FEDERAL UNIVERSITY OF PARAIBA
INFORMATICS CENTER
POSTGRADUATE PROGRAM IN INFORMATICS

A STUDY ON THE GENERATION OF EXPLANATIONS BASED ON ONTOLOGIES - A CASE STUDY IN mHEALTH

ISABELA NASCIMENTO CAVACO

JOÃO PESSOA - PB
July-2021

FEDERAL UNIVERSITY OF PARAIBA
INFORMATICS CENTER
POSTGRADUATE PROGRAM IN INFORMATICS

A STUDY ON THE GENERATION OF EXPLANATIONS BASED ON ONTOLOGIES - A CASE STUDY IN mHEALTH

ISABELA NASCIMENTO CAVACO

JOÃO PESSOA - PB
July-2021

Catalogação na publicação
Seção de Catalogação e Classificação

C376s Cavaco, Isabela Nascimento.
A study on the generation of explanations based on ontologies : a case study in mHealth / Isabela Nascimento Cavaco. - João Pessoa, 2021.
71 f. : il.

Orientação: Clauirton de Albuquerque Siebra.
Dissertação (Mestrado) - UFPB/CI.

1. Informática - Ontologia. 2. Saúde - Dados. 3. Aplicações mobile. I. Siebra, Clauirton de Albuquerque. II. Título.

UFPB/BC CDU 004:111(043)

Elaborado por WALQUELINE DA SILVA ARAUJO - CRB-15/514

ISABELA NASCIMENTO CAVACO

A STUDY ON THE GENERATION OF EXPLANATIONS BASED ON ONTOLOGIES - A CASE STUDY IN mHEALTH

DISSERTATION PROPOSAL SUBMITTED TO THE POSTGRADUATE PROGRAM IN INFORMATICS AT THE FEDERAL UNIVERSITY OF PARAIBA, AS PART OF THE REQUIREMENTS TO OBTAIN THE MASTER'S DEGREE IN INFORMATICS.

Supervisor: Ph.D. Prof. Clauirton de Albuquerque Siebra

JOÃO PESSOA - PB

July-2021

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Ata da Sessão Pública de Defesa de Dissertação de Mestrado de Isabela Nascimento Cavaco, candidata ao título de Mestre em Informática na Área de Sistemas de Computação, realizada em 30 de julho de 2021.

Aos trinta dias do mês de julho do ano de dois mil e vinte e um, às quatorze horas, por meio de videoconferência, reuniram-se os membros da Banca Examinadora constituída para julgar o trabalho da srª. Isabela Nascimento Cavaco, vinculada a esta Universidade sob a matrícula nº 20191001153, candidata ao grau de Mestre em Informática, na área de "Sistemas de Computação", na linha de pesquisa "Sinais, Sistemas Digitais e Gráficos", do Programa de Pós-Graduação em Informática, da Universidade Federal da Paraíba. A comissão examinadora foi composta pelos professores: Clauirton de Albuquerque Siebra (PPGI), Orientador e Presidente da banca; Danielle Rousy Dias Ricarte (UFPB), Examinadora Externa ao Programa; e Cecilia Neta Alves Pegado Gomes (FAMENE), Examinadora Externa à Instituição. Dando início aos trabalhos, o Presidente da Banca cumprimentou os presentes, comunicou a finalidade da reunião e passou a palavra a candidata para que ela fizesse a exposição oral do trabalho de dissertação intitulado: "A Study on the Generation of Explanations Based on Ontologies - A Case Study in mHealth". Concluída a exposição, a candidata foi arguida pela Banca Examinadora que emitiu o seguinte parecer: "***aprovado***". Do ocorrido, eu, Ruy Alberto Pisani Altafim, Coordenador do Programa de Pós-Graduação em Informática, lavrei a presente ata que vai assinada por mim e pelos membros da banca examinadora. João Pessoa, 30 de julho de 2021.

Prof. Dr. Ruy Alberto Pisani Altafim

Prof. Clauirton de Albuquerque Siebra
Orientador (PPGI-UFPB)

Prof. Danielle Rousy Dias Ricarte
Examinadora Interna (PPGI-UFPB)

Prof. Cecilia Neta Alves Pegado Gomes
Examinadora Externa à Instituição (FAMENE)

ABSTRACT

While mobile health (mHealth) applications provide a proper way to continuously assess data about the health conditions of their users, machine learning (ML) is the main technique used to process such data by means of inductive reasoning. However, ML algorithms do not usually give any explanation concerning the rationale of their produced outputs due to the black-box feature of such algorithms. This study analyzed 120 mHealth applications to create an integrated ontology that represents the health condition of mobile users and can be used as background knowledge to generate explanations for inductive reasoning. The integrated ontology involved several quality of life (QoL) dimensions (e.g., diet, physical activity, emotional, etc.), enabling the specification of a holistic process of reasoning that can improve the effectiveness of interventions. Therefore, the main contributions of this study are (1) the proposal of a strategy to create background knowledge for mHealth applications that support holistic reasoning and explanations regarding the results obtained by means of inductive reasoning, (2) evaluation of a description logics based approach to generate explanations using a simplified version of the ontology, and (3) discussions about important elements that can affect the readability and accuracy of explanations, such as the use of unnamed classes and configuration of the explanation algorithms.

Keywords: Health care, Knowledge modeling, Mobile Applications

ABSTRACT

Enquanto os aplicativos mobile de saúde (mHealth) fornecem uma maneira adequada de avaliar continuamente os dados sobre as condições de saúde de seus usuários, o aprendizado de máquina (ML) é a principal técnica usada para processar esses dados por meio do raciocínio indutivo. No entanto, os algoritmos de ML geralmente não fornecem nenhuma explicação sobre a lógica das saídas produzidas devido a tais algoritmos serem caixa preta. Este estudo analisou 120 aplicativos mHealth para criar uma ontologia integrada que representa a condição de saúde dos usuários destas aplicações e pode ser usada como conhecimento de fundo para gerar explicações para o raciocínio indutivo. A ontologia integrada envolveu várias dimensões da qualidade de vida (QoL) (por exemplo, dieta, atividade física, emocional, etc.), permitindo a especificação de um processo de raciocínio holístico que pode melhorar a eficácia das intervenções. Portanto, as principais contribuições deste estudo são (1) a proposta de uma estratégia para criar conhecimento para aplicativos mHealth que suportam raciocínio holístico e explicações sobre os resultados obtidos por meio de raciocínio indutivo, (2) avaliação de uma abordagem baseada em lógicas descritivas para gerar explicações usando uma versão simplificada da ontologia e (3) discussões sobre elementos importantes que podem afetar a legibilidade e precisão das explicações, como o uso de classes sem nome e configuração dos algoritmos de explicação.

