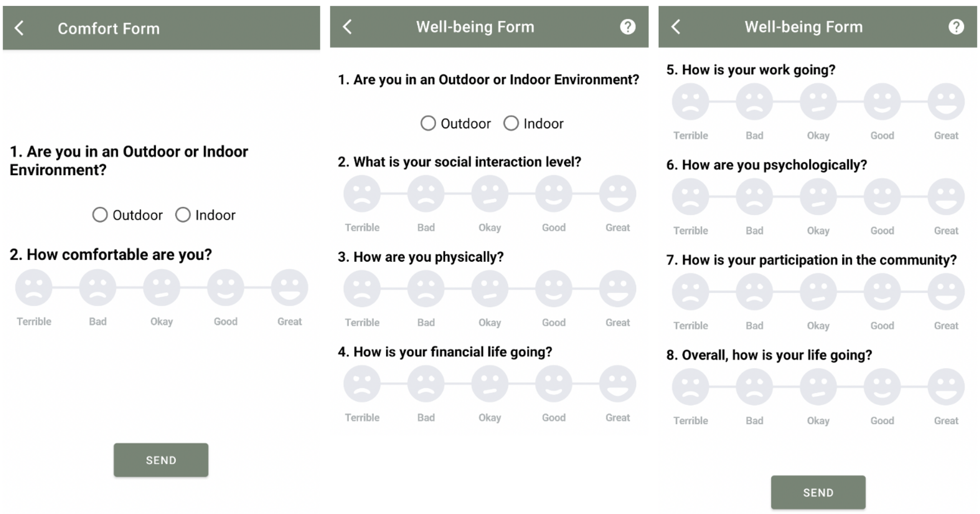


Figure 12: Questions about comfort and well-being

However, the fact that no question is asked does not mean that data are not being collected anyway. Between each question are performed background operations that continue to store information regarding the context in which the user is inserted. This case only has the particularity that the comfort or well-being questions has the value ″-1″ which means no question has been asked. This could be useful because we can interpret this ″-1″ value as a non-alteration on comfort or well-being.

The background activity turn's on and off automatically. As can be seen in Fig. 11 is asked to the user the time that they usually wake up and go to sleep. When a user does not answer a question after the defined sleep time, is assumed that he is sleeping and for that motive, the background sensing is turned off. When is already wake-up time, the background sensing turned on automatically and the questions start again to show up.

Throughout the development of the entire application, the challenges presented were always taken into account. In that sense, some measures were taken to combat these eventual problems. One of them was to, at no time, use any type of information that identifies the user. Thus, any contact he has with the application becomes anonymous. In fact, no email is requested, or any form of user identification. Regarding battery issues, an attempt was made to optimize the battery, avoiding access to some sensors when it was not necessary.

4.3 SMART APPLICATION STRAND

With the building of machine learning models will be possible to make the application more "intelligent" and, furthermore, will be viable the creation of some smart solutions. On one hand, this prediction of comfort and well-being will result in less interaction of the user, so, it will be possible to get a major number of contributions with less effort. This is crucial since with more data the creation of a map where conditions of comfort and well-being could be seen by region or place is achievable. On the other hand, with those models it will be possible to react to user behavior and thus give some smart advices.

With this new strand, the application begins to work based on two modes: Capture Mode and Prediction Mode. The main difference between these is that Capture Mode as the name implies is used only to capture data. In this mode, the user only will receive notifications with the question enunciated before. In its turn, the Prediction Mode, if enabled, will give the application the capacity to predict comfort and well-being. However, the background activity is similar to what happens in captured mode. The major difference is that the application does prediction about user comfort and well-being reducing in that way the number of presented questions. Furthermore, it is in this mode that smart advice will be provided. It is important to note that not all types of well-being are being predicted. In fact, some types of well-being have been selected which, according to the data used, would be possible to obtain better results. However, through a more extensive collection, the other dimensions can be explored.

When changing the selected mode, the application will adapt (Fig. 13). The home page gives priority to what users want to see in the first place. Depending on what mode is on, this page shows a summary of the numbers related to some of the user's interaction. In the case of well-being, since there are many types, the home screen will privilege the type chosen under the settings page. In the next subsections will be described in detail some of the extra functionalities added to the initial application, where the capture of data was the only objective. The settings page (which is responsible for mode change) could be consulted in the Appendix C.

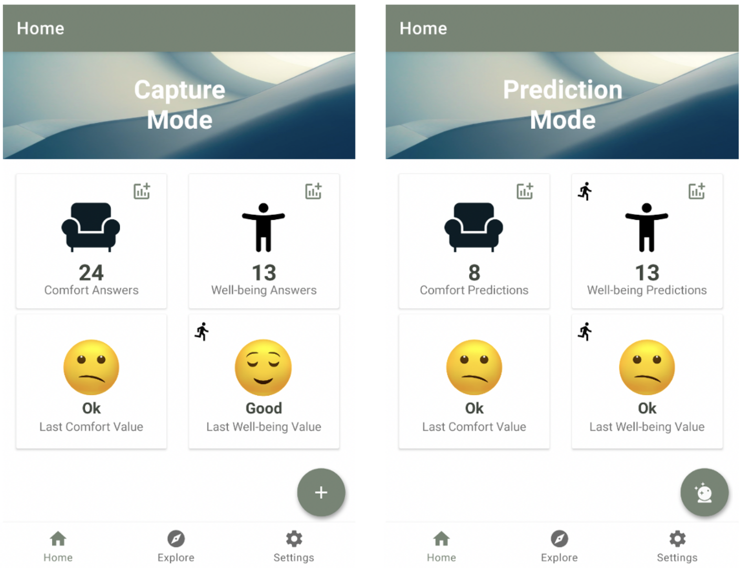


Figure 13: Capture Mode vs Prediction Mode

4.3.1 *Wellness Place Map*

One of the objectives of this dissertation was to create a map capable of showing both well-being and comfort. In this sense, some changes were made to the initial data collection application. Thus, each time the user answers a question, the data is now also saved as an entry on the map. In addition to the latitude, longitude, country, postal code and administration area, a set of values are also saved that allow a characterization of the place. To this set of values belongs the ambient temperature, AQI, a general and specific category of the place, humidity, light level, magnetic field value, atmospheric pressure, the intensity of ultraviolet rays, temperature, and wind speed. These values are being updated not only through the response to the notifications provided, but also through the predictions that are being made in the background. In practice, what happens is that the developed algorithm retrieves from the database all existing records for the region in which a user is located, through a filter by postal code. Then it goes through all these records and those that are less than a certain distance from the place where this user is, are updated with the captured values, otherwise, a new entry is created in the database (Algorithm 1). In this way, whenever users who are in the same place answer a question (or a prediction is made) an average of the conditions and questions answered by the two or more users is kept. In these cases, it can always happen that several users simultaneously access the same fields in the database. To avoid inconsistencies, transactions are used, this allows users to always have an updated version of the data.

Algorithm 1 Update map entries

```
X ← false
if data exists on database then
  for Location L in data do
    Distance ← L.distanceTo(MyLocation)
    if Distance < 100 then
      Calculate mean
      Update data on database
      X ← true
    end if
  end for
end if
if X == false then
  Create new entry on database, with the new conditions
end if
```

Since the efficiency of the application is one of the concerns, the use of a recursively calculated average was considered so that it could be stored in the database. So whenever there was a new contribution, the average would simply be updated. For this reason, when it was necessary to make a display of the results, no extra calculation was necessary since everything that is eventually needed would be stored. Two different types of average were considered. A flat average and an weight moving average. The flat average is the commonly used. For greater precision, this average implies keeping the sum of the values until the moment and the total number of contributions. The other strategy considered was the use of a weight average (Eq. (2)) in which the new values had a greater contribution to the calculation of the average (Moreno, 2020). This was therefore the strategy used since it makes sense that the most recent values were more important.

$$S_t = \begin{cases} X_1, & t = 1 \\ \alpha * X_t + (1 - \alpha) * S_{t-1}, & t > 1 \end{cases} \tag{2}$$

Where:

α: Constant smoothing factor between 0 and 1;

X_t: Is the value at a time period t;
S_t: Is the value of the moving average at any time period t.

For the visualization of these data, two maps were built, one in the android application using the apk of google (Fig. 14), and another web-based, using JavaScript and the leaflet library (Fig. 15). However, both maps have their similarities. In both cases, the records found in the database are drawn with the help of markers. Markers are icons that allow indicating specific places on a map. Thus, each marker has a color and an icon that varies according to the specific average value of a given area and the type of factor that is being demonstrated. It was also added circles to the thee markers that allow representing the wide area. In addition, these markers are grouped into different clusters. This choice was made because a high number of markers is expected, so the clusters allow an efficient control of these. In addition, depending on the chosen zoom level, the clusters are grouped and allow an overview of comfort and well-being, either by area or country. In fact, invisible borders have also been created that allow the values of one country not to mix with another. Some extra features have also been added to the map, such as the possibility of applying filters. These include the possibility for users to see only the data added for a particular day (a day ago or the day itself) and not an aggregate total of several days. For greater consistency, and for greater coherence, users can choose to consult a joint average, obtained by answering questions and predictions, or, they can choose to consult an individual average (only from answers to questions, or obtained only from predictions).

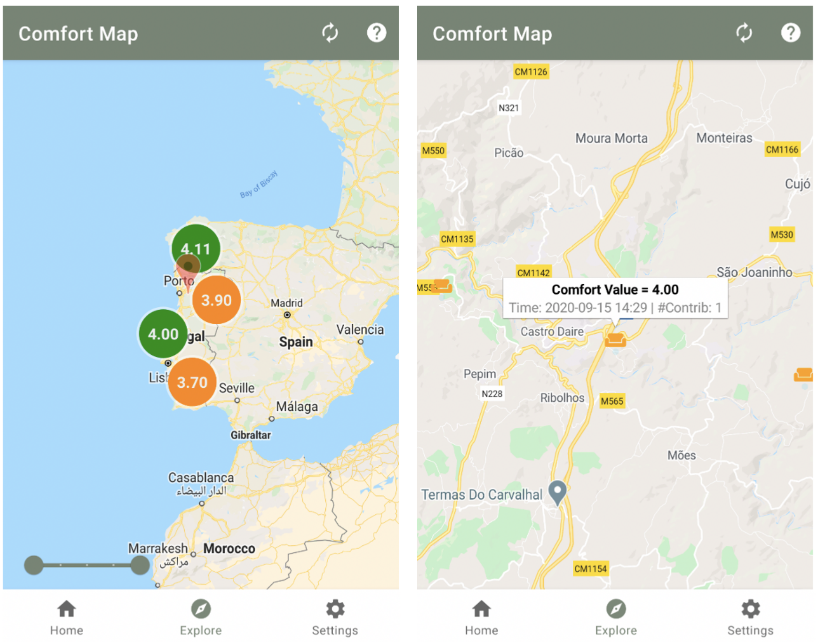


Figure 14: Android Map Example

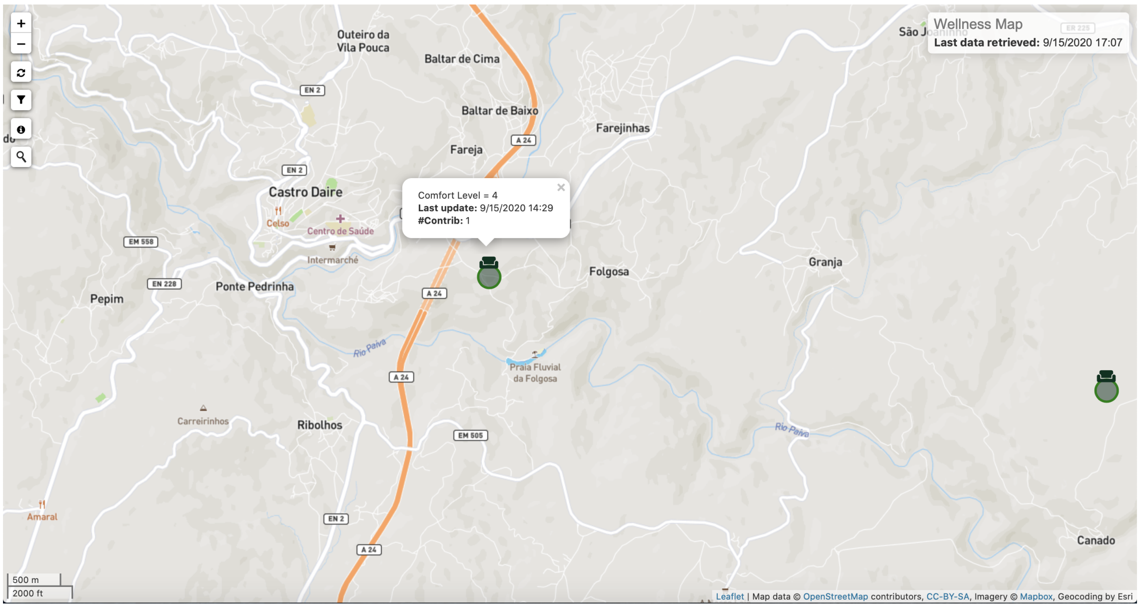


Figure 15: Javascript Map Example

4.3.2 *Smart Advice*

With the development of machine learning models, it was possible to create an application capable of predicting comfort and well-being. Thus, with the help of these models, it was possible to create ways of providing intelligent advice.

Whenever the user is found to have a poor well-being or comfort value, an advice is generated and consequently an notification shows up. It was defined that the advice of the same type can not occur less than two hours of difference so that users have time to act.

In the case of well-being, since there are many types the user can choose in what of these types wants to be advised. This helps to simplify user's interaction with the application.

Two types of models were developed, one in which the algorithm used was a traditional machine learning algorithm and another in which the algorithm used was deep learning, more specifically the LSTM. According to the model developed for a certain type of well-being, the application will follow different behaviors. In case of the models that use LSTM what will happen is that the user, even in the prediction mode, will have an early stage of receiving notifications with questions related to their well-being, with more frequency. This is because these types of models presuppose the existence of a history of records in order to be able to make a prediction. Even after having collected the necessary number of records, users, over time, continued to receive questions, although less frequently, to update some aspects that cannot be collected automatically. When all the necessary premises for the given of advice are verified, the strategies applied follow the presuppositions presented in Fig. 16. There are two different ways to provide advice. One is due to a change of location and the other is due to the suggestion of changing some conditions, be they environmental or personal. The advice provided is chosen randomly. In the next subsections, different forms of counseling will be discussed in more detail.

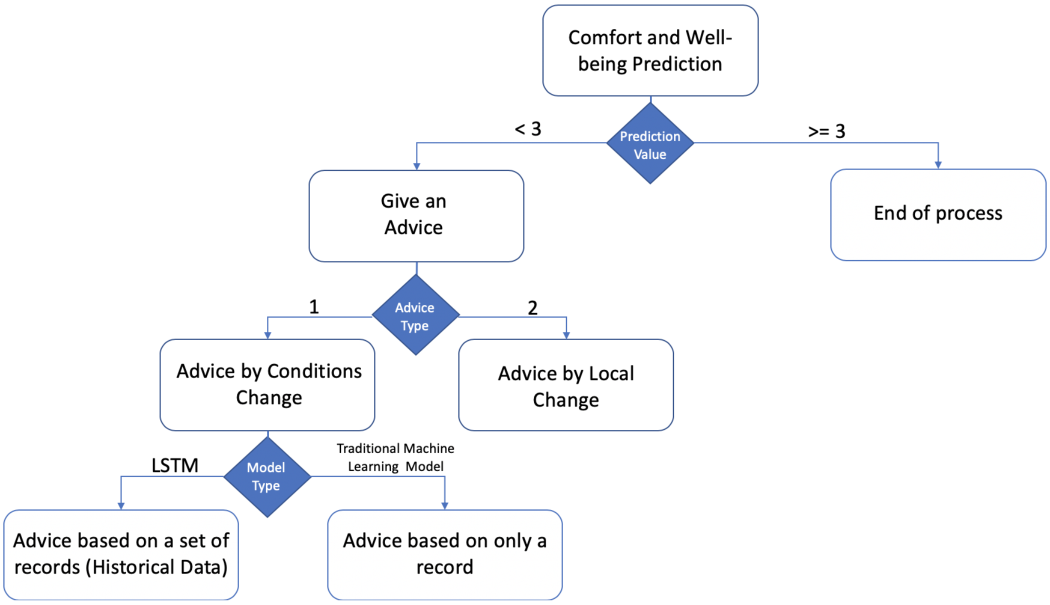


Figure 16: Example Advice Process

Advice by Local Change

One of the strategies is to suggest a place close to the one where the user is, that makes him feel comfortable or well, depending on what is being evaluated at the moment. Using the attribute of the postal code it is possible to quickly search for the nearest place with a high degree of comfort or well-being. To avoid suggest a residential area, a filter is also applied to the location category.

After choosing the location, the user receives a notification that informs him that an advice has been generated regarding his comfort or(and) well-being (if the estimated values are low). When clicking on that notification the user is redirected to a page where the last generated advice is found. If this advice is actually for exchange of places, a link will be presented which, when clicked, will direct the user to google maps, where he will be able to consult the suggested zone and the routes to get there (Fig. 17). However, since the user may not have google maps, they can still consult this location on the map described above.

When a location advice is impossible, because of the nonexistence of records, the system gives an advice by condition change.

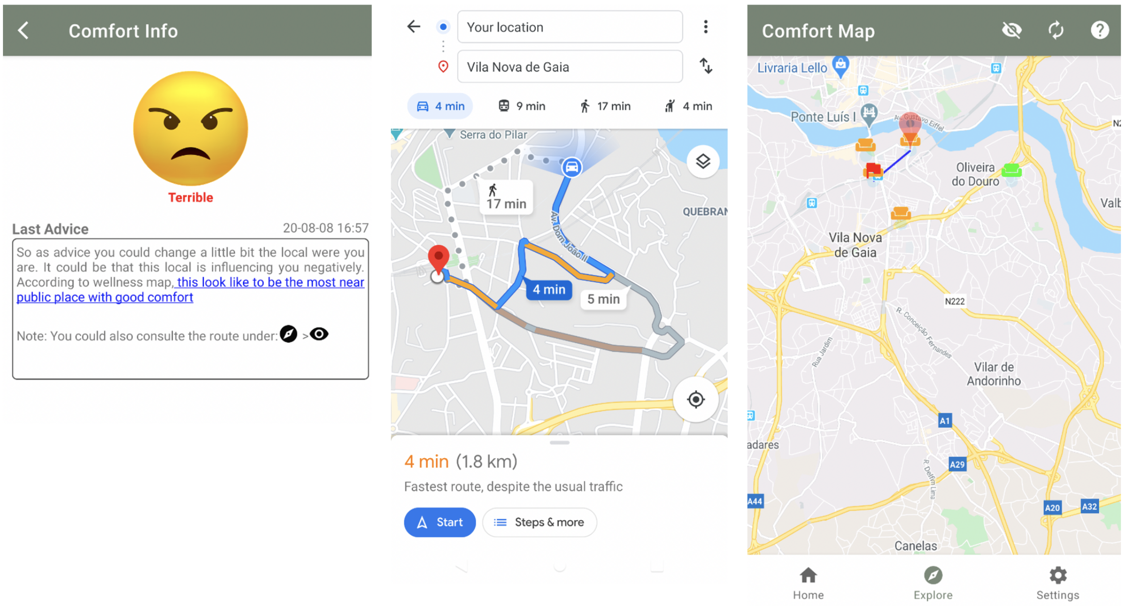


Figure 17: Advice by local change

Advice by Conditions Change

Another strategy is to provide advice according to conditions similar to the user who needs to be advised (Fig. 18). In this sense, this involves searching the database for the record with the conditions most similar to those users, but which has a good classification of the indicator that he wants to be advised. By comparison, we can understand what is causing bad comfort or well-being. This comparison is made using the equation present on Eq. (3). The objective is to subtract condition (personal or environmental) and registration information between two users. The record whose subtraction value is lower is the most similar. Some aspects could be adapted. By giving more weights to parcels, is possible to make a parcel more, or less important.

Thus, it is possible to suggest aspects that the user can possibly improve. Although this technique can be applied to both types of models, traditional machine learning models, and deep learning models, there is a small difference the last one. In this case, whenever a user of the application has a good prediction, an average of the records used in that prediction is stored in the database, since the deep learning models use a set of records for the classification. Thus, whenever another user (or himself) gets a bad rating, the database is searched for records whose conditions are similar to those that were captured but with a good rating. Then, the advice occurs in a similar way to what happened previously and the user is informed of changes that he can make in his routine to improve the degree of comfort or well-being.

$$Similar_{val} = X \times Conditions_{(1,2)} + Y \times RegisterInfo_{(1,2)} \tag{3}$$

$$Conditions_{(1,2)} = lightLevel_1 - lightLevel_2 + pressureLevel_1 - pressureLevel_2 + humidity_1 - humidity_2 + ...$$

$$RegisterInfo_{(1,2)} = Age_1 - Age_2 + Nationality_1 - Nationality_2 + Job_1 - Job_2 + ...$$

Where X and Y are real numbers. These values are the weights given to each of the parcels.

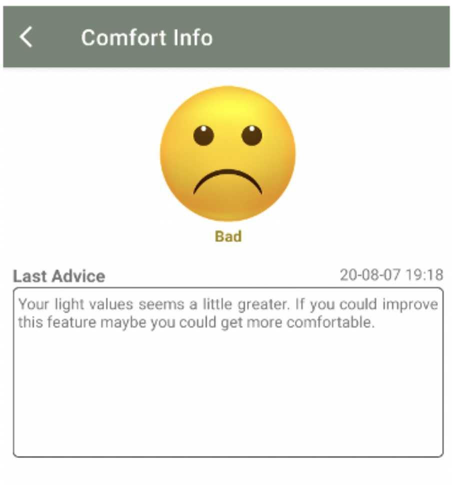


Figure 18: Advice by conditions change

Dashboards

So that users can follow the evolution of their comfort and well-being over time, some charts have been created that allow them to visually demonstrate what is happening (Fig. 19). In these graphs, users can see what the last ten values of comfort or well-being were, or they can also access the most frequent types of classes. The data represented in the dashboards could be of three different types: mixed data, only data that come of the answer to questions, and only data that come of a prediction. The user can always change the type that wants to display.

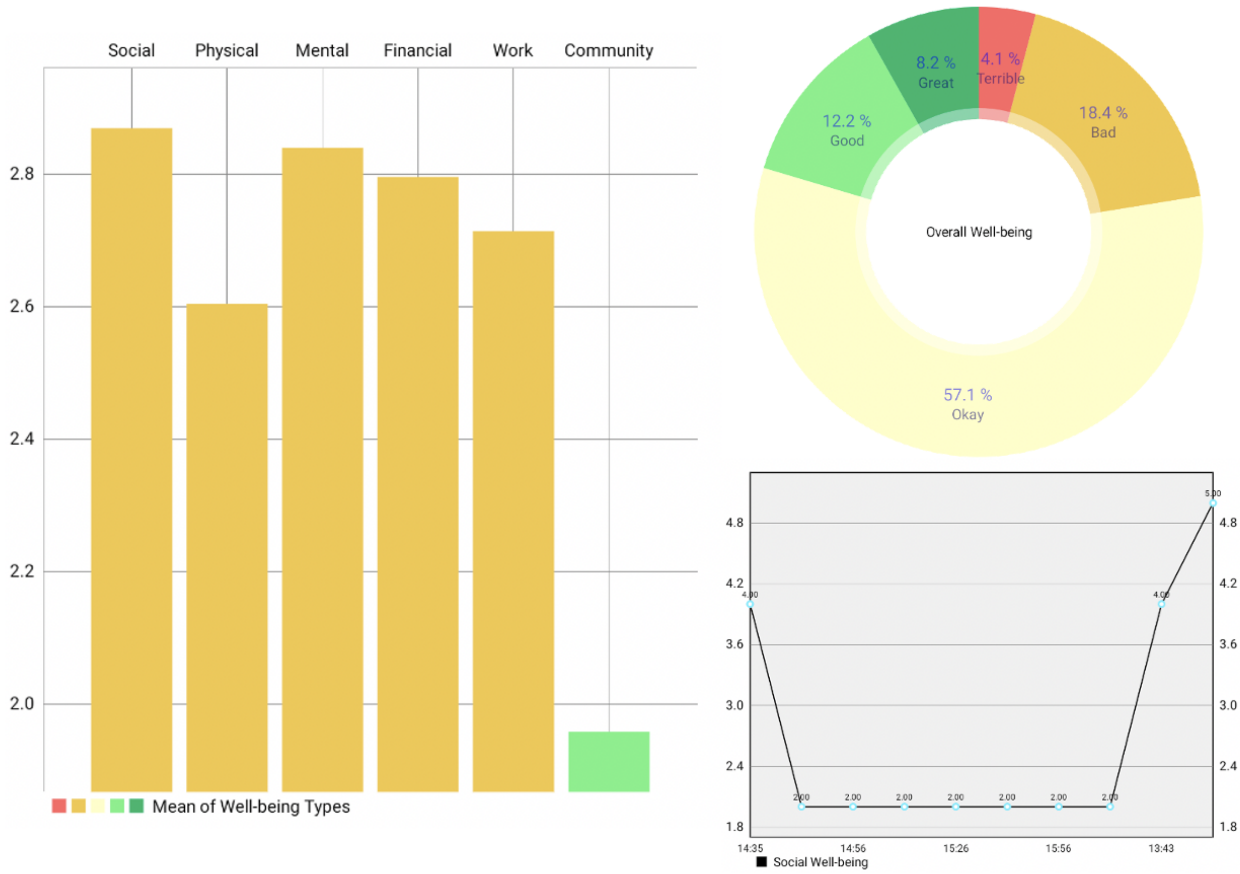


Figure 19: Example of Well-being Dashboard

5

CASE STUDY

In this chapter some experiences will be presented. These can be included in the learning workflow strand which needs to be carried out in order to the smart application features be applicable. In fact, since one of the objectives of this work is to provide some sort of intelligence is necessary to create a model capable to make predictions. In this case, will be used the collected data to train different models. As the architecture suggests this model will be developed in python once this has a lot of libraries oriented to machine learning. These tests were made on the google colab platform. Google Colab is a free cloud service, with GPU support. The virtual machine used have 13 GB of RAM and 107.77 GB of disk and a 2vCPU@2.2GHz processor. Follow this premises, two experiments will be presented. One with a study done with data collected during the quarantine and another with data collected after the quarantine.

5.1 FIRST CASE

The experiment that will be presented next involves the application enunciated before and count with the participation of ten users who used the application during the period from 15 March 2020 to 20 April 2020. It was possible to obtain 1954 lines. The application used by these volunteers was only for data collection purposes and could be inserted in the first strand described before. Users are located in Portugal and due to this factor, the data collection period coincided with a mandatory quarantine period due to the global pandemic.

5.1.1 *Data Analysis*

Once the data was saved in the firebase database, they were in JSON format. In this sense, in order to simplify all the work, it was tried to convert JSON into CSV. A python script was developed for this purpose.

In this sense, since data are the main foundations in artificial intelligence, some investigations were carried out to understand the quantity and quality of the collected information.

In Fig. 20 immediately was noticeable that the dataset had missing values. The bars that are not fully filled means that some of the features have missing values. The exact number of captured data for each feature is above each bar. This can be easily explained by the fact that the Android system has limitations and restrictions on the background collection, especially when the device is in Sleep mode. In fact, due to battery optimization issues, sometimes the operating system ends up blocking some actions (such as access to certain types of sensors) when the device is not in use for long periods of time, resulting in a lack of values. Furthermore, not all phones have the same type of sensors which make impossible to retrieve some type of data. Despite everything, care was taken to combat these restrictions as much as possible. In addition, there were some other features exclusively for devices that were connecting to the internet via WiFi (as is the case with the router information specified above), so for this same reason whenever the device was not connected via WiFi ended then resulting in a lack of information.