Keywords: Saúde, Modelagem, Aplicações Mobile

FIGURE LIST

TABLE LIST

LIST OF ABBREVIATIONS AND ACRONYMS

AI	*Artificial Intelligence*
ARFF	*Attribute-Relation File Format*
ANN	*Artificial Neural Network*
CELOE	*Class Expression Learning for ontology Engineering*
DL	*Description Logics*
ELTL	*EL Tree Learner*
GPL	*General Public License*
HAR	*Human Activity Recognition*
IoT	*Internet of Things*
MHealth	*Mobile Health*
ML	*Machine Learning*
OWL	*Web Ontology Language*
QoL	*Quality of Life*
QS	*Quantified Self*
RBF	*Radial Basis Function*
SVM	*Support Vector Machine*
WHO	*World Health Organization*

SUMMARY

1 INTRODUCTION

The use of mobile health (mHealth) applications according to Sim (2019) has grown in the last decade and several studies are presenting proposals to use mobile devices to support real-time and continuous monitoring of their users. These proposals range, for example, from cardiac rehabilitation support programs as Rosario et al (2018) described, and some others have a real-time detection of cough events as demonstrated by Hoyos-Barceló and others (2018). Mobile technology is very useful for such a domain because it can generate proper just in time (right support at the right moment/context and in the right amount), and customized (based on an individual´s performance and goals) assessments and interventions.

According to Istepanian and Al-Anzi (2018), over this process, the task of health assessment usually generates a significant amount of data so the studies in the health area typically rely on machine learning (ML) techniques to obtain insights into the underlying problems. However, the main ML techniques, such as neural networks and support vector machines (SVM), execute as black boxes in the sense they do not give any direct indication for why an output was generated by their internal models. Neural networks, for example, rely on the distributed nature of the information encoded in the set of the network weighted connections. Thus, the rationale concerning the mapping from inputs to outputs is not “human-readable”, which is generally a compulsory requirement for safety-critical applications. In the medical domain, for example, it is important that the healthcare personnel comprehend and could explain how such algorithms arrived at their predictions.

Considering such a limitation, studies are investigating how interpretable explanations can be automatically generated for black-box ML methods, claimed by Melis and Jaakkola (2018) . The study of Fong and Vedaldi (2017), for example, proposed a general framework for learning different kinds of explanations for any black-box algorithm. Their main idea is to provide interpretable rules that describe the input-output relationship captured by the learning functions. Similarly, other studies conducted by Lehmann, Bader and Hitzler (2010) are also based on rule extraction, which employs propositional sentences as target formalism for creating the explanations. As this logic family is limited in terms of expressiveness, the outcomes also remain restricted in terms of explanations that can be generated.

A different approach to generate explanations is to use background knowledge, which describes concepts and their relations in the form of structured data. Ontologies are the main current approach to represent this knowledge. In order, they are commonly used in the health domain such as the studies shown by the authors Cornet and Keizer (2008) and Mastropietro and others (2018), due to their advantages as consistency, reuse, and easy extensibility. This dissertation considers this form of representation to create background knowledge based mainly on the information extracted from 120 Android and iPhone mHealth applications that are publicly available in app stores.

1.1 MOTIVATIONAL SCENARIO

Nowadays, mobile technologies are everywhere and, according to Pew Research Center (Pew Research Center, 2019), it is currently estimated that more than 5 billion people have mobile devices and over half of these connections are smartphones. Android and IOS, from Google and Apple respectively, are the main operating systems for smartphones in this market. Both platforms together provide a total of 4.9 million apps, being 2.2 million apps from iOS and 2.7 million from Android (Pew Research Center, 2019). With this enormous variety of mobile applications, users can choose from thousands of apps of diverse categories, which range from games and office tools to health and wellbeing apps that provide tips for exercises, information about heart rate, rhythm, pace, travelled distance, nutrition tips, and others.

In this scenario, the growth of mobile applications has led users to increasingly look for technological solutions to assist them to monitor and maintain their health conditions and, consequently, improve their Quality of Life (QoL). The World Health Organization (WHO)[1] defines this important concept (QoL) as an "*individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns*". This definition involves several dimensions, such as physical health, psychological state, personal beliefs, social relationships, and relationship to salient features of the environment. As a strategy to assess the QoL of individuals, WHO has developed and evaluated a scale (WHOQOL) that presents a more granular set of QoL subdimensions (e.g., activities of daily living, sleep and rest, etc.) that can be measured. Therefore, the goal of the majority of the mHealth applications is exactly to assess data that is associated with these QoL dimensions.

As smartphones are close to their users at least half of the time, they are likely the best tools to assist health improvement and QoL over everyday activities, as affirmed by Wac and others (2015). According to Wac (2018), it was estimated that 60% of the US population in 2013 tracked some aspect of their life (e.g., weight, exercise, mood), 33% of adults tracked health indicators or symptoms (e.g., blood pressure, blood sugar, headaches, or sleep patterns), and 12% tracked a health indicator on behalf of someone they care for. It was also evidenced by Chiarini and others (2013) that mobile technology improved healthcare for the elderly and patients with chronic conditions by providing solutions in terms of self-healthcare, assisted healthcare, supervised healthcare, and continuous monitoring.

Considering the emotional QoL dimension as an example, as stated by Ciman, Matteo, and Wac (2016), people are always in stressful situations, especially in the past few decades, and this situation negatively influences their lives in the long term. According to the author, research has shown how stress relates to several other people's health parameters. For example, stress is correlated with heart, respiration rates and skin response values. Meanwhile, we're living in a time where it is possible to easily check our health with the use of wearable devices. By exploiting sensors in wearables and smartphones, apps are giving users new powerful mobile experiences that have the potential to change the way users live and interact with each other. Therefore, as the authors Lane and Georgiev (2015) evidenced, users can quantify their sleep and exercise patterns, monitor personal commute behaviors, and track their emotional state. According to Jiang, Wenchao and Zhaozheng (2015), some examples of wearable sensors are smartwatches and sport bracelets, which use embedded accelerometers, gyroscopes, microphones and other types of sensors. "*The use of sensors in wearables and smartphones, apps are giving users experiences that have the potential to change the way users live and interact with each othe*r, *offering interventions and plans of action to improve QoL*" (Lane et. al, 2015).

[1] https://www.who.int/healthinfo/survey/whoqol-qualityoflife/en

This mHealth scenario is an interesting case study for our dissertation since it involves the capture of several pieces of information, which can be used as input for inductive systems. These systems are powerful resources to generate conclusions and support the definition of health interventions. However, they should also generate explanations regarding their conclusions to better guide the definition of such interventions.

1.2 PROBLEM

There are many gadgets and apps focused on monitoring QoL dimensions, such as nutrition, physical activity, stress, and so on, but they're not linked or integrated. Consider, for example, an app for physical activity intervention that can indicate a QoL plan for losing weight. Studies [9,10] demonstrated that even when exercise energy expenditure is high, a healthy diet is still required for weight loss to be successful. Thus, the assessment of the physical activity and nutritional dimensions should be conducted together. Unfortunately, this may not still be enough because, even if individuals have good behavior in both dimensions, a stressful life associated with psychological conditions liberates an extra amount of cortisol that impacts the body's metabolism to store fat.

As these applications are not integrated, then we also do not have a proper holistic representation of data that could be used as background knowledge for inductive systems. Then, our first problem is how to create this knowledge and, after that, evaluate the use of this knowledge so we can identify aspects that could improve the quality of the explanations. This contribution extends the state of the art in explanations because the literature does not present guidelines on how to create ontologies to better support the generation of explanations.

1.3 OBJECTIVE

The main objective of this study is to create a knowledge representation, in the form of an ontology, which integrates different QoL dimensions that can support the generation of explanations for inductive reasoning processes. So, the main contributions of this study are (1) the proposal of a strategy to create background knowledge for *mHealth* applications that support holistic reasoning and explanations regarding the results obtained by means of inductive reasoning, (2) evaluation of a description logics based approach to generate explanations using a simplified version of the ontology, and (3) discussions about important elements that can affect the readability and accuracy of explanations, such as the use of unnamed classes and configuration of the explanation algorithms.