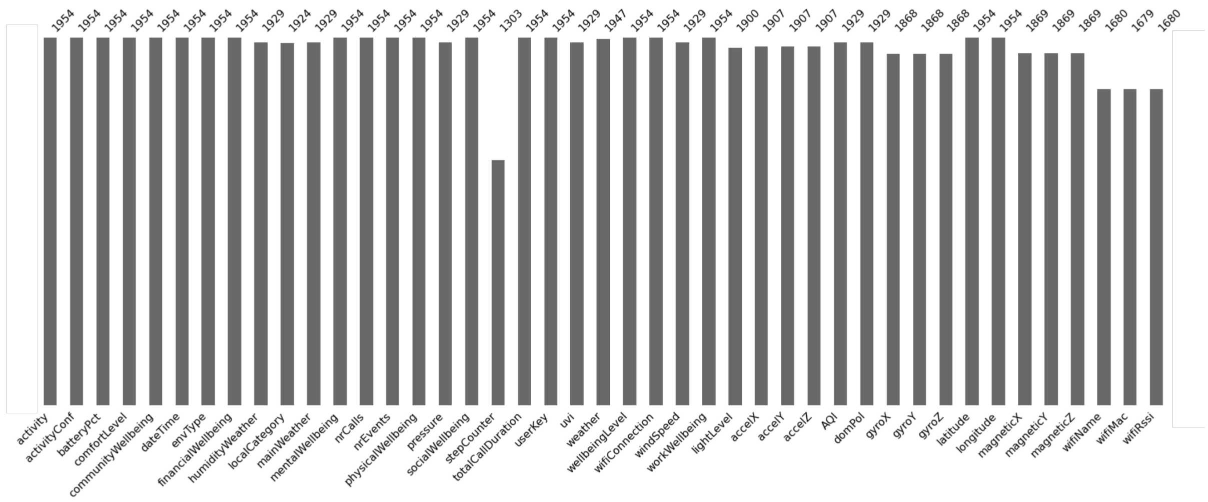


Figure 20: Number of data captured in the different features

From this initial analysis, it was also noticed that 8 of the 43 variables were non-numeric, which would indicate some treatment in the future, and of the remaining (numeric) the intervals they had were quite heterogeneous.

Since there are many features, to understand how they correlate with each other a heatmap was drawn. The results could be seen in the attachment (Appendix D).

A brief analysis was also carried out regarding the classes used by users in the answer they gave to the various questions. This step was important to understand how the dataset was balanced.

Regarding the -1 value, it is expected that it will be the most accentuated class as it concerns the values captured in the background. In addition, users generally tend to avoid extreme values, which is true in this case.

Thus, class 1 has not been chosen practically once. However, in relation to the other classes, it is worth noting that they are not particularly balanced.

For example, in the case of comfort it was noticed that class 1 was not used but in relation to the others, the most accentuated was class number 5 and the least accentuated was class number 2, so these were not balanced in terms of general point of view (Fig. 21). This suggest that latter some balancing techniques should be used.

```
Class=-1, Count=1456, Percentage=74.514%
Class=4, Count=184, Percentage=9.417%
Class=5, Count=253, Percentage=12.948%
Class=2, Count=14, Percentage=0.716%
Class=3, Count=47, Percentage=2.405%
```

Figure 21: Comfort Classes Analysis

In this analysis, the presence of outliers in the data was also studied. This fact may be due to the small amount of data. However, the chosen made was not to remove these outliers as this could cause relevant data to be deleted.

5.1.2 *Model Building*

After an initial analysis of the data, the basic procedures of machine learning and, more particularly deep learning, are used to build predictive models. Such procedures result from the application of frameworks that arise from different studies. The approach followed was the one defined in state of art. This involves, after the data collection, a pre-processing phase. Then the objective is to choose a model and train it, trying to improve it in relation to a base value. Finally is necessary to evaluate the model by making optimizations regarding its hyper-parameters. The studies applied below could be applied to others indicators. For making the analysis more embracing possible, next will be described different studies applied to different indicators.

Traditional Machine Learning Models for Comfort

In order to limit the classifications that could be made by lower and upper intervals, started up by resorting to the use of a *DummyClassifier*. This function provided by the *SKLearn* library in Python allows to establish a baseline accuracy for the dataset used. As baseline accuracy it was possible to obtain a value of 47.28%. So, to start studying how comfort can be predicted, I started by studying the effect that various types of pre-processing could have on the accuracy of a machine learning model, used only for this initial comparison (in this case the base model chosen was SVM). In this first phase, some decisions were made due to the general perception of the problem: “Predict comfort and well-being”.

Thus, the main operations carried out in pre-processing were as follows:

- **Time tag treatment with its transformation into cyclic values:** In this case, the day of the week, month, hour, minutes, and seconds were used. This is due to the fact that over the course of a day, due to the routine we can adopt the values of comfort and well-being can be influenced. This can happen not only by the time of day but also by the day of the week (since, naturally, at the weekend there are always some routine changes);
- **Elimination of some irrelevant variables:** Variables like information related to WiFi have been eliminated as the number of missing values is quite high. The user's key, which is not really relevant to the prediction, was also deleted;
- **Handling of missing values:** Since there were a lot of missing values, I started by using some methods for their treatment. What seemed more adequate was the replacement of these values by the immediately previous value. This was due to the fact that once the data is captured sequentially, this technique could result in less introduction of variability in the dataset. In addition, in the case of some features, such as the case of meteorology, for example, it is expected that they will not have major variations over the course of a day. However, other techniques were also explored here, for example, the use of the mean and the median of the column. Since the number of lines in the dataset is not very high, it was decided not to eliminate lines with missing values in order not to further reduce the sample number;
- **Treatment of non-numeric variables:** Several features of the dataset were represented by strings. Since machine learning models only accept numbers, it was necessary to proceed with some type of pre-processing. For this purpose, the *OneHotEncoder* function was used;
- **Data Normalization:** In order to reduce the discrepancy, some techniques for framing the values were also applied, among them the Standardization and Normalization functions were tested;
- **Target Encoding:** Since the target presents values in a certain way sorted from 1 to 5, a *LabelEnconding* technique was used to normalize these values (thus transforming these values into classes 0 to 4).

Before any procedure, the dataset was separated into training and testing. This division was carried out so that the test set is not biased by the processing performed. It should also be noted that in order to reduce overfitting, cross-validation techniques were used. Of the various variants, the *stratifiedKFold* method was used, which is the most suitable for classification problems, especially when the classes are not balanced. Furthermore, because of the randomness of the process, the function *seed* was used to make the results reproducible.

Then, some tests were carried out (note that the train's accuracy values between the various folds were also noted to detect any overfitting/underfitting problems).

The first step of this study was to understand what missing values treatment experiments was better for the capture data. The results could be seen in Table 1.

Table 1: Missing Values Treatment Experiments

Approach Type	Train Accuracy (Deviation)	Test Accuracy (Deviation)
Using only forward propagation	49.520% (0.247%)	49.200% (2.177%)
Using mean values	49.520% (0.247%)	49.200% (2.177%)
Using median values	49.520% (0.247%)	49.200% (2.177%)

When analyzing the values, it was decided to choose the first technique that assumes the forward propagation of all rows. However, there wasn't, in fact, a difference regarding accuracy in the case of the test.

The next step was to understand which of the techniques, whether standardization or normalization, would allow better accuracy. In this sense, the results are present on Table 2. Each step taken used the best approach for the previous experiments.

Table 2: Values Normalization Experiments

Approach Type	Train Accuracy (Deviation)	Test Accuracy (Deviation)
Standardization (*StandardScaler*)	85.986% (0.344%)	77.847% (2.786%)
Normalization (*MinMaxScaler*)	81.020% (0.485%)	78.156% (3.738%)

Whichever technique is chosen, the results are better than those presented previously. Normalize is required when the features are in different ranges. This could explain the bigger difference in the results because the dataset has various features that are in different ranges. In fact, if we have a feature within larger ranges it will have more influence int the model, however, that does not mean, that this type of feature is more important. Furthermore, these techniques can help to accelerate the optimization of the model. As can be seen from the table above, the results in the case of normalization (*MinMaxScaler*) are better. Although the accuracy in the case of training has dropped significantly about standardization, the accuracy of the test, being the most relevant for decision making, has increased. Also, the difference between training and test accuracy has been reduced, which may indicate a fight against overfitting.

As mentioned initially, users did not choose the same number of classes in the same way. It was then that a resampling technique was also explored so that more values of the missing classes could be generated. For this, there is a function in Python (SMOTE) that allows for obtaining the desired effect. The results obtained using this technique could be seen in Table 3.

Table 3: Resample Experiments

Approach Type	Train Accuracy (Deviation)	Test Accuracy (Deviation)
Without SMOTE	81.020% (0.485%)	78.156% (3.738%)
With SMOTE	87.553% (0.387%)	85.480% (2.638%)

This technique proved to be quite beneficial, as it allowed a high improvement in the accuracy values. This happens because the model will have more data to train for those classes who have fewer cases and for this motive consequently this helps the model generalize better. However, it should also be noted that techniques like these can sometimes lead to overfitting, however, by analyzing the accuracy of training and testing this does not seem to be the case.

Since there a large number of features, it is important to understand which ones best contribute to a better performance of the model. A large number of features resulted essentially from the application of a *OneHotEncoder* process. In this sense, the Table 4 was elaborated where the results are aggregated according to different feature selection methods. The objective is also to try to understand if with a lower number of features it is possible to obtain similar performance.

Table 4: Feature Selection Experiments

Approach Number	Algorithm	Feature Number	Train Accuracy (Deviation)	Test Accuracy (Deviation)
1	f_classif	7	79.888% (0.738%)	79.218% (2.431%)
2	f_classif	10	78.789% (0.417%)	78.257% (2.951%)
3	f_classif	40	86.181% (0.332%)	84.268% (2.498%)
4	f_classif	45	86.816% (0.362%)	84.853% (2.419%)
5	f_classif	50	87.414% (0.387%)	85.354% (2.402%)
6	f_classif	68	87.748% (0.399%)	85.647% (2.738%)
7	Mutual_info_classif	7	78.761% (0.287%)	78.423% (2.207%)
8	Mutual_info_classif	10	80.069% (0.311%)	79.343% (2.485%)
9	Mutual_info_classif	40	88.513% (0.428%)	84.979% (2.919%)
10	Mutual_info_classif	45	87.386% (0.475%)	84.561% (2.784%)
11	Mutual_info_classif	50	88.119% (0.462%)	85.062% (2.279%)
12	Mutual_info_classif	68	87.363% (0.453%)	85.188% (2.979%)
Extra	f_classif	66	87.832% (0.437%)	85.688% (2.686%)
Extra	f_classif	70	87.632% (0.411%)	85.605% (2.674%)
Extra	Mutual_info_classif	66	87.410% (0.434%)	85.313% (2.870%)
Extra	Mutual_info_classif	70	87.581% (0.408%)	85.563% (2.743%)
Best	**f_classif**	**66**	**87.832% (0.437%)**	**85.688% (2.686%)**

With only 66 features is possible to improve a little bit the test accuracy.

The next step was to choose the most suitable model. For this purpose, several Machine Learning models were tested. A visual summary of what has been done can be seen on Fig. 22. This graph shows the spread of the accuracy scores across each cross validation fold for each algorithm.

```
(LR) Test: 0.810563 (0.031158) - Train: 0.828186 (0.003679)
(LDA) Test: 0.812232 (0.030125) - Train: 0.822853 (0.003888)
(KNN) Test: 0.883593 (0.019761) - Train: 0.925385 (0.003982)
(CART) Test: 0.896097 (0.022249) - Train: 1.000000 (0.000000)
(RFC) Test: 0.934911 (0.017358) - Train: 1.000000 (0.000000)
(NB) Test: 0.582650 (0.027162) - Train: 0.586718 (0.003841)
(SGC) Test: 0.812235 (0.027745) - Train: 0.820626 (0.015680)
(SVM) Test: 0.856885 (0.026860) - Train: 0.878316 (0.004368)
```

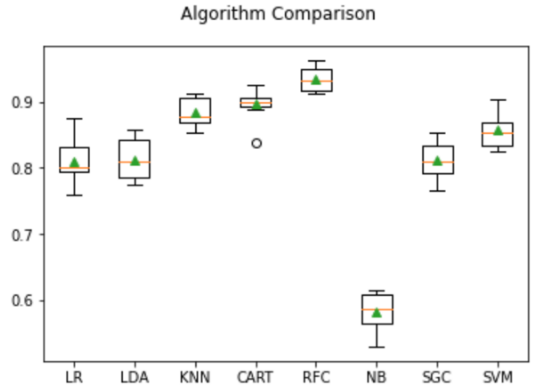


Figure 22: Model Check

In this sense, according to the results presented above, the model that seemed to be the most suitable was then the *RandomForestClassifier*. Although the results are already favorable, the next steps were to try to improve them.

For the model tuning was used a gridsearch technique. This technique is known for a process of performing hyper parameter tuning to determine the optimal values for a given model. It is a expensive process, since this method will generate all the combinations possible for a given set of values (Lutins, 2017). In this case, this technique was used on the model selected before, *RandomForestClassifier*, in order to find the best parameters. The learning curve for the final model could be consulted on Fig. 23

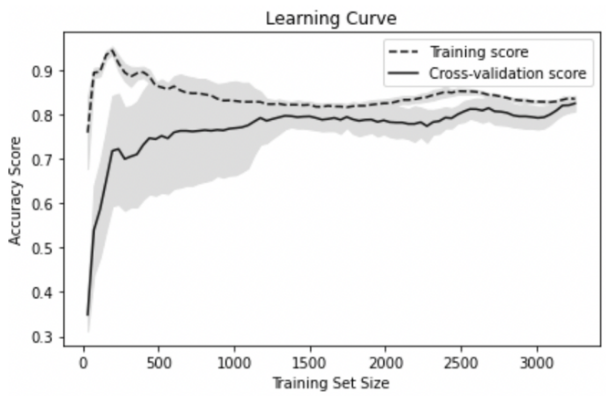


Figure 23: Gridsearch Results

Accuracy has decreased compared to the previous step. However, this technique allows us to find the best hyper-parameters. The fact the default parameters of *RandomForestClassifier* were being used could lead to problems such as underfitting and overfitting. In this case, despite the accuracy that has decreased, is important to note that these problems are also being avoided. As can be seen, there was an approximation between training and test accuracy.

	precision	recall	f1-score	support
0	0.96	0.95	0.96	57
1	0.64	0.84	0.73	86
2	0.67	0.67	0.67	197
3	0.89	0.82	0.85	305
accuracy			0.79	645
macro avg	0.79	0.82	0.80	645
weighted avg	0.80	0.79	0.79	645

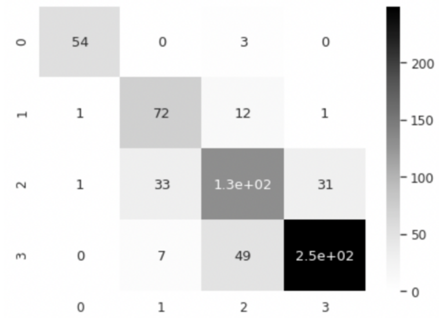


Figure 24: Model Validation

From the report (Fig. 24), it is possible to see that the model generally performs well for all the classes that he was trained for. The major problem is that the model was not trained for class number one. This was because the users, don't choose this class. The main problems with these approaches are that sometimes users tend to evict some extremes values.

This can be explained simply by the fact that in the period when the data was collected, people were at home, due to mandatory confinement. In this sense, it is natural that the lower comfort values have not been so chosen since being at home users can more easily control the environment around them.

Deep Learning Model for Physical Well-being

In the case of well-being, the same premise could be applied. However, I want to study if the previous data are relevant to the prediction. For that case were used RNN, more in particular LSTM. As this is a neural network where the order of the data is quite relevant, no shuffle has been done. Although, other special precautions regarding the way the data had to be processed were taken. In addition to pre-processing, the data were transformed according to a sliding window principle. Thus, the previous data are organized so that they serve as input to predict the target of the next line. A pratical example, with a dataset size of 10 and a window size of 2 could be seen in Fig. 25.

Since the goal is to classify the well-being some precautions have to be taken. As this is a multiclass classification problem (although it can be seen as an ordinal classification problem, this is one of the ways that this type of problem is usually treated), the loss function used was therefore *categorical_crossentropy*. Furthermore, in the final layer, a *softmax* activation

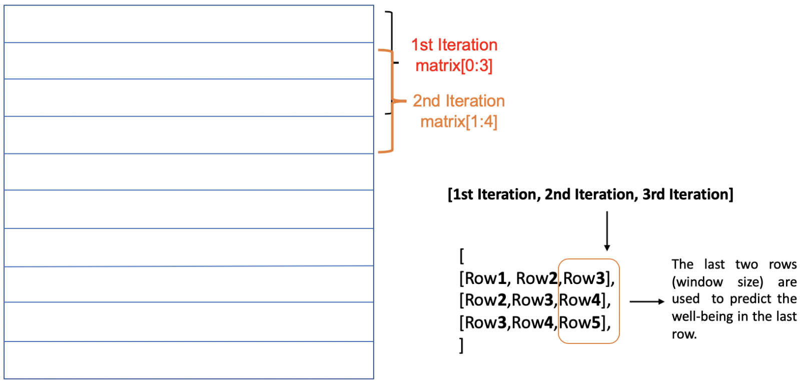


Figure 25: Sliding Window

function was used. Here it is very important to take into account the type of this activation function and the loss function, as incorrect use of these can lead to false results.

Some studies were carried out in order to understand what is the method that could give best results. Essentially two approaches were built. The first one defends that since each user is different and the data they have consequently corresponded to different periods, it was initially hypothesized that it would be better to separate the users. Thus, the goal was to train a neural network for a single user and then try to understand if it was possible to generalize the results. In the second one all the data were used together. Furthermore, this data was ordered by the user key, and for date-time. Next could be summarize this two approaches:

1. **LSTM using data from a single user**
   a) Only one user was used for training the rest were used for testing.
2. **LSTM without separating by user**
   a) There was no separation of each user (only some users are saved for the test);
   b) The transformation into a sliding window occurred in all data, so the data was sorted by user key and date-time in order to reduce inconsistencies.

The following are some of the results obtained in each of the approaches. Similar to what was done in the case of the traditional machine learning models, here too I tried to follow an approach based on the attempt. It is although important to refer that the following experiments were made on physical well-being. However, the study could be extended to other well-being types.

The same process was applied here and again, a study was done on what should be the best treatment for missing values (Table 5).

Table 5: Missing Values Treatment

Missing Values Treatment	Approach Type	Train Accuracy (Loss)	Test Accuracy (Loss)
Using only forward propagation	1	78.96% (0.5802)	64.42% (1.2970)
	2	77.04% (0.7422)	28.70% (1.8702)
Using mean values	1	78.96% (0.6429)	64.42% (1.2372)
	2	41.99% (1.0705)	30.26% (2.0822)
Using median values	1	78.96% (0.5680)	64.42% (1.3350)
	2	41.56% (1.0673)	30.26% (2.0683)

Comparing the two approaches is possible to see that the first one, gives in a first instance better result than the second one. This could be for the fact that the first one, is simpler and the data is less. However, test accuracy does not have big changes in the two approaches.

In these experiments seem like using the mean values to solve the missing values problem is the most appropriate in the first approach (less loss) and the median is better in the second approach.

With the use of values normalization techniques, it was possible to obtain a big improvement in the accuracy of the second approach, as shown in Table 6. However, in the first one, no improvement has been achieved. Especially, in this case, *StandardScaler* technique seems to be causing an overfitting phenomenon, which is fought with *MinMaxScaler*. Despite all, generally, the *MinMaxScaler* technique was the best one, so this was the chosen one. Once again the improvements could be explain for the fact that the input variables have different units witch lead to a unscaled data. This can result in a unstable learning process, as explained before.

Table 6: Value Normalization

Value Normalization	Approach Type	Train Accuracy (Loss)	Test Accuracy (Loss)
Standardization (*StandardScaler*)	1	100% (9.6612e-04)	23.93% (3.3128)
	2	99.57% (0.0147)	65.91% (2.1979)
Normalization (*MinMaxScaler*)	1	83.23% (0.4657)	64.42% (1.2368)
	2	98.97% (0.0265)	88.52% (0.6357)

The next step was to test resampling mechanisms. In this case, resampling involves changing the frequency of the time series observation. This can help to latter define a concrete windows size. The method varies depending on the approach. In the first one, the data was resampling in an interval of 15 minutes because it was the minimum time that notification is thrown. In the case of the second approach, the resampling was done taking into account some considerations. First, the data were grouped by user, after that the resampling was applied for each user individually. This is, for every user was generated samples that correspond to the interval of 15 minutes. However, in this case, with the algorithm studied, this method seems to not contribute for a accuracy. This may happen because the data generated introduces more variability making the process harder to learn. The results could be consulted in Table 7.

Table 7: Resampling

Resampling	Approach Type	Train Accuracy (Loss)	Test Accuracy (Loss)
Without Resampling	1	83.23% (0.4657)	64.42% (1.2368)
	2	98.97% (0.0265)	88.52% (0.6357)
With Resampling	1	99.70% (0.0140)	82.16% (0.9723)
	2	99.61% (0.0133)	62.91% (0.8457)

The final step was to validate and tuning the model. In these approaches the objective was to experiment some combinations in order to find a good fit. The number of layers, the number of neurons, the windows size, epochs, batch size, among others, were tested together. Some of the values were chosen based on some heuristics resulting from the application of neural networks in certain studies. In fact, according to Jeff Heaton on the book Introduction to Neural Networks for Java (Heaton, 2008) the number of hidden neurons should be:

- Between the size of the input layer and the size of the output layer;
- 2/3 the size of the input layer, plus the size of the output layer;
- Less than twice the size of the input layer.

Furthermore, sometimes just one layer is enough for the majority of the problems. However, the best method could be achieve by trial and error since there is no exact formula to known was it the best for some type of problem.

In this sense, I also tried to resort to the alteration of other hyper-parameters to combat this problem. The adjustments made were mainly about the size of the window, the batch size, and the number of epochs. Other numbers of features were also tested. These adjustments were not only applied to obtain better results but also to reduce the instability of the solution and, consequently, reduce any problems as is the case of overfitting that could be happening. In fact, LSTM is a neural network with a propensity for overfitting (Brownlee, 2019), so a good fit is not that one that has high accuracy. Also to avoid this problem, some dropouts layers were been added. Beyond that, some regularizers were explored in order to improve the results. In fact, after the application of some of these techniques, it was possible to improve the results and consequently combat the previous problems. In the first approach, the model seems to converge after a few epochs. The learning curves of the best model could be seen in Fig. 26. The windows size used was 8 and the batch size was 6.

Despite these results, this approach has some problems. The main one is in choosing a user who is the most possible representative of the problem. For example, if user A is used for training and user B is used for testing the model will fail if the classes in B are not the

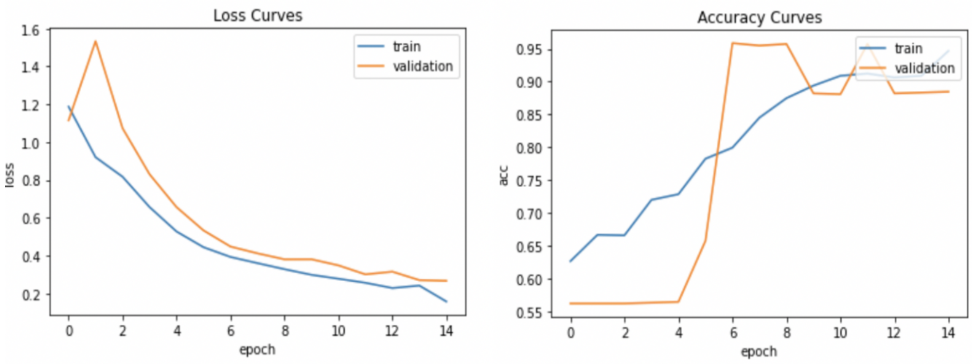


Figure 26: Approach 1: Learning Curves

same as in A (user A has chosen 1 and 2 for physical well-being, and user B has chosen 2 and 3, as the model was not trained in class 3 this could be a problem). Furthermore, each user individually has few rows of data. This makes the model be more fallible in a real scenario since it will depend on that one user used for the training. The next image (Fig. 27) represents two confusion matrices from two different users. The one on the right (Fig. 27 b)) has a higher accuracy than the one on the left (Fig. 27 a)). This could be explained for what was said before.

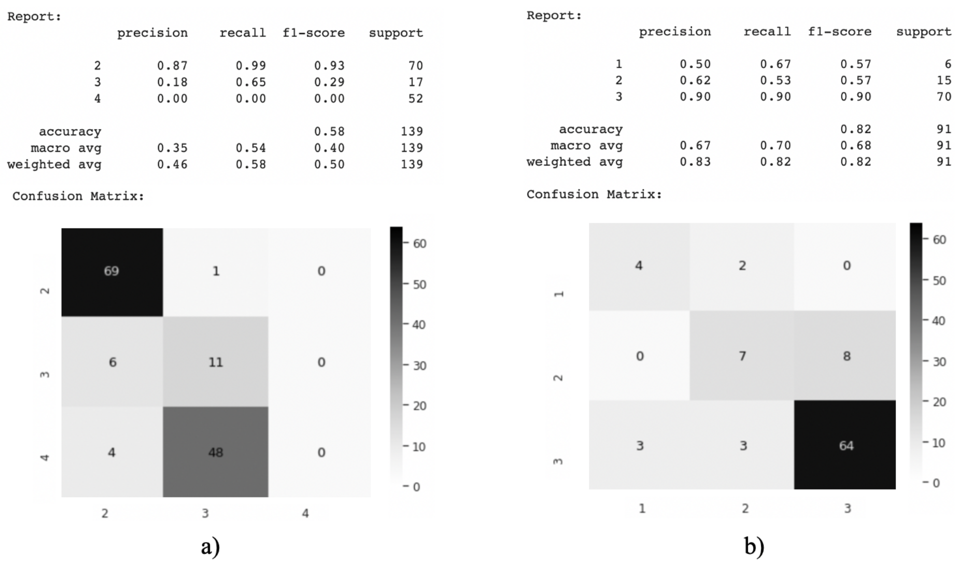


Figure 27: Approach 1: Confusion matrixes

The validation of the model and consequently the final accuracy was obtained by using all the remaining users from the dataset. The mean accuracy obtained was 84.22%. In a search for finding a good model the same steps following before were applied in the second approach. The final results could be seen on Fig. 28.

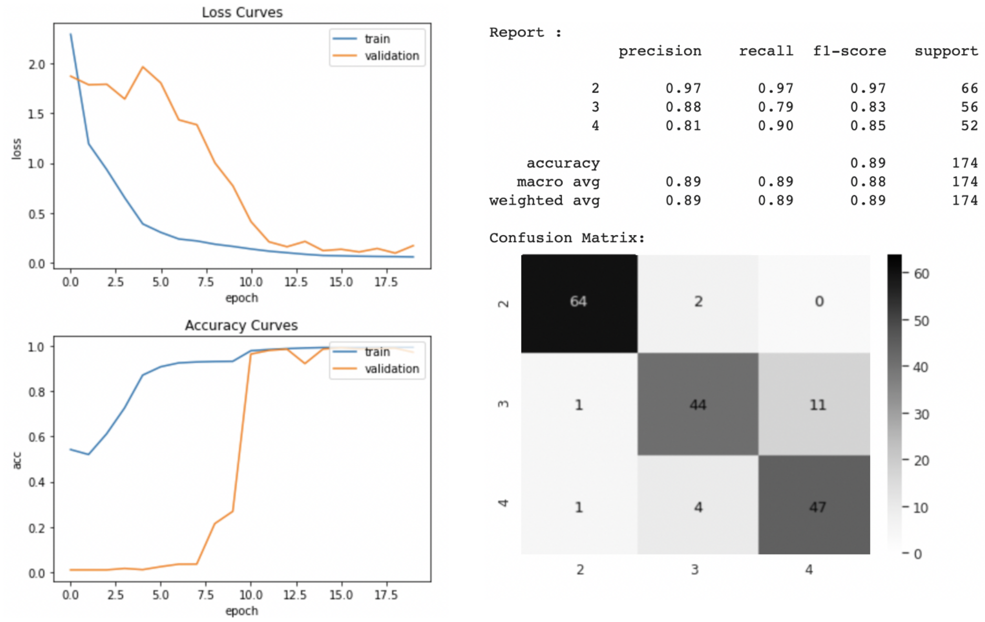


Figure 28: Approach 2: Final Model

The windows size used was 12. The best batch size found was 14. Although this approach would have the same problem as the first one, since the number of data is bigger the probability of that problem happen is lower. Despite all this, this approach has also some other problems. First, by joining the data of all users in the same file, I may be creating some inconsistencies mainly about the creation of sliding windows. Second, the classes are unbalanced, this combined with the impossibility to shuffle the dataset could lead to a leak of performance. However, the second approach seems to be one with better results.

5.2 SECOND CASE

Since the data collected initially focused mainly on a period of mandatory quarantine, a new data collection was carried out after this period. However, due to time and resources, the data collected remains scarce. In this case, the dataset has a total of 2459 lines, in a set of 43 features. In addition, with the experience obtained through the first capture, some improvements were made in the initial application that allows a more efficient data collection. With the experience gained in the first case, the same procedures were applied in this case.