1.4 DISSERTATION STRUCTURE

The remainder of this document is organized as follows: Chapter 2 discusses the theoretical fundamentals of QoL, mobile health, wearables, ontology and inductive learning. Chapter 3 presents the research method, detailing how the apps review related to the QoL dimensions was conducted, the strategy to compose the holistic ontology, the configuration of a neural network to act as an accuracy baseline, and the evaluation process used to compare ontological scenarios. Chapter 4 presents the results which are: apps review, the final version of the ontology, the case study involving the neural network, and the accuracy comparison between neural network and DL-Learner experiments. Finally, Chapter 5 concludes this work with the main contributions, limitations and research directions.

2 THEORETICAL FOUNDATIONS

This chapter presents the theoretical basis used to develop the research, which involve the concepts of quantified self, quality of life, mobile health, ontology and inductive methods (Supervised Learning and Neural Networks).

2.1 QUANTIFIED SELF

Mobile technologies are able to provide solutions to ultimately improve QoL dimensions, which are evaluated in terms of *Quantified Self* (QS). According to Wac (2018), this term is defined as a trend, where individuals focus on tracking their own state and behavioral patterns with the help of personalized devices (wearables and smartphones) for continuous, ideally unobtrusive tracking. Still according to Wac (2018), users expect that digital data from QS technologies, which embrace wearables, applications, and self-reports, enable them to track different aspects of their physical and psychological health, social interactions, and environmental conditions. These four aspects constitute the main domains of the *Quality of Life* (QoL) literature.

2.2 QUALITY OF LIFE

According to Wac (2018), *Quality of Life* (QoL) is defined by the World Health Organization (WHO) as the "individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns". The four QoL domains are illustrated in Figure 1.

Figure 1: QoL domains.

QoL Domain	Facets incorporated within QoL domains	CODE
Physical Health	Activities of daily living	phy-adl
	Dependence on medicinal substances, medical aids	phy-meds
	Energy and fatigue	phy-energy
	Mobility	phy-mobility
	Pain and discomfort	phy-pain
	Sleep and rest	phy-sleep
	Work capacity	phy-work
Psychological Health	Bodily image and appearance	psy-bodyimage
	Negative feelings	psy-negativefeel
	Positive feelings	psy-positivefeel
	Self-esteem	psy-selfesteem
	Spirituality/religion/personal beliefs	psy-beliefs
	Thinking, learning, memory and concentration	psy-thinking
Social relation.	Personal relationships	soc-relationships
	Social support	soc-support
	Sexual activity	soc-sex
Environment	Financial resources	env-finances
	Freedom, physical safety and security	env-freedom
	Health and social care: accessibility and quality	env-healthcare
	Home environment	env-home
	Opportunities for acquiring new information and skills	env-info
	Participation in/opportunities for recreation/leisure	env-leisure
	Physical environ. (pollution / noise / traffic / climate)	env-environ
	Transport	env-transport

Source: (Wac, 2018).

A new perspective on diseases can be obtained by focusing on the individuals' own views of their wellbeing. For example, it’s known that diabetes involves poor body regulation of blood glucose, but the effect of the illness on the perception that individuals have of their social relationships, working capacity, and financial status has received little systematic attention. QoL instruments developed by WHO[2] (*WHOQOL-100* and *WHOQOL-BREF*) are tools that enable this type of research to be carried out. These tools also inquire how satisfied the patients are with their functioning and with the effects of treatment (WHOQOL: Measuring Quality of Life, 2020). As stated by Sungmee and Jayaraman (2003), technology is the key to enhancing the quality of life for everyone, from newborns to senior citizens. Moreover, they affirm that technology is indeed the catalyst that can rapidly transform healthcare, and by doing so, minimize the loss of human life and enhance the QoL.

2.3 MOBILE HEALTH

“Mobile health technologies (*mHealth*) are playing an instrumental role in serving patients by making healthcare more affordable, accessible and available” (Chiarini et. al, 2013). The study from Peart and others (2017) reviews several *mHealth* apps that collect physiological and anatomical measurements such as heart rate, range of motion and physical performance measurements such as vertical jump height, barbell velocity and contact times. Examples of such apps are: *RRate, HRV4Training, Instant Heart Rate, PostureScreen, PowerLift, My Jump* and others.

Mobile apps also enable higher self-monitoring for individuals with chronic diseases such as obesity, cardiovascular disease, and type-2 diabetes. These apps contribute, for example, to nutrition care, such as the study from Fallaize and others (2019) that evaluates the extent to which popular nutrition-related apps measure the energy, macronutrient, and micronutrient consumed. Examples of apps assessed by them are *MyFitnessPal*, *Samsung Health, FatSecret* and others.

Some activities, however, are not fully measured only with the use of mobile apps and smartphone embedded sensors. According to Fallaize and others (2019), even though smartphones include accelerometers and it’s possible to count the steps without the use of external devices, a study has shown that these apps lack accuracy in comparison with a professional pedometer, probably due to the low quality of the accelerometers included in smartphones. The usability of wearable sensors is intended, for example, for human activity recognition (HAR), which according to Jiang, Wenchao and Zhaozheng (2015) is the recognition of body states such as standing, walking, running and falling. Such recognition can be applied to many application fields just as human-computer interaction and surveillance. The application of wearable sensors is broad, but their traditional use is mostly focused on healthcare, personal training, military, diet management, and security.

In order, as stated by Chen, Yuqing and Yang (2015), due to significant improvements made in sensor and processor technologies over the past decades, it is possible to achieve more accurate sensors in a smaller size and faster processors with lower power consumption. According to Lane and others (2015), this is essential to the development of resources that can use sensors to assess and process data associated with the user behavior and ambient context .

[2] https://www.who.int/tools/whoqol

Moreover, still according to Lane and others (2015), there is currently an increasing market for applying *mHealth* methods in our daily lives than years ago. Considering the growth of mobile technologies, especially smartphones that come with several built-in types of sensors, it is possible to supply this market with applications that do not require extra hardware apart from the one already embedded in the smartphones.

2.4 ONTOLOGY

As mentioned before, the QoL research often involves knowledge representation in the form of ontology. As described by Snae, Chakkrit, and Brüeckner (2007), the term ontology has been widely used in the field of Artificial Intelligence, computer and information science, especially in domains such as cooperative information systems, intelligent information integration, information retrieval and extraction, knowledge representation, and database management systems. According to Gruber (1993), ontology is an explicit formal specification of the terms in a domain and relationships among them. The authors Noy and McGuinness (2001) of a guide to create ontology models, using Protegé-2000[3] editing environment, that is a feature rich ontology editing with an user interface, with full support for the OWL Web Ontology Language, and direct in memory connections to description logic reasoners and tracking down inconsistencies. On the the authors provided the following reasons to develop ontologies:

- To share common understanding of the information structure among people or software agents;
- To enable the reuse of domain knowledge;
- To make domain assumptions explicit;
- To separate domain knowledge from the operational knowledge;
- To analyze domain knowledge.

Moreover, the guide defines ontology as a formal explicit description of concepts in a domain (classes, sometimes called concepts), properties of each concept describing various features and attributes of the concept (called roles or properties), and restrictions on slots (facets, also called role restrictions).