5.2.1 *Data Analysis*

Data analysis occurs in the same way as the first experiment. In that sense, the first step was then to see if there were missing data. A visual graph about the missing values in each feature could be seen in Fig. 29.

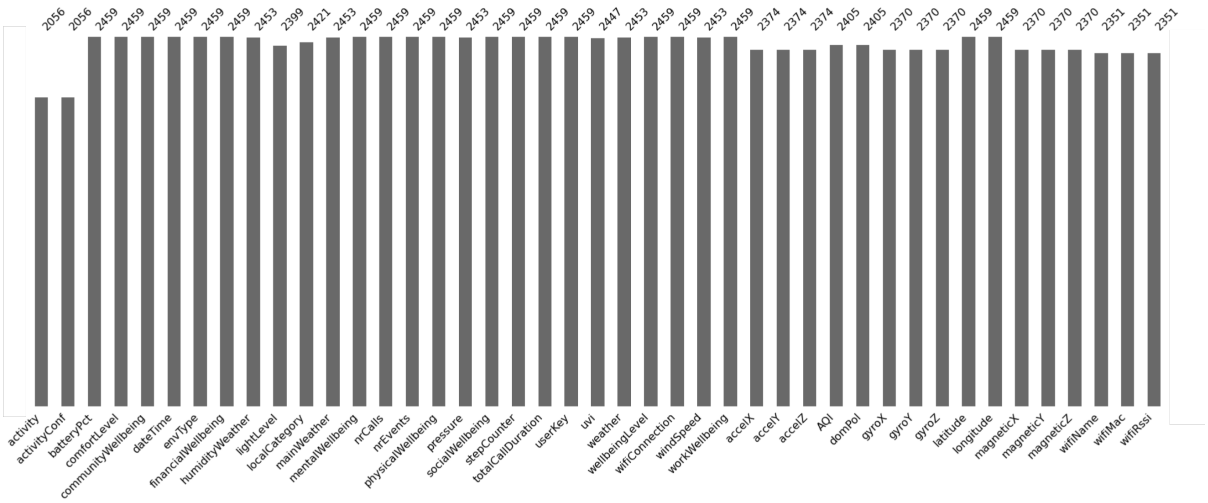


Figure 29: Number of data captured in the different features

In this case it is possible to verify that there are missing values mainly in the case of activity detection. This may have been due to the fact that people have more heterogeneous routines which led to some flaws in the detection of activity.

Once again an heatmap with the relation of all features was drawn. The results could be consult in the appendix for the sake of simplification (Appendix D).

Regarding the balancing of classes, the same happened in this new data collection (Fig. 30). In fact, whether in the classes that concern comfort or in the classes that concern well-being, users still do not use the lower class.

```
Class=-1, Count=1937, Percentage=78.772%
Class=2, Count=7, Percentage=0.285%
Class=4, Count=277, Percentage=11.265%
Class=3, Count=50, Percentage=2.033%
Class=5, Count=188, Percentage=7.645%
```

Figure 30: Comfort Balance

This data still has outliers, however, once again the chosen made was not to remove these outliers as this could cause relevant data to be deleted.

5.2.2 *Model Building*

In this experiment, the same processes, used in the first one, were used. Thus, the following descriptions will be more succinct just stating the results achieved.

Traditional Machine Learning Models

Applying the same principle used previously, here again it is concluded that the algorithm that allows to obtain better results is the Random Forest. According to Fig. 31 it is possible to conclude that in this case, for comfort it is possible to obtain a final accuracy of 81 %.

	precision	recall	f1-score	support
0	0.64	0.91	0.75	23
1	0.47	0.55	0.51	47
2	0.74	0.56	0.64	201
3	0.87	0.92	0.90	541
accuracy			0.81	812
macro avg	0.68	0.74	0.70	812
weighted avg	0.81	0.81	0.81	812

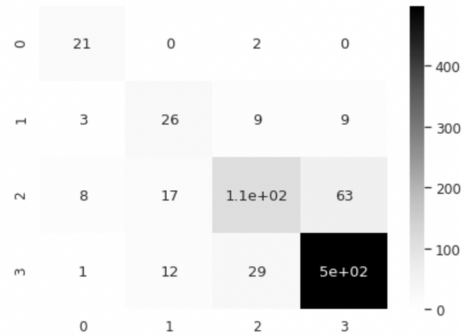
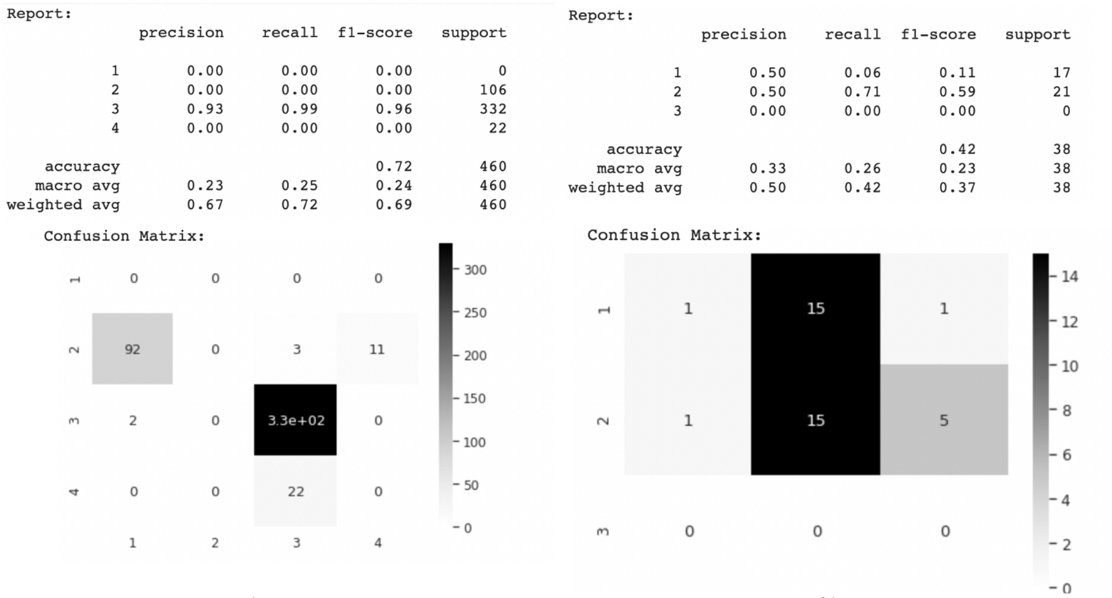


Figure 31: Model Validation

Deep Learning Models

With regard to deep learning models, the process was quite similar. In fact, also in this case were tested the two hypotheses initially proposed. In the first case, the results are not that good. In fact, if I took two different users as a test case, like in the first experiment is possible to see that the results really depend on the classes chosen by those users (Fig. 32). This could be for what is explained in the first case. The train set could not be sufficiently representative of the problem. Using all the test set the mean accuracy obtained is 42.65 %

The second approach still seems to be the best one. In fact, the accuracy obtained by the first approach is not good. By changing several parameters it was possible to achieve some favorable results. Thus, the Fig. 33 concerns a window of 6 and a batch size of 26.


Report:

	precision	recall	f1-score	support
1	0.00	0.00	0.00	0
2	0.00	0.00	0.00	106
3	0.93	0.99	0.96	332
4	0.00	0.00	0.00	22
accuracy			0.72	460
macro avg	0.23	0.25	0.24	460
weighted avg	0.67	0.72	0.69	460

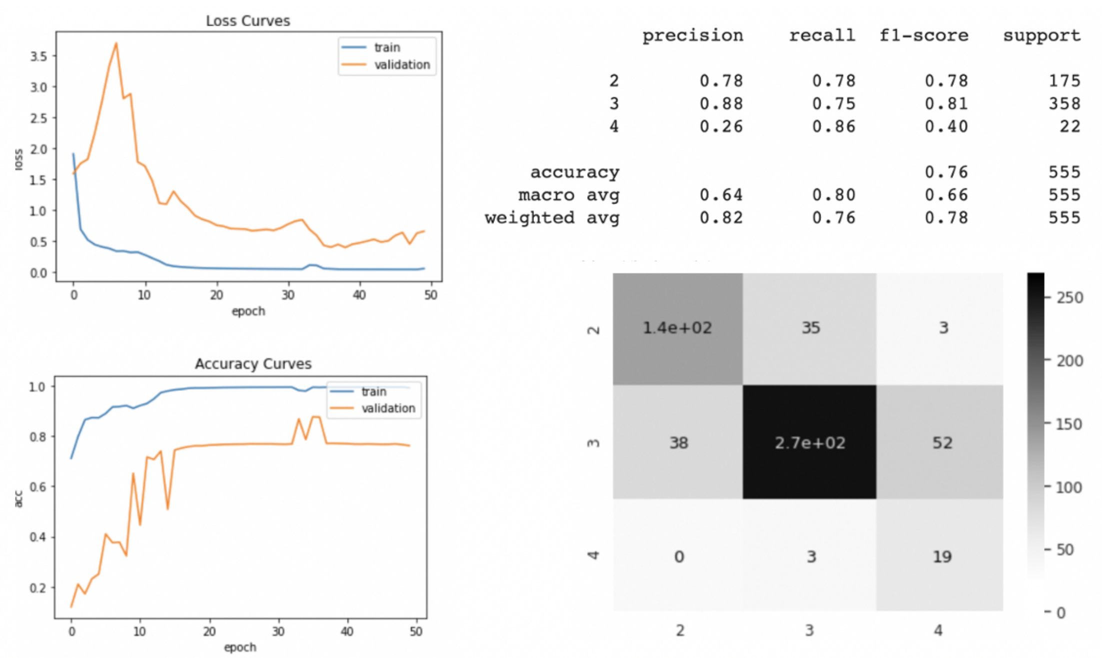
a)

Report:

	precision	recall	f1-score	support
1	0.50	0.06	0.11	17
2	0.50	0.71	0.59	21
3	0.00	0.00	0.00	0
accuracy			0.42	38
macro avg	0.33	0.26	0.23	38
weighted avg	0.50	0.42	0.37	38

b)

Figure 32: Approach 1: Confusion matrixes



	precision	recall	f1-score	support
2	0.78	0.78	0.78	175
3	0.88	0.75	0.81	358
4	0.26	0.86	0.40	22
accuracy			0.76	555
macro avg	0.64	0.80	0.66	555
weighted avg	0.82	0.76	0.78	555

Figure 33: Approach 2: Final Model

5.3 SIMULATION

Using the android studio simulator it was possible to simulate how the application would behave in a real scenario. The simulator allows using a virtual device on a computer. Thus, it provides the possibility to change the values of base sensors of a device making it possible to create situations that can easily be felt in a real case.

Since it is hard in the context of this work to get the help of multiple users distributed by the country, it was used the location change capacities of the simulator to create a map with data from different regions. So, instead of using a real physical location, the simulator was used to make it look like the user was in several other parts of the country.

So, being as comprehensive as possible, I went, on the one hand, to answer some questions related to comfort and well-being, and, on the other, I waited that some predictions would be made in order to have the most heterogeneous map possible. This simulation was created after the application was fully developed and aimed to understand how it would behave when used by users in different areas. The results can be seen on Fig. 34 and these will be discussed in the next sub-section.

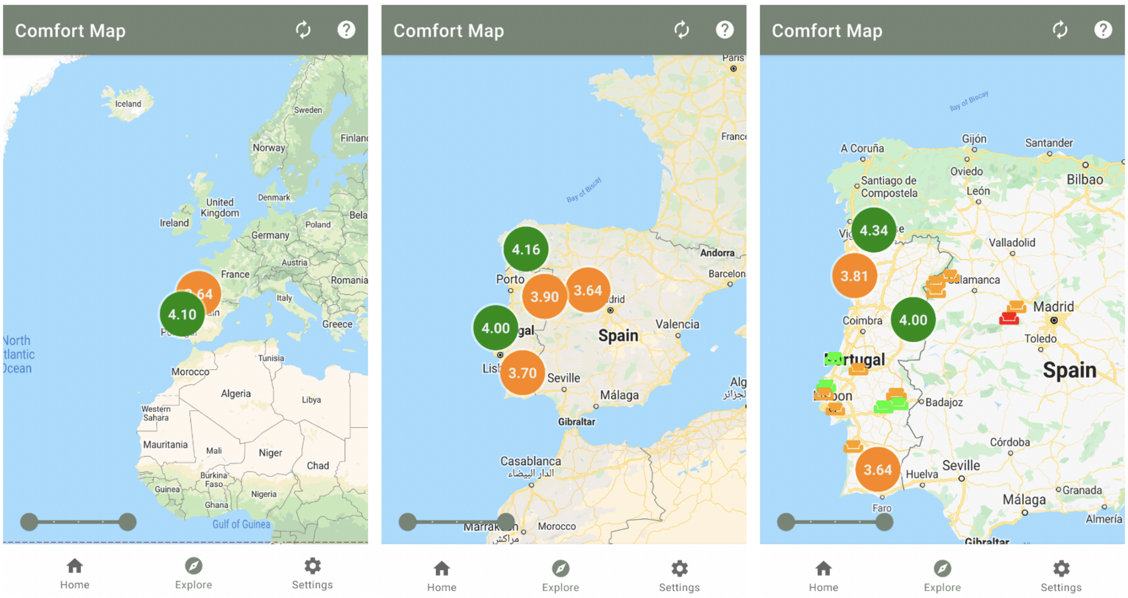


Figure 34: Simulated Map

5.4 RESULTS ANALYSIS

In general, the deep learning models with data after the quarantine have lower performance when compared to the models of the same subject but with the data collected during the quarantine. This may be due to the fact that the data are more heterogeneous, since people could already leave the house freely, and thus, the model had more difficulty in generalizing. However, it is not possible to make a direct comparison as the conditions are not exactly the same as the initial conditions. In fact, there has been an improvement in the way that data is being captured. The initial application has thus undergone some changes resulting from the first attempt to collect data, making the collected data effectively more refined. In fact, the data collection carried out in the second experiment already consisted of a more stable version of the application.

It was expected that the data captured after the demanded quarantine have more frequency in high classes with some types of well-beings like for example, physical well-being. However, that is not verified. This can happen because of the low number of user's. Consequently, the results obtained with the deep learning models are not as reliable as the data has to maintain a logical organization and only a few records from the selected users data are evaluated.