An ontology together with a set of individual instances of classes constitutes a knowledge base. Still according to the guide authors Noy and McGuinness (2001), the development of an ontology involves:

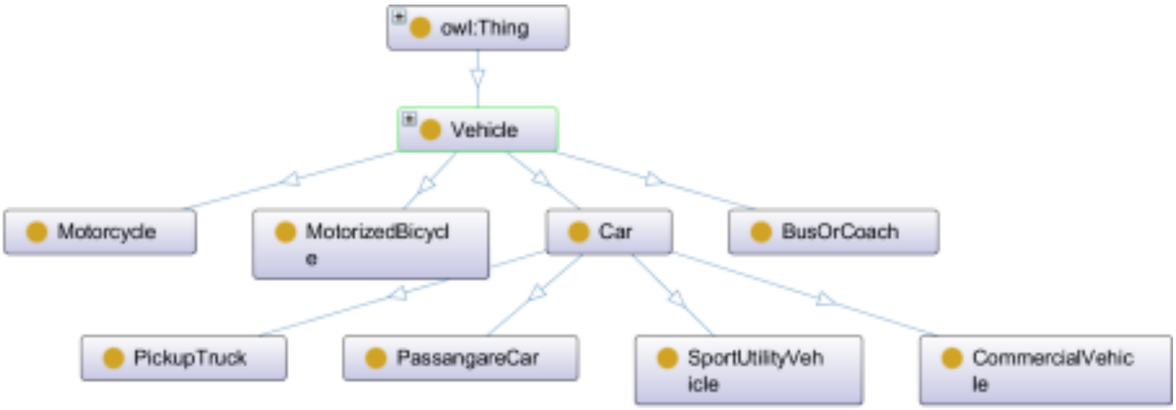
- Defining classes in the ontology;
- Arranging the classes in a taxonomic (subclass–superclass) hierarchy;
- Defining properties and describing allowed values for these properties;
- Filling in the values for properties with instances.

Figure 2 illustrates a very simple example of ontology, designed for *Car* domains. Apart from the class *Thing*, which is the most general concept of an ontology, the example describes

[3] https://protege.stanford.edu

the class *Vehicle* as the top-level concept, while its subclasses *MotorCycle*, *MotorizedBicycle*, *Car* and *BusOrCoach* are general middle level concepts. The other subclasses of *Car* - *PickupTruck, PassangareCar, SportUtilityVehicle* and *CommercialVehicle* are the bottom level concepts.

Figure 2: Representation of a Car ontology, with classes and subclasses.



Source: (Automotive Ontology Community Group Wiki, 2016).

In this example, the arrows represent the hierarchy between classes and subclasses (from top-down), which are also called *concepts*. These concepts can have slots (data properties) as presented in Automotive Ontology Community Group Wiki (2016). For example, the concept *Vehicle* could have the slots *speed, torque, weightTotal* and *accelerationTime*. “A slot should be attached at the most general class that can have that property” (Noy and McGuinness, 2001). Thus, the choice to add these properties means that they’re common to all instances of *Vehicle*. The ontology above shows an “*is a*” relationship between concepts, from subclasses to classes, such as: *PickupTruck is a Car*, which *is a Vehicle*, and so on.

About knowledge-engineering methodology, the guide quotes that “there is no one correct way or methodology for developing ontologies” (Noy and McGuinness, 2001). The modeling decisions that describe an iterative approach to ontology development can follow a few steps below (that revise and refine the evolving ontology and fill in the details). Therefore, deciding for what purpose the ontology will be used, how detailed or general the ontology is going to be, and many other questions, can follow some guide, helping matching decisions among several viable alternatives, as described next:

Step 1. Determine the domain and scope of the ontology;

Step 2. Consider reusing existing ontologies;

Step 3. Enumerate important terms in the ontology;

Step 4. Define the classes and the class hierarchy;

Step 5. Define the properties of classes-slots;

Step 6. Define the facets of the slots;

Step 7. Create instances.

Moreover, as stated by Maxat et al (2020)., ontologies have long been employed in the life sciences to formally represent and reason over their domain knowledge. According to such a study, ontologies have also the potential to provide constraints that improve machine learning models and, thus, be an important symbolic resource to generate explanations for inductive methods, as detailed in the following section.

2.5 INDUCTIVE METHODS

According to Michalski (1983), the ability that people have to make accurate generalizations from a few scattered facts or to discover patterns in seemingly chaotic collections of observations holds the key to improving methods used by computers to obtain knowledge. This ability is achieved by a process called inductive learning, which is an inductive inference from facts provided by "instructors" or environments. Still according to Michalski (1983), this form of learning is one of the central topics of machine learning. The author Nicoletti (1994) describes inductive learning as the acquisition of concepts through a set of examples (training set) and the concept of induction corresponds to a search in a space of hypotheses in order to find those that best classify the examples, in terms of precision and understandability.

To represent the training set, inductive learning systems apply an attribute-based language. An attribute is a possible feature relevant to the concept being learned. Training examples are described as the attribute-value of a class, and the attribute set to describe the examples is fixed, with each example belonging to a single class. As mentioned by Nicoletti (1994), the purpose of such inductive systems is to find the rule that predicts the class based on attributes and values, in relation to what the system has learned with the training sets.

The author Montavon (2017) states that, although highly successful in terms of performance, inductive learning systems have a drawback of acting like a black box in the sense that it is not clear how and why they arrive at a particular decision.

As stated by Mantaras and Armengol (1998), inductive learning methods can be classified according to the following perspectives: Supervised/Unsupervised learning, Single/Multiple concept learning, and Propositional/ Relational learners. From these methods, our approach focuses on supervised learning for classification, which is explored in the next subtopic.

2.5.1 SUPERVISED LEARNING - CLASSIFICATION

As stated by Sathya and Abraham (2013) classification is one of the most frequently encountered decision-making tasks of human activity and problems occur when an object needs to be assigned into a predefined group or class based on a number of observed attributes related to that object. Still according to them, there are many industrial problems identified as classification problems, such as: Stock market prediction, Weather forecasting, Bankruptcy prediction, medical diagnosis, Speech recognition, Character recognition and much more.

According to Bhavsar and Ganatra (2012), classification is a supervised learning method, since the instances are given with known labels, in contrast to unsupervised learning in which labels are not known. In other words, supervised learning is based on training a data sample from a data source with correct classification already assigned, Sathya and Abraham (2013). Some examples of machine learning classification algorithms were described by Mohammad and others (2006) such as *Support Vector Machine (SVM), Decision Tree*, *Multilayer Perceptron (Neural Networks)*, *K-Nearest Neighbors, Naive Bayes and Radial Basis*

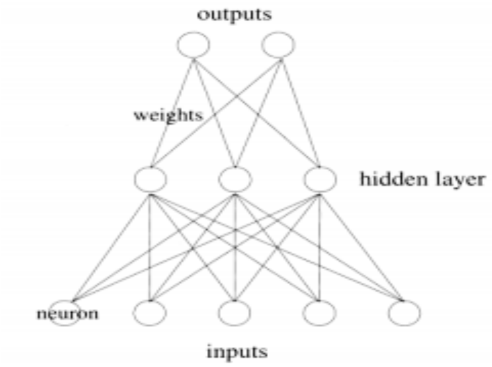
Function (RBF) Network. As previously mentioned, we use a neural network to generate an accuracy baseline for comparisons with the accuracy of explanation algorithms.

2.5.2 NEURAL NETWORKS

Also according to Bhavsar and Ganatra (2012), an *Artificial Neural Network (ANN)* is a set of connected input/output units in which each connection has a weight associated with it, during the learning phase, to predict the class label of the input sample, the network learns by adjusting its weights. The supervised learning paradigm of an ANN following Kulkarni and Joshi (2015) is efficient and finds solutions to several linear and non-linear problems such as classification, plant control, forecasting, prediction, robotics and others.

Based on the human central nervous system, Neural Networks are networks with points that connect, as a small abstraction of how the brain works. According to [35], ANN consists of an input layer of neurons (or nodes, units), one or more hidden layers of neurons, and a final layer of output neurons. Figure 3 shows a typical ANN architecture, with lines representing the connection between neurons.

Figure 3: Architecture of a Neural Network.



Source: [37].

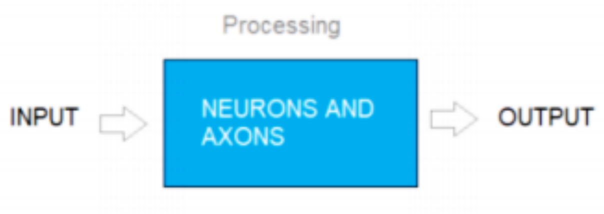
Each connection is associated with a weight. The output h_i of a neuron i in the hidden layer is calculated with the formula:

$$h_i = \sigma\Big(\sum_{j=1}^{N} V_{ij}x_j + T_i^{hid}\Big),$$

where there is an activation (or transfer) function σ, being N the number of input neurons, V the weights, x inputs from the input neurons, and T the threshold terms of the hidden neurons [35].

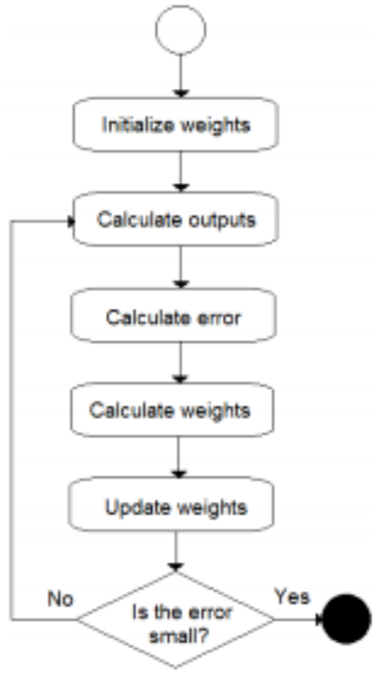
Based on the biological fundamentals of how the brain works in information processing, the neuron fires if the input is greater than a defined number (then the neuron can be activated or not). Then, in brief, an input value is provided, the network processes, and returns a response based on the data that was inputted before. This process is summarized in Figure 4:

Figure 4: Information processing of a Neural Network.



For a Neural Network to find the best weights for the attributes, its error margin must be small. To achieve that, the weights are updated until the error margin decreases, by recalculating all nodes with their input values at every new weight update, as it's illustrated in Figure 5.

Figure 5: Neural Network Algorithm.



The simplest algorithm to calculate the error is described below:

$$Error = Correct\ response - Calculated\ response$$

Moreover, to calculate the adjusting weight to decrease the error rate, the following calculation is used:

$$Weight_{(n+1)} = Weight_{(n)} + (Learning\ rate * Input\ value * Calculated\ error\ value)$$

The calculations presented above can also be used for multilayer networks. A summary of what happens in the algorithm used in Neural Networks is:

- Initialize the weights with random values;
- Based on the data, perform the calculations with the weights and calculate the error;
- Calculates changes in weights and updates them (following the backpropagation methodology);
- The algorithm ends when the margin of error is inside of an acceptable margin, otherwise the calculations are remade using new weights.

This is a simplification of the ANN processing and several other possibilities could be used to calculate the intermediate parameter values over the training stage. The discussion of this process is important to stress that this and other inductive methods execute as black boxes, in the sense that they do not give any direct indication for why an output was generated by their internal models. Neural networks, in particular, rely on the distributed nature of the information encoded in the set of the network weighted connections. Thus, the rationale concerning the mapping from inputs to outputs is not “human-readable", which is generally a compulsory requirement for safety-critical applications. In the medical domain, for example, it is important that the healthcare personnel comprehend and could explain how such algorithms arrived at their predictions.

3 RESEARCH METHOD

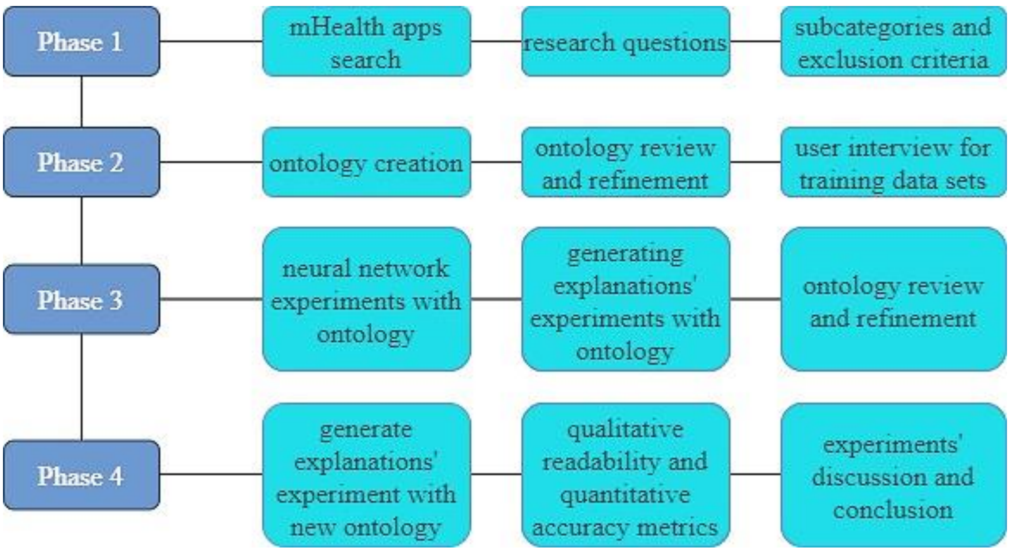
The research method of this study was divided into 4 phases (Figure 6). First, we conducted a review and categorization of 120 applications found in the app stores of mobile operating systems (iOS and Android), with further refinement and analysis on the input and output data required by these applications. After this categorization, we adopted an exclusion criterion to define which scope (categories) should be adopted to create the ontology.

Phase 2 accounted for the ontology design using Protegé-2000, a tool with a graphic user interface used to model ontologies, following the steps presented in Section 2.4. This design was important to better understand the connections between the areas and their mutual influences. Thus, we could also adjust the ontology using other resources of the academic literature [20,50]. Then, we conducted an interview with 20 apps users to create a dataset for tests, considering the concepts and properties of a subpart of the ontology. This dataset was further used to validate our proposals.

Phase 3 involves the execution of two experiments. The first one using Neural Networks through Weka[4], (tool with graphic user interface, data analysis and predictive modeling) and the second one using DL-Learner[5] (a tool for learning concepts in Description Logics (DLs) from user-provided examples), refining the ontology further.