Furthermore, the fact that enough attributes have been collected allows a high number of studies to be carried out. Through the simple exchange of targets or the features involved, other outputs can be explored and, consequently, other studies can be presented. Although only some of the studies carried out are represented above, many others can be aggregated following the same procedures.

In this section, two case studies were presented, in both, comfort and physical well-being were the targets. However, since, in the case of comfort, only the traditional machine learning algorithms were used and in the case of physical well-being, only the LSTM was used, the following table shows the results through the application of the same previous procedures using different targets. Thus, the first two lines of the Table 8 refer to the first case study, that is, with data captured during the quarantine, while the last two lines show results associated with the data from the second case study, captured after the quarantine. In these cases, it is then tested whether using the LSTM in the case of comfort and the traditional algorithms in the case of physical well-being if is possible to obtain improved results. As a general case, traditional machine learning algorithms shows better results. In fact, according to the results obtained, traditional machine learning models are usually sufficient to respond to the problem. Effectively, in both cases, *RandomForestClassifier* can prove that is possible to predict comfort and well-being using data captured by the smartphone. The full results, associated to this study can be consulted in the Appendix E.

Regarding the simulation performed, the results came as expected. It was possible to see, with the zoom at the country level, the results aggregated for that country. Every time a zoom-in is made, the clusters spread out, and it is possible to see the results by region. If we continue zooming in, the cluster spread until only a unique marker is seen. The same thing happens on the web map. In this way is possible to see with this simulation that the behavior is the one idealized. As expected, the markers from different countries do

Table 8: Comparation between the application of traditional machine learning algorithms and deep learning

Data	Target	Selected Model	Test Accuracy
Captured during the quarantine	Comfort	DL-LSTM	59%
Captured during the quarantine	Physical Well-being	ML-Random Forest	88%
Captured after the quarantine	Comfort	DL-LSTM	61%
Captured after the quarantine	Physical Well-being	ML-Random Forest	76%

not mix. This makes it possible to take a more concrete view and even compare the level of comfort and well-being in each country (Fig. 34). However, this simulation has shown some weaknesses. For one side, if for some reason the place doesn't have a postal code it is impossible to assert a comfort or well-being value to that place. On the other side, it is important that Foursquare API can identify correctly the place category since this will have implications in user advice.

6

CONCLUSIONS

In this chapter, a description of all the work will be done in a first instance, it will also be present a scientific contribution made. Finally, some aspects that can be improved and some ideas for future work will be discussed.

6.1 WORK SUMMARY

Throughout the work, the objectives initially defined were always present. In this sense, the fulfillment of these was the main priority. However, throughout the development, some restructuring was necessary to organize and make the entire study more coherent. The work done can thus be divided, in more general terms, into two parts. A more theoretical part, where the main focus was on the perception of the theme and another more practical with the application of some of the concepts learned.

Regarding theoretical development, I started by collecting relevant studies for the topic in question. Initially, the main difficulty was in defining the state of the art, mainly in defining the concepts of comfort and well-being. These are subjective terms and there was a need to explore a large number of articles to reach some kind of consensus. At this stage, the existence of projects similar to the one proposed was also one of the main focuses, as, in a way, they would serve as inspiration.

After clarifying concepts and analyzing the related work, I moved to a more practical environment with the definition and construction of the project. Thus, an architecture was defined that could be followed to respond to the main objectives. The purpose of this step was to establish an overview of the project and to present some of the technologies that were used to build the application initially proposed. So, in this step, it was possible to develop an application capable of not only collecting data but also predict comfort and well-being. The main challenge was the heterogeneity existing between the various devices of different users, which ended up making the whole process a little harder. Also, the fact that it is an academic project makes a collection with a high number of participants more complicated and highly expensive. However, as intended, the application continuously has collected data

associated with different users that were then used to create the different machine learning models.
The challenge here was also to understand which would be the most suitable algorithms and which type of strategy would be the most reliable. Indeed, the fact that there are different dimensions of well-being made this process and the whole application more complex and implied a high number of possibilities for the creation of machine learning models. It is in the case study part that some possibilities were explored.
In fact, some experiences were made. The first one happens in a period of mandatory quarantine. That leads to the necessity to make a second data captured that happens after this period. This was made for captured data more heterogeneous and that fits a normal context. However, is important to refer that despite this was made, the global pandemic influences always the people's quotidian, so getting a "normal" context is not possible.

Once these models were built, it was possible to create an application capable of predicting comfort and well-being and with the data captured by user's responses to various questions, and the predicted data, it was possible to create a map that allows consulting comfort and well-being, by location and region. Not only is this visualization possible to be consulted through the android application, but also it is possible to be viewed through a web page created for this purpose.
The difficulty in this phase was to decide what was the most appropriate strategy to guarantee a correct and efficient visualization of the values. In fact, the choice of an adequate average calculation mechanism was one of the main challenges of this phase. In addition, at this stage, it was still possible to build some dashboards that gave the user a way of consulting his comfort and well-being over time.

Therefore, all the objectives initially established have been met. Data were captured using a mobile application, thus responding to the first sub-objective. Then, with the data captured, models of machine learning and in particular also of deep learning were built, which mainly encompassed physical well-being and comfort. Thus, it can be concluded that the second sub-objective has also been completed. After these models were elaborated, it was then possible with their use to predict well-being and comfort, which allowed generating intelligent advice through suggestions for changing location and conditions. In this sense, the third sub-objective was also achieved. And finally, a map was created where it is possible to check comfort and well-being. In fact, all the proposed objectives were achieved, and as a final result, it is possible to evaluate, monitor and predict comfort and well-being.

In short, a prototype was built with favorable results, given the conditions felt, and with all the initially established premises achieved.

6.2 Relevant Work

As mentioned in the presented work plan, a scientific article was developed in parallel with the project. This in turn was published and consequently accepted by the 21st International Conference on Intelligent Data Engineering and Automated Learning (IDEAL):

- Sousa D., Silva F., Analide C. (2020) Learning User Comfort and Well-Being Through Smart Devices. In: Analide C., Novais P., Camacho D., Yin H. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2020. IDEAL 2020. Lecture Notes in Computer Science, vol 12489. Springer, Cham. https://doi.org/10.1007/978-3-030-62362-3_31

As relevant work in the article, some aspects discussed in the case of machine learning models, also presented here, are highlighted. This fact allowed not only to deepen scientific concepts and approaches but also allowed to disseminate the project to the scientific community.

6.3 Future Work

Although the work satisfies all the conditions initially defined, there is always room for improvement. The first aspect that can be improved is the machine learning algorithms developed. In fact, with a new data collection and with access to a greater number of users, the results can be improved by making the predictions more accurate. Furthermore, the fear of the population to leave home could influence some of the results.
In this field, there is also the possibility of exploring different algorithms and alternatives. In addition, there may be an adjustment of the features to be captured through access to a larger number of data. An example that can be given is the collection of data related to fitness conditions captured by other applications, such as Google Fit. However, data of this nature, mainly from large companies such as Google, required some bureaucracies related to the privacy of users. Furthermore, some of the Apis used has some limitations because of the fact that were been used the free services. An interesting improvement is trying to use more reliable services.

Another aspect to be improved is some details in the application related to the way that advice was provided. The fact that the same set of advice is being provided regardless of the level and type of well-being can cause advice of little relevance, depending on what is being assessed. This was one of the main issues discussed and needs to be improved in the future. A suggestion that is given is to understand the most suitable types of advice according to user feedback. Thus, through this feedback it is possible to improve this part of the application in the long term, giving more and more appropriate advice. The study

provided here is a study of open discovery in which it is intended to study what may or may not have implications for comfort and well-being. Effectively, comfort and well-being, as already mentioned, are quite complex terms and are influenced by quite different factors that in some parts are difficult to perceive using only with the use of intelligent devices. However, the present work provides a possible way of perceiving other factors that are sometimes not taken into account when changing comfort and well-being. It is these factors and the way they are captured that in future work can be improved.
Still, regarding the advice, one possible idea is also to make filters to the map records to give advice more related to the context felt. For example, by filtering, the records by time and local category are possible in the lunch hour advice the user to the nearest, comfortable restaurant. However, this will only be feasible in the future with more users and consequently with more data. In fact, with a high number of filters, the results will be less and the probability to find a location with all the conditions is lower. In addition, over time, it will be possible to improve in the formula how the average is calculated and how the data is demonstrated so that this fact reflects better the real world.

In short, despite everything, there are always changes that can be made in the future in order to improve the whole system. Also, the extension, for example, of the application to a more business domain may be something that arouses some interest, giving to the work a whole new scope. In this sense, ”the door is open” for improvements.

BIBLIOGRAPHY

Tracking the Transit Rider Experience: Using Smartphones to Measure Comfort and Well-Being Throughout the Trip, Washington, D.C., 01/2015 2015. Transportation Research Board. URL https://trid.trb.org/view.aspx?id=1339490.

Indice de bem estar, 2017. URL https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=315227706&DESTAQUESmodo=2. Accessed: 2019-10-10.

AFSCME. Factors that affect air quality, 2019. URL https://www.afscme.org/news/publications/workplace-health-and-safety/safe-jobs-now-a-guide-to-health-and-safety-in-the-workplace/chapter-9-indoor-air-quality/factors-that-affect-air-quality. Accessed: 2020-01-13.

Pragya Agarwal. How does lighting affect mental health in the workplace, 12 2018. URL https://www.forbes.com/sites/pragyaagarwaleurope/2018/12/31/how-does-lighting-affect-mental-health-in-the-workplace/. Accessed: 2020-02-22.

AirNow. Air quality index (aqi) basics, 2019. URL https://airnow.gov/index.cfm?action=aqibasics.aqi. Accessed: 2019-11-28.

Elena Alatartseva and Galina Barysheva. Well-being: Subjective and objective aspects. *Procedia - Social and Behavioral Sciences*, 166:36 – 42, 2015. ISSN 1877-0428. doi: https://doi.org/10.1016/j.sbspro.2014.12.479. URL http://www.sciencedirect.com/science/article/pii/S1877042814066178. Proceedings of The International Conference on Research Paradigms Transformation in Social Sciences 2014 (RPTSS-2014).

Vito Albino, Umberto Berardi, and Rosa Dangelico. Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology*, 22:2015, 02 2015. doi: 10.1080/10630732.2014.942092.

Asier Aztiria, Alberto Izaguirre, and Juan Augusto Wrede. Learning patterns in ambient intelligence environments: A survey. *Artif. Intell. Rev.*, 34:35–51, 06 2010. doi: 10.1007/s10462-010-9160-3.

Michael R. Berthold, Nicolas Cebron, Fabian Dill, Thomas R. Gabriel, Tobias Kötter, Thorsten Meinl, Peter Ohl, Kilian Thiel, and Bernd Wiswedel. Knime - the konstanz information miner: Version 2.0 and beyond. *SIGKDD Explor. Newsl.*, 11(1):26–31, November 2009. ISSN 1931-0145. doi: 10.1145/1656274.1656280. URL https://doi.org/10.1145/1656274.1656280.

Jason Brownlee. How to use weight regularization with lstm networks for time series forecasting, 2019. URL https://machinelearningmastery.com/use-weight-regularization-lstm-networks-time-series-forecasting/. Accessed: 2020-05-15.

CDC. Well-being concepts, 2018. URL https://www.cdc.gov/hrqol/wellbeing.htm. Accessed: 2019-10-15.

Rohit Chaudhri, Waylon Brunette, Mayank Goel, Rita Sodt, Jaylen VanOrden, Michael Falcone, and Gaetano Borriello. Open data kit sensors: Mobile data collection with wired and wireless sensors. In *Proceedings of the 2nd ACM Symposium on Computing for Development*, ACM DEV '12, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450312622. doi: 10.1145/2160601.2160614. URL https://doi.org/10.1145/2160601.2160614.

Thais C. Morata Chuck Kardous, Christa L. Themann and W. Gregory Lotz. Understanding noise exposure limits: Occupational vs. general environmental noise, 2016. URL https://blogs.cdc.gov/niosh-science-blog/2016/02/08/noise/. Accessed: 2019-11-27.

Diane Cook, Juan Augusto Wrede, and Vikramaditya Jakkula. Review: Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5:277–298, 08 2009. doi: 10.1016/j.pmcj.2009.04.001.

UMass Dartmouth. Decision-making process, 05 2018. URL https://www.umassd.edu/fycm/decision-making/process/. Accessed: 2019-11-04.

Tchiki Davis. What is well-being? definition, types, and well-being skills, 2019. URL https://www.psychologytoday.com/us/blog/click-here-happiness/201901/what-is-well-being-definition-types-and-well-being-skills. Accessed: 2019-10-10.

Rachel Dodge, Annette P. Daly, Jan Huyton, and Lalage D. Sanders. The challenge of defining wellbeing. *International Journal of Wellbeing*, 2, 2012. doi: 10.5502/ijw.v2i3.4. URL https://internationaljournalofwellbeing.org/ijow/index.php/ijow/article/view/89.

Quan Z. Sheng Edward Curry, Schahram Dustdar and Amit Sheth. Smart cities – enabling services and applications. *Journal of Internet Services and Applications*, 7, 2016. doi: https://doi.org/10.1186/s13174-016-0048-6.

Emma Ellis. Diagnosing (and dealing with) your low-battery anxiety, 06 2019. URL https://www.wired.com/story/diagnosing-and-dealing-with-your-low-battery-anxiety/. Accessed: 2020-02-22.

Mental Health Foundation. What is wellbeing, how can we measure it and how can we support people to improve it?, 2015. URL https://www.mentalhealth.org.uk/blog/what-wellbeing-how-can-we-measure-it-and-how-can-we-support-people-improve-it. Accessed: 2019-11-17.

Jake Frankenfield. Artificial neural networks (ann) defined, 2018. URL https://www.investopedia.com/terms/a/artificial-neural-networks-ann.asp. Accessed: 2019-11-20.

Bruno S Frey. *Happiness and economics*. Princeton University Press, 2002. ISBN 9780691069982.

Michael Garbade. Clearing the confusion: Ai vs machine learning vs deep learning differences, 2018a. URL https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb. Accessed: 2020-01-26.

Michael Garbade. Understanding k-means clustering in machine learning, 2018b. URL https://towardsdatascience.com/understanding-k-means-clustering-in-machine-learning-6a6e67336aa1. Accessed: 2020-01-26.