Finally, in Phase 4, we conducted experiments again using DL-Learner with the new refined ontology. In that phase, the purpose was the generation of explanations, represented as a logical formula close to natural language. All experiments were evaluated in terms of qualitative readability of explanations and quantitative accuracy metrics.

Figure 6: Research methodology in phases.



We discuss each phase in detail in the following subsections: 3.1 mHealth Apps Review, 3.2 Research Questions, 3.3 ontology Design, 3.4 User Interview for Training Data

[4] https://www.cs.waikato.ac.nz/ml/weka

[5] https://dl-learner.org

Sets, 3.5 Neural Networks experiment with ontology Data, 3.6 Generating explanations with ontology Data and 3.7 Evaluation Process.

3.1 mHEALTH APPS REVIEW

We conducted a review on current mobile applications that could improve the quality of life of their users, considering the properties and definitions of QoL and its state of the art. To this initial stage, 5 research questions were raised about the applications, bringing relevant aspects such as advantages and strengths of using this type of application, disadvantages and difficulties, among other information that may be relevant to identify the real impact and effect of using mobile applications for this purpose.

The first part of this review was focused on the better understanding of the QoL concepts and its properties. Therefore, the research involved studies retrieved from popular digital libraries (e.g., Google Scholar[6] and ACM Digital Library[7]) with the generic search string "Quality of Life or QoL". With the main QoL concepts in mind, we selected 120 applications for Android and iOS platform in mid-December of 2019. They were divided into 12 theme groups, such as Physical Health, Health care, Emotional and Social, among others, as described in Table 1. The selection of applications by group was made according to the main function of the application. For example, if the application was more focused on diet, then it was included into the group related to diet; if it was aimed at relieving the emotional impacts of stress and anxiety, it was added to the emotional group. For each group, 10 applications were selected (5 for each operating system - Android and iOS). All classified groups are detailed in Appendix I.

Table 1: Application groups (groups not considered).*

Diet	Educational	Spiritual
Physical Health	Emotional	Recreation*
Health Care	Business*	Social
Entertainment	Productivity*	Tools*

Initially, a general overview of groups and types of applications was conducted. Then, according to the research progress, it was necessary to refine and decide which groups of applications were essential to remain. Some groups were removed because they only have indirect connections with the main focus of this study (quality of life). For example, the Productivity group that contains applications like Trello (shares and monitors the execution of activities). Other groups removed were Business, Tools and Recreation. Business brought applications directly linked to the business world, such as the Whatsapp Business and Google My Business. The Tools group brought applications to support people's daily lives like Color Note for notes or Dropbox for file storage. Finally, Recreation reflects applications that could assist people's leisure like Airbnb and Booking. These groups were removed over the refinement, but they are still present in the Appendix I for possible future references.

[6] https://scholar.google.com

[7] https://dl.acm.org

Tables 2, 3 and 4 present examples of dimension groups and their corresponding apps. The apps are followed by the star rating at the time of this survey.

Table 2: Physical Health application group.

IOS	Android
Nike Training Club - 5 starts	Stronglifts 5x5 - Weight Lifting & Gym Workout Log - 4.9 stars
8fit Workouts & Meal Planner - 4.7 stars	Step Counter - 4.7 stars
Seven, 7 Minute Workout - 4.8 stars	Legs & Butt Workout - 4.6 stars
Freeletics - Workout & Fitness - 4.6 stars	30 Day Fitness Challenge - 4.8 stars
Sworkit Fitness & Workout - 4.7 stars	Daily Workouts - 4.7 stars

Table 3: Health care application group.

IOS	Android
Sleepzy - 4.5 stars	Blood Pressure - 4.2 stars
Kardia - 4.9 stars	PsicoTests - 4.6 stars
Dental Drugs - 4.8 stars	Headache Log - 4.6 stars
Pregnancy + - 4.8 stars	Period Tracker - 4.9 stars
BabySparks - 4.7 stars	My Pregnancy - 4.4 stars

Table 4: Spiritual application group.

IOS	Android
Zen - 4.9 stars	Headspace: Meditation & Sleep - 4.3 stars
Calm - 4.7 stars	Law of Attraction Space - 4.7 stars
Colorfy - 4.6 stars	5' Minutes, I meditate - 4.4 stars
Daily Spiritual Quotes - 5 stars	Spiritual Me: Masters Edition - 4.6 stars
Headspace: Meditation & Sleep - 4.9 stars	Spiritual Transformation Daily - 4.5 stars

Some criteria were considered when selecting apps. First, they should have between 4.0 and 5.0 classification stars on the App Store[8] (iOS platform) and Google Play[9] (Android platform). At first, it was defined that only applications with rates higher than 4.5 stars would be selected but, for some groups, finding such top-rated applications proved to be difficult. Thus, the rating expectation was reduced to 4.0 stars. Open-source applications would also have priority, but since they represented small quantities, at the end this requirement was not considered as a priority. Other criteria adopted was that 5 IOS platform apps were ranked first (for all groups), and then the other 5 applications from Android platform. This strategy avoids choosing repeated applications for Android and IOS groups. There was also no ranking and priority about the language. Due to the difficulty to analyze subscription-based features, only free apps were selected, which can limit the view about the type of information that can be obtained by means of this technology.

After the selection of these apps, they have been analyzed in terms of data input required from users, such as age, current weight, frequency of exercises, and others. This analysis assisted in determining the relationship between domain areas, which is similar to the information requested as data entry in the selected mobile applications.

3.2 RESEARCH QUESTIONS

The next research questions were elaborated as a strategy to organize the information about the *mHealth* applications and stress the assessments that are in fact important to their users. The main objective of doing this review was to identify the classes and relationships that could be part of the ontology's creation, and to highlight what kind of information collaborates with what is requested by applications as inputs. These questions are:

A. What are the advantages and disadvantages of applications that aid a user's life, according to user ratings and related studies?

B. Is there any application that connects more than one QoL dimension?

C. What are the attributes most used by applications that link more than one dimension? Which inputs were most important to (example: heartbeat)?

D. Is there any real impact reported by the use of these applications in daily life?

These questions were also important to highlight that these applications lack in providing more information to their users regarding the rationale of the interventions generated. This means that they do not justify the reasons for users to follow their directives (e.g., why are applications telling their users to take certain action or deliver such information?). The results and discussions for this stage of the research are explored in Section 4.

3.3 ONTOLOGY DESIGN

The background knowledge was developed as an ontology that integrates the information about diverse health dimensions and enables the specification of a holistic process

[8] https://www.apple.com/ios/app-store

[9] https://play.google.com/store?hl=en

of reasoning that can improve the accuracy and reliability of health interventions. One of the limitations of the current approaches, regarding their knowledge representations and reasoning strategies, is the lack of integration among the different health dimensions as Sivilai and others (2012) explained. The design of this study was constructed based on the guide from Noy and McGuinness (2001) and the tool that helped us to build the ontology was the Protegé-2000 editing environment.

To design the ontology model, we analyzed the user input requested by the 40 applications, 10 per area (*Diet*, *Physical Health*, *Health care* and *Emotional*). The class and subclass names of the modeled domain areas are similar to the information requested as data input from the selected mobile applications. For example, some of the classes and subclasses of the ontology are (class - subclass, respectively): *Body* - *InitialStateOfPhysicalConditioning* and *PhysicalExercise* - *BasicExerciseDefinition*.

The names followed this nomenclature to remain similar to how the apps inquire this information. Regarding the data related to user profile, the following variables were defined for the *Person* class: age, gender, weight and height. It’s important to observe that when the ontology model was designed, the most popular names for variables and classes with the semantics were considered. Thus, some of the names presented in the figures described in this section may not be in the ontology explicitly. For example, Figure 7 presents 2 screenshots from the “Freeletics” and “30 Day Fitness Challenge” apps, which indicate the same body parts, but using different terms to work on (e.g., full body and all body).

Figure 7: A screenshot from the app “Freeletics (left), and a screenshot for the app “30 Day Fitness Challenge” (right).

Source: On the left, “Freeletics”[10] adapted to English. On the right, “30 Day Fitness Challenge”[11].

---

[10] https://play.google.com/store/apps/details?id=com.freeletics.nutrition&hl=en

[11] https://play.google.com/store/apps/details?id=com.popularapp.thirtydayfitnesschallenge&hl=en

As the focus of this present study is on *mHealth* applications, the strategy proposed was analyzing 120 health-related apps that are available in the app's stores. These applications could be divided into 12 dimensions, and we initially decided to embrace 3 of these dimensions (as explained in Section 3.1). Then, for each of these groups, we selected 10 apps with the best evaluations so all the data assessed with these apps could be represented as elements of the ontology. As an example of the knowledge specification, consider the ontology fragment related to a person (*Person* class), where apps requested personal information related to their *DietObjective*, *Body* and *ExerciseFrequency*.

Figure 8 illustrates an overview of the general model, with classes and subclasses that were created based on the app analysis. Academic studies were also used to adjust some concepts and properties of the ontology, such as the NESTORE models used by Mastropietro and others (2018) and the e-NUTRI project used by Fallaize and others (2019). Another important point of this work is how the dimensions are linked. The complete ontology model will be introduced and reviewed in Section 4.2.

Figure 8: Ontology model overview.

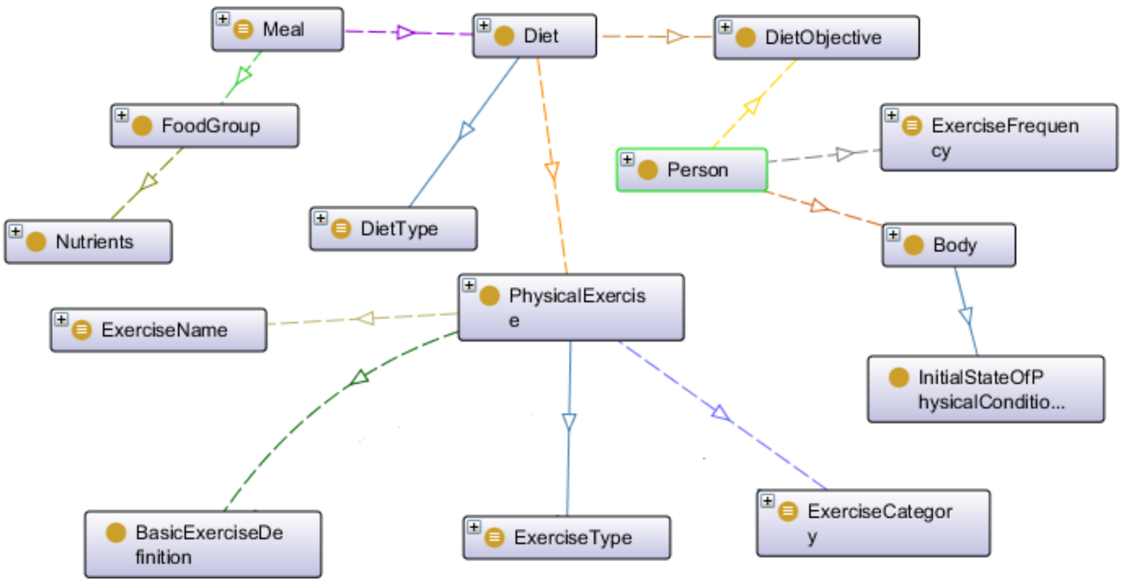
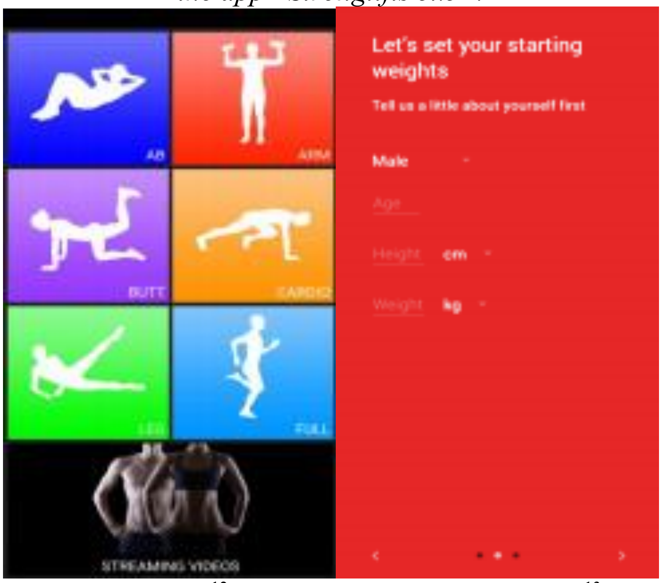


Figure 9 demonstrates some of the properties and classes used in our model. On the left, a screenshot from the app "Daily Workouts", with the properties defined in the class *Body*. On the right, a screenshot from the app "Stronglifts 5x5", with common properties asked for the user profile, defining the *Person* class (like age, gender, height and weight), which was the most common request for almost all applications.

Figure 9: On the left, a screenshot from the app "Daily Workouts". On the right, a screenshot from the app "Stronglifts 5x5".



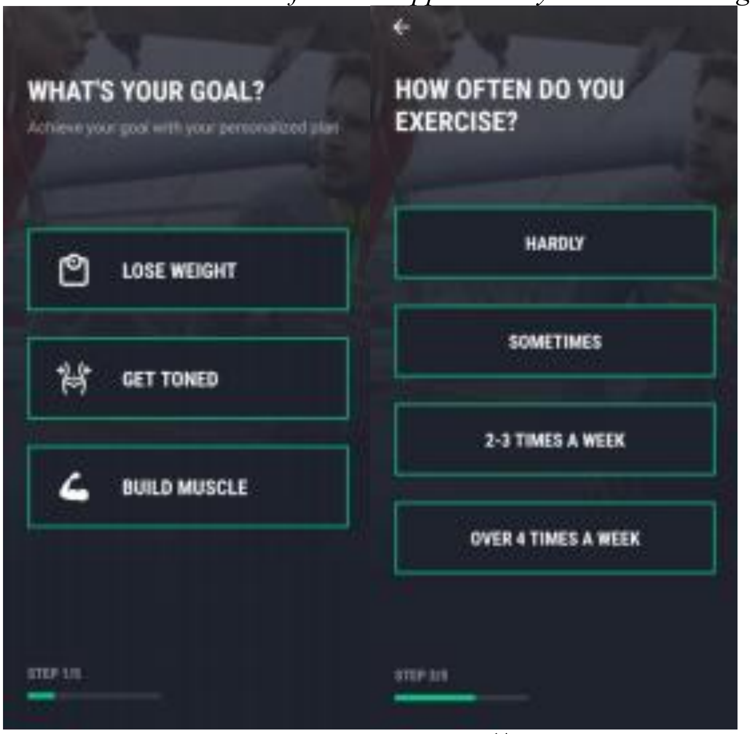
Source: "Daily Workouts"[12] app on the left. "Stronglifts 5x5"[13] on the right.