Sebastian Guenther. What is pmv? what is ppd? the basics of thermal comfort, 2019. URL https://www.simscale.com/blog/2019/09/what-is-pmv-ppd/. Accessed: 2019-11-22.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: An update. *SIGKDD Explor. Newsl.*, 11(1):10–18, November 2009. ISSN 1931-0145. doi: 10.1145/1656274.1656278. URL https://doi.org/10.1145/1656274.1656278.

Barry P. Haynes. *The impact of office comfort on productivity*. Emerald Publishing Limited, 02 2008.

Health and Safety Executive. Thermal comfort: The six basic factors, 2019. URL https://www.hse.gov.uk/temperature/thermal/factors.htm. Accessed: 2019-11-22.

Jeff Heaton. *Introduction to Neural Networks for Java, 2nd Edition*. Heaton Research, Inc., 2nd edition, 2008. ISBN 1604390085.

Dominique Heger. Big data analytics—where to go from here. *International Journal of Knowledge, Innovation and Entrepreneurship*, 2:38–54, 2014. URL http://ijkie.org/IJKIE_December2014_DOM%20HEGER.pdf.

Mark Hung. Leading the iot, 2017. URL https://www.gartner.com/imagesrv/books/iot/iotEbook_digital.pdf. Accessed: 2019-10-6.

Mohsen Joshanloo, Parviz Rastegar, and Ali Bakhshi. The big five personality domains as predictors of social wellbeing in iranian university students. *Journal of Social and Personal Relationships*, 29(5):639–660, 2012. doi: 10.1177/0265407512443432. URL https://doi.org/10.1177/0265407512443432.

Radovan Kuchta. Smart city concept, applications and services. *Journal of Telecommunications System Management*, 03, 01 2014. doi: 10.4172/2167-0919.1000117.

Nicholas Lane, Mashfiqui Rabbi, Mu Lin, Xiaochao Yang, Hong lu, Shahid Ali, Afsaneh Doryab, Ethan Berke, Tanzeem Choudhury, and Andrew Campbell. Bewell: A smartphone application to monitor, model and promote wellbeing. *Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare*, 01 2011. doi: 10.4108/icst.pervasivehealth.2011.246161.

Evan Lutins. Grid searching in machine learning: Quick explanation and python implementation, 2017. URL https://medium.com/@elutins/grid-searching-in-machine-learning-quick-explanation-and-python-implementation-550552200596. Accessed: 2020-05-15.

Calum Maclean. Thermal comfort and wellbeing, 01 2017. URL https://www.bsria.com/uk/news/article/thermal-comfort-and-wellbeing/. Accessed: 2020-01-10.

Cory Maklin. Mean shift clustering algorithm, 2018. URL https://towardsdatascience.com/machine-learning-algorithms-part-13-mean-shift-clustering-example-in-python-4d6452720b00. Accessed: 2020-01-26.

Stephen Marsland. *MACHINE LEARNING An Algorithmic Perspective*. CRC Press, 2015. ISBN 978-1-4665-8333-7.

Aditi Mittal. Understanding rnn and lstm, 2019. URL https://towardsdatascience.com/understanding-rnn-and-lstm-f7cdf6dfc14e. Accessed: 2020-01-26.

Amanda Iglesias Moreno. Moving averages with python, 2020. URL https://towardsdatascience.com/moving-averages-in-python-16170e20f6c. Accessed: 2020-08-28.

Chris Nicholson. A beginner's guide to generative adversarial networks (gans), 2019. URL https://pathmind.com/wiki/generative-adversarial-network-gan. Accessed: 2020-01-26.

James J. Park, Qun Jin, Martin Sang-soo Yeo, and Bin Hu. *Human Centric Technology and Service in Smart Space: HumanCom 2012*. Springer Publishing Company, Incorporated, 2012. ISBN 9400750854.

Jamie Pearce and Sarah Curtis. Why places matter for our wellbeing, 04 2017. URL https://cresh.org.uk/2017/04/24/why-places-matter-for-our-wellbeing/.

Tom Rath and Jim Harter. *The economics of Wellbeing*. Gallup press, 2010.

Sunil Ray. Commonly used machine learning algorithms, 2017. URL https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/. Accessed: 2020-01-26.

Sebastiaan Rothmann. Job satisfaction, occupational stress, burnout and work engagement as components of work-related wellbeing. *SA Journal of Industrial Psychology*, 34:11 – 16, 00 2008. ISSN 2071-0763. URL http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2071-07632008000300002&nrm=iso.

Angela Sanguinetti, Marco Pritoni, Kiernan Salmon, and Joshua Morejohn. Thermoostat: Occupant feedback to improve comfort and efficiency on a university campus. 08 2016.

Warwick Medical School. About wemwbs, 2019. URL https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/about/. Accessed: 2019-11-17.

NoAA SciJinks. How is air quality measured?, 2019. URL https://scijinks.gov/air-quality/. Accessed: 2019-11-28.

Fábio Silva and Cesar Analide. Tracking context-aware well-being through intelligent environments. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 4(2), 2015. ISSN 2255-2863. URL http://campus.usal.es/~revistas_trabajo/index.php/2255-2863/article/view/ADCAIJ2015426172.

Fábio Silva and Cesar Analide. Computational sustainability and the phess platform: Using affective computing as social indicators. *Future Generation Computer Systems*, pages 329–341, 10 2018. doi: https://doi.org/10.1016/j.future.2018.10.006.

Simplilearn. Deep learning algorithms you should know about, 2019. URL https://www.simplilearn.com/deep-learning-algorithms-article. Accessed: 2020-01-26.

Emma Smith. Dal researchers create app that monitors mental health by measuring user's emotions, 2020. URL https://www.cbc.ca/news/canada/nova-scotia/dalhousie-university-researcher-prosit-study-smartphone-app-mental-health-treatment-1.5653705. Accessed: 2020-08-22.

Debbie L Stoewen. Dimensions of wellness: Change your habits, change your life. 58, 08 2017. Accessed: 2020-07-28.

Dataflair Team. 11 top machine learning algorithms used by data scientists, 2019. URL https://data-flair.training/blogs/machine-learning-algorithms/. Accessed: 2020-01-26.

Expert System Team. What is machine learning? a definition, 2017. URL https://expertsystem.com/machine-learning-definition/. Accessed: 2019-11-20.

Engineering ToolBox. lluminance - recommended light level, 2004. URL https://www.engineeringtoolbox.com/light-level-rooms-d_708.html. Accessed: 2019-11-25.

Claudia Trudel-Fitzgerald, Rachel A. Millstein, Chanelle J. Howe Christiana von Hippel, Linda Powers Tomasso, Gregory R. Wagner, and Tyler J. VanderWeele. Psychological well-being as part of the public health debate? insight into dimensions, interventions, and policy. 2019. doi: https://doi.org/10.1186/s12889-019-8029-x.

Elizabeth Tutton and Kate Seers. An exploration of the concept of comfort. *Journal of Clinical Nursing*, 12(5):689–696, 2003. doi: 10.1046/j.1365-2702.2003.00775.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2702.2003.00775.x.

Mark van Rijmenam. 7 steps to machine learning: How to prepare for an automated future, 2019. URL https://medium.com/dataseries/7-steps-to-machine-learning-how-to-prepare-for-an-automated-future-78c7918cb35d. Accessed: 2020-04-30.

Peter Vink and Susan Hallbeck. Editorial: Comfort and discomfort studies demonstrate the need for a new model. *Applied ergonomics*, 43:271–6, 07 2011. doi: 10.1016/j.apergo.2011.06.001.

Katrina Wakefield. A guide to machine learning algorithms and their applications, 2019. URL https://www.sas.com/en_ie/insights/articles/analytics/machine-learning-algorithms.html. Accessed: 2020-01-26.

WHO. Ultraviolet radiation (uv), 2020. URL https://www.who.int/uv/intersunprogramme/activities/uv_index/en/index3.html. Accessed: 2020-04-30.

Mark R. Young, Eve Rapp, and James W. Murphy. Action research: enhancing classroom practice and fulfilling educational responsibilities. *Journal of Instructional Pedagogies*, 06 2010. URL http://www.aabri.com/manuscripts/09377.pdf.

Hui Zhang, Edward Arens, and Wilmer Pasut. Air temperature thresholds for indoor comfort and perceived air quality. *Building Research & Information*, 39(2):134–144, 2011. doi: 10.1080/09613218.2011.552703.

A

REFERENCE VALUES FOR LIGHT

Condition	Illumination	
	(ftcd)	*(lux)*
Sunlight	10000	107527
Full Daylight	1000	10752
Overcast Day	100	1075
Very Dark Day	10	107
Twilight	1	10.8
Deep Twilight	0.1	1.08
Full Moon	0.01	0.108
Quarter Moon	0.001	0.0108
Starlight	0.0001	0.0011
Overcast Night	0.00001	0.0001

Figure 35: Outdoor light reference levels (ToolBox, 2004)

Activity	Illuminance *(lx, lumen/m^2)*
Public areas with dark surroundings	20 - 50
Simple orientation for short visits	50 - 100
Areas with traffic and corridors - stairways, escalators and travelators - lifts - storage spaces	100
Working areas where visual tasks are only occasionally performed	100 - 150
Warehouses, homes, theaters, archives, loading bays	150
Coffee break room, technical facilities, ball-mill areas, pulp plants, waiting rooms,	200
Easy office work	250
Class rooms	300
Normal office work, PC work, study library, groceries, show rooms, laboratories, check-out areas, kitchens, auditoriums	500
Supermarkets, mechanical workshops, office landscapes	750
Normal drawing work, detailed mechanical workshops, operation theaters	1000
Detailed drawing work, very detailed mechanical works, electronic workshops, testing and adjustments	1500 - 2000
Performance of visual tasks of low contrast and very small size for prolonged periods of time	2000 - 5000
Performance of very prolonged and exacting visual tasks	5000 - 10000
Performance of very special visual tasks of extremely low contrast and small size	10000 - 20000

Figure 36: Indoors light reference levels (ToolBox, 2004)

B

THERMAL COMFORT TABLE

PMV	Thermal sensation	Stress sensation
< -3.5	Very cold	Extremely cold
$[-3.5, -2.5]$	Cold	Strong cold
$[-2.5, -1.5]$	Cool	Moderate cold
$[-1.5, -0.5]$	Slightly cool	Slight cool
$[-0.5, 0.5]$	Neutral	No thermal
$[0.5, 1.5]$	Slightly warm	Slight warm
$[1.5, 2.5]$	Warm	Moderate warm
$[2.5, 3.5]$	Very hot	Strong warm
> 3.5	Very hot	Extremely warm

Figure 37: Thermal Comfort Values (Silva and Analide, 2018)

C

APPLICATION INTERFACE

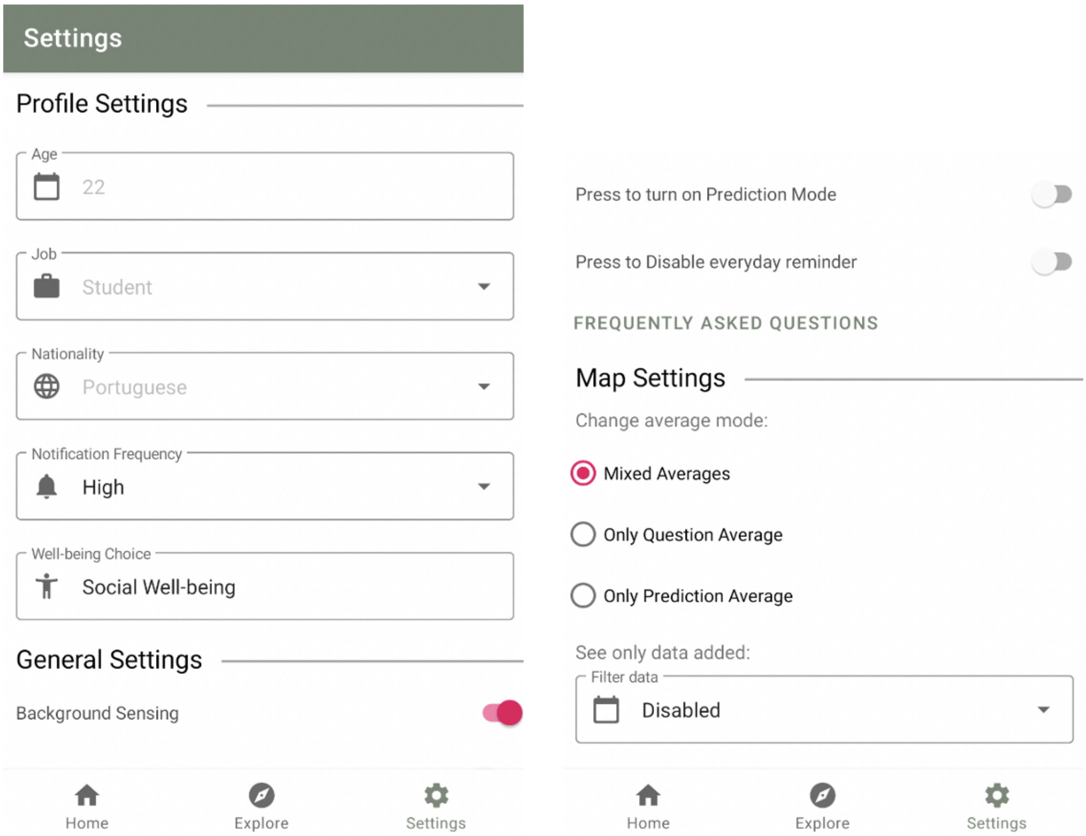


Figure 38: Settings Page

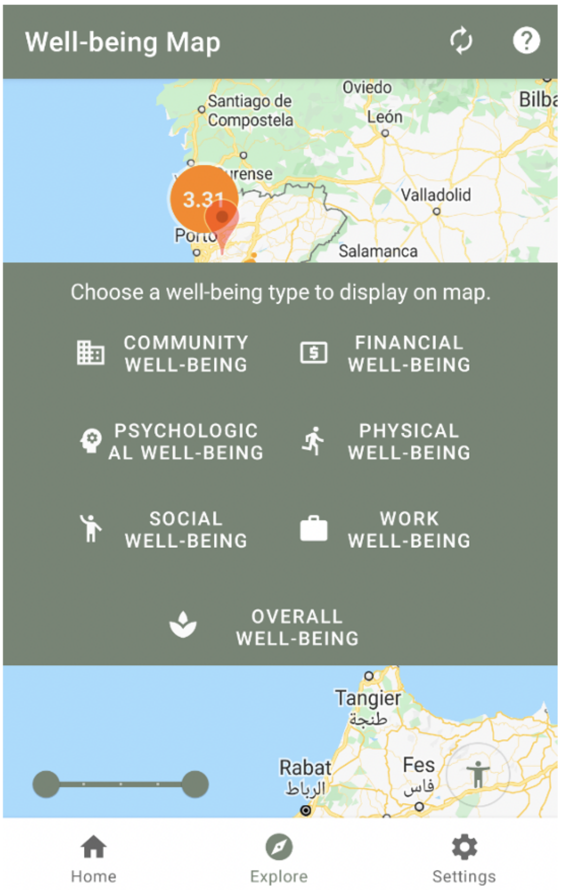


Figure 39: Well-being Map

D

DATA ANALYSIS

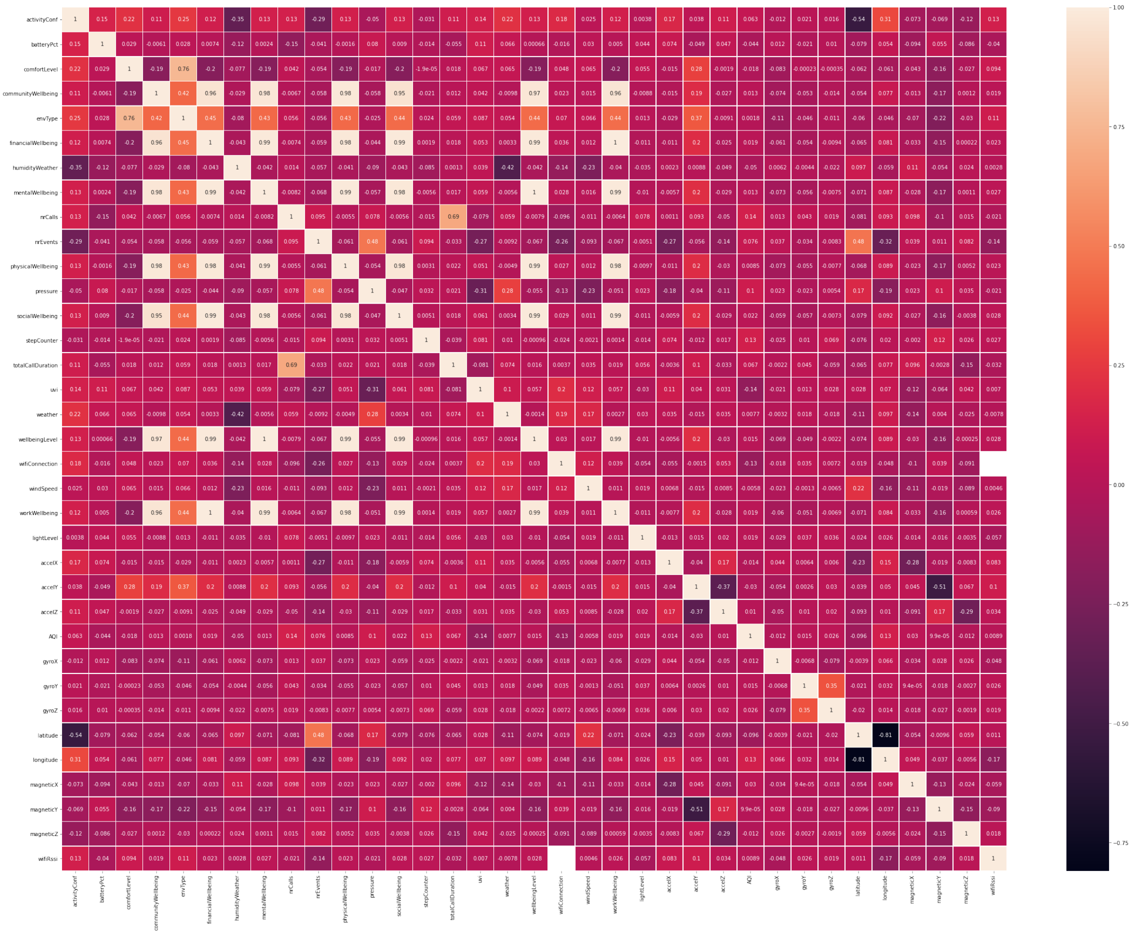


Figure 40: Data Captured During the Quarantine - heatmap

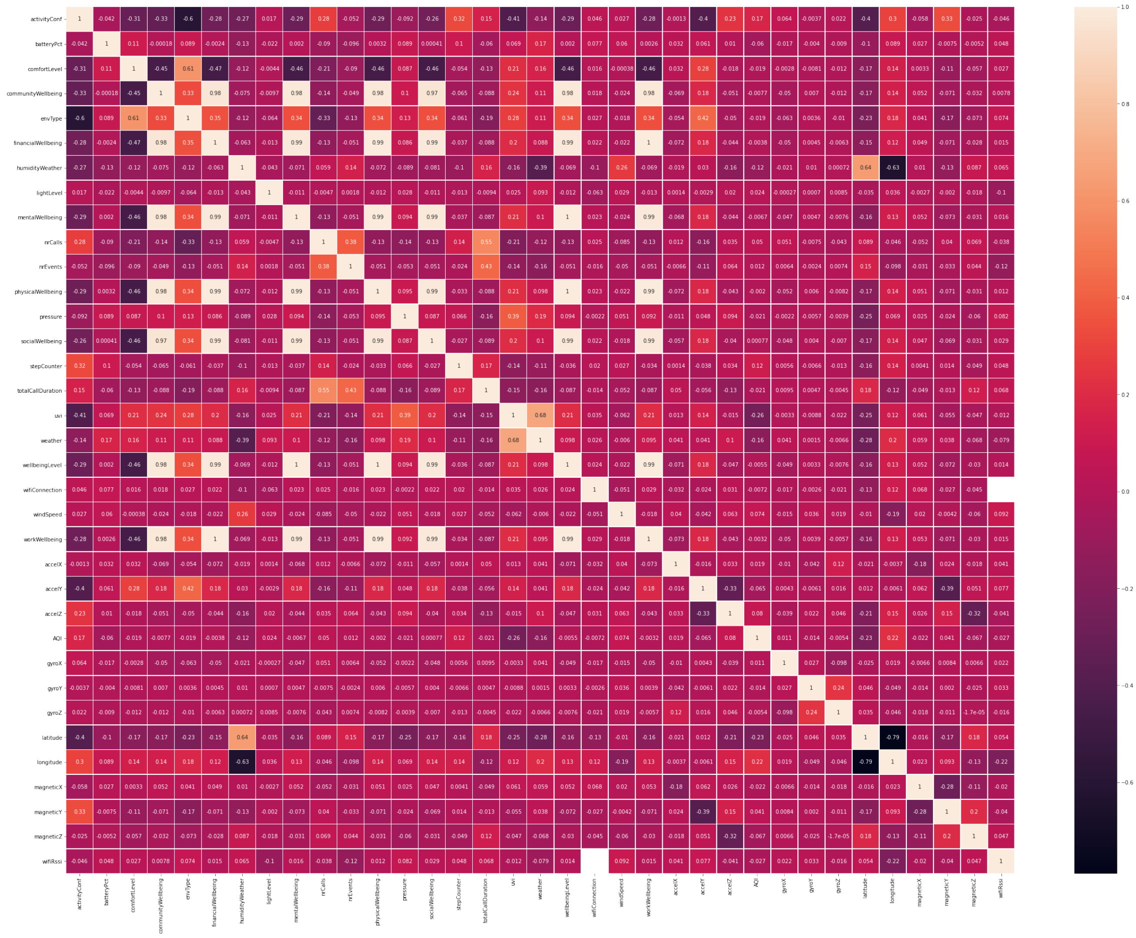


Figure 41: Data Captured After the Quarantine - heatmap

E

STUDIES RESULT

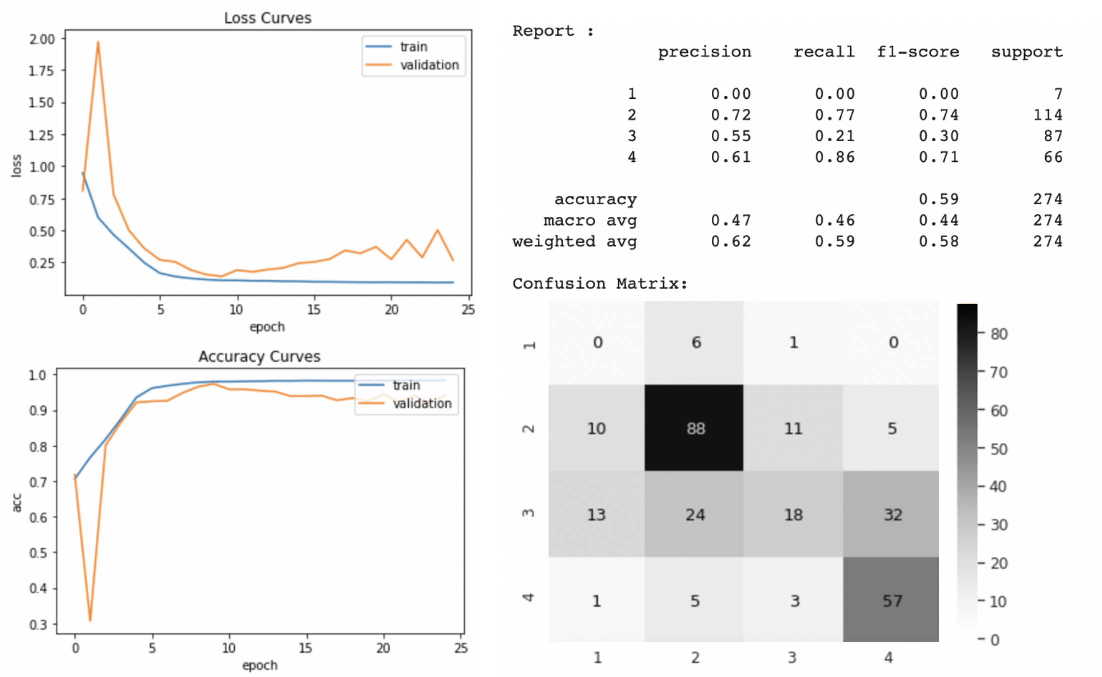


Figure 42: Comfort - Data Captured During the Quarantine - LSTM

	precision	recall	f1-score	support
0	0.74	0.93	0.82	120
1	0.92	0.88	0.90	196
2	0.84	0.73	0.78	99
3	0.97	0.93	0.95	230
accuracy			0.88	645
macro avg	0.87	0.87	0.86	645
weighted avg	0.89	0.88	0.88	645

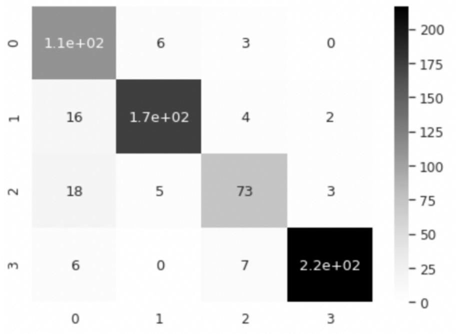
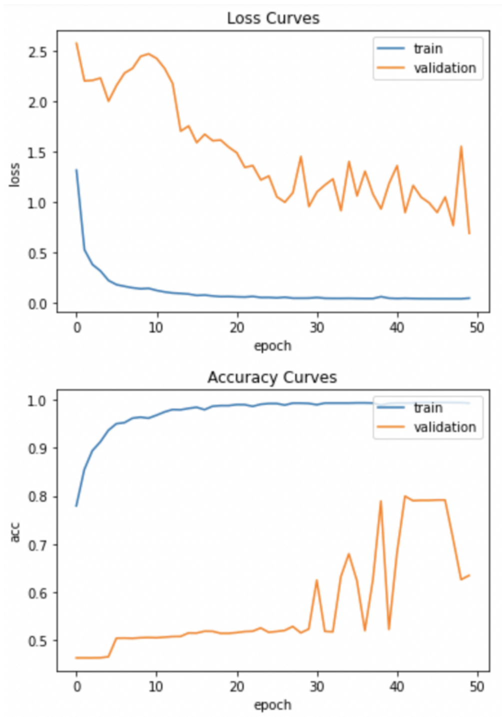


Figure 43: Physical Well-being - Data Captured During the Quarantine - RandomForest



Report :

	precision	recall	f1-score	support
1	0.17	0.75	0.27	4
2	0.29	0.14	0.19	63
3	0.83	0.63	0.72	390
4	0.40	0.86	0.54	98
accuracy			0.61	555
macro avg	0.42	0.59	0.43	555
weighted avg	0.69	0.61	0.62	555

Confusion Matrix:

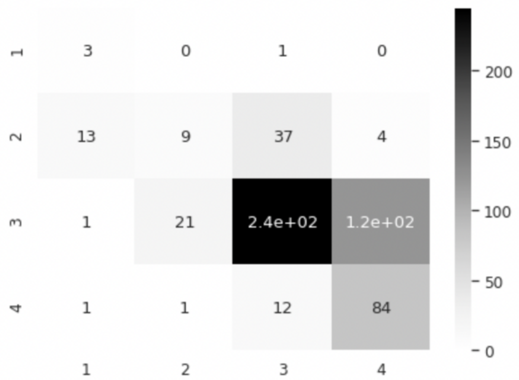


Figure 44: Comfort - Data Captured After the Quarantine - LSTM

	precision	recall	f1-score	support
0	0.54	0.91	0.68	56
1	0.69	0.51	0.58	211
2	0.60	0.66	0.63	195
3	0.95	0.94	0.95	350
accuracy			0.76	812
macro avg	0.69	0.75	0.71	812
weighted avg	0.77	0.76	0.76	812

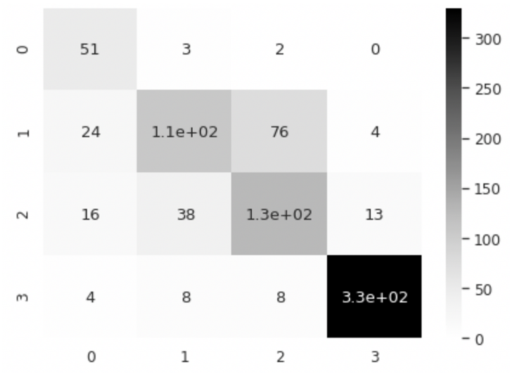


Figure 45: Physical Well-being - Data Captured After the Quarantine - RandomForest

NB: place here information about funding, FCT project, etc in which the work is framed. Leave empty otherwise.