Another example in Figure 10 shows two screenshots from the app "30 Day Fitness Challenge". On the left, there are some of the properties defined in the class *DietObjectiveChosen* (e.g., *loseWeight* and *buildMuscleMass*). On the right, there are some properties used to inspire the definitions of the class ExerciseFrequency (e.g., *occasionallyExercise*, *oftenExercise* and *ratherExercise*), related to which frequency users practice any type of physical exercise.

[12] https://play.google.com/store/apps/details?id=com.tinymission.dailyworkoutsfree&hl=en

[13] https://play.google.com/store/apps/details?id=com.stronglifts.app&hl=en

Figure 10: Two screenshots from the app "30 Day Fitness challenge".



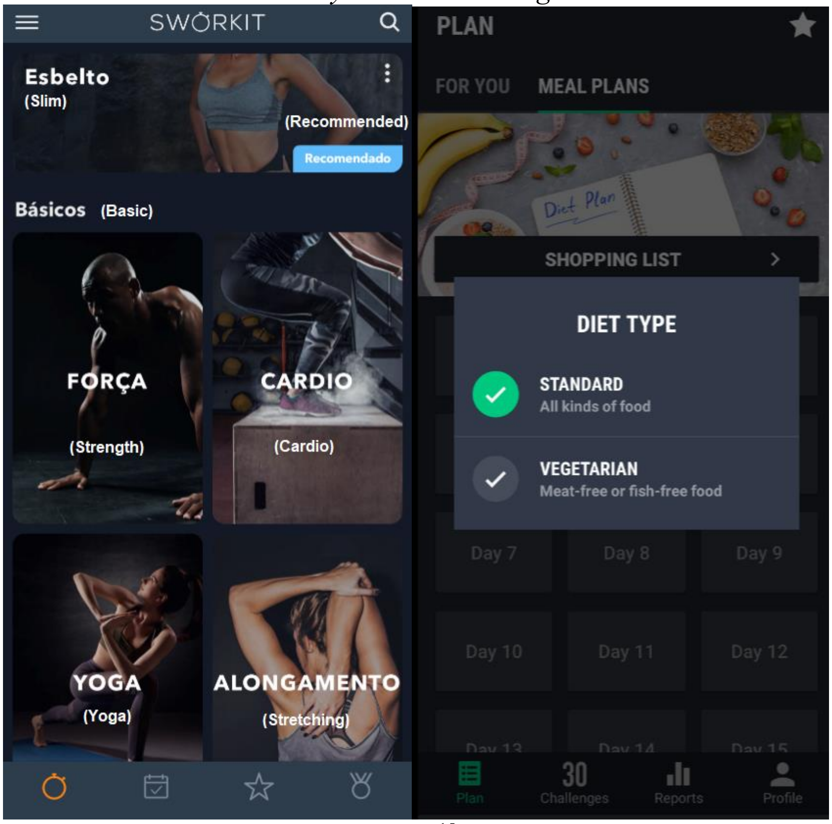
Source: "30 Day Challenge"[14] app.

The multidimensional feature is a key aspect of health research, and the health dimensions usually have a high mutual impact and cannot be analyzed in isolation. For example, an app for physical activity intervention (apps that inspired the class *PhysicalExercise*) can indicate a health program for losing weight, usually associated with a diet objective (*DietObjectiveChosen class*), which could also be to maintain or increase weight. This objective helps to define the type of exercise (*ExerciseType*) that could be performed in order to help users to achieve a diet objective. Moreover, a person's motivation will have an impact in controlling the diet (*Diet*), affecting the latter's initial state, in alignment with the diet goal (*DietObjective*) and affecting the exercise frequency (*ExerciseFrequency*).

Figure 11 exemplifies types of exercise, which include strength, cardio, yoga, stretching and others.

[14] https://play.google.com/store/apps/details?id=com.popularapp.thirtydayfitnesschallenge&hl=en

Figure 11: On the left, a screenshot from the app "Sworkit". On the right, a screenshot from the app "30 Day Fitness Challenge".



Source: On the left, adapted from the IOS app "Sworkit"[15]. On the right, the app "30 Day Fitness Challenge".

Regarding diet, another class that was analyzed within workout apps is *Meal*, which includes the type of meal, food group and nutrients. It serves to prepare a diet plan to help achieve body goals, as in Figure 12. The presence of nutrition concerns within workout apps shows how these two areas are intertwined. Thus, it makes sense to consider this part of the user's life as previously discussed, since Miller and others (1997) and Caudwell and others (2009) demonstrated that even when exercise energy expenditure is high, a healthy diet is still required for weight loss to occur in many people. Figure 12 also shows other dietary concerns within the workout app.

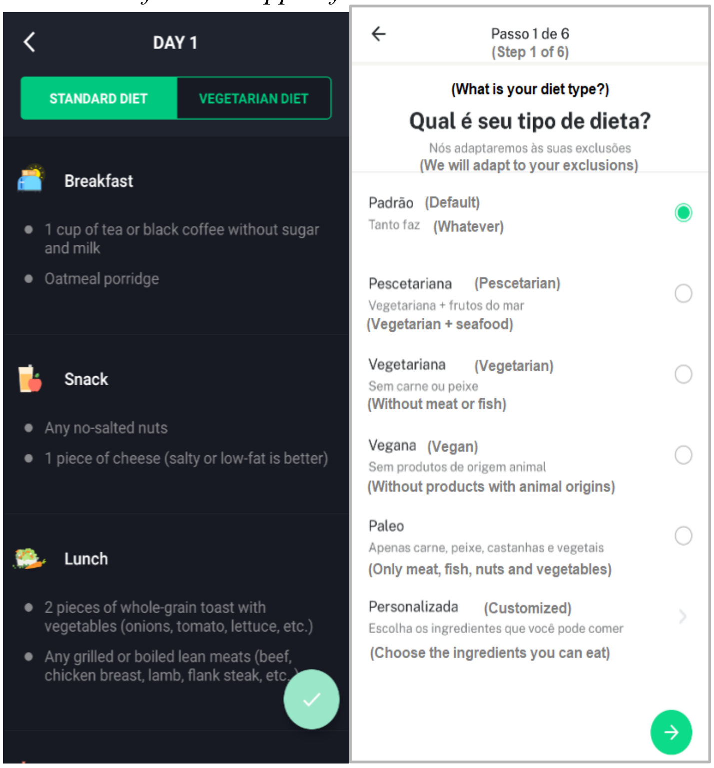
In Figure 12, on the left, it's possible to see some of the properties defined to the class *Meal*. The same figure on the right presents a screenshot from the app "8fit Workouts & Meal Planner", which includes a meal planning based on the type of diet (class *DietType*) that the user follows; and how physical exercises can be correctly balanced with right portions of meals throughout the day. The app "8fit Workouts & Meal Planner" really focuses on the importance of the nutritional part in training, offering a good number of options to users, such as number of meals per day, types of recipes preferred, and types of foods that should be avoided. Figure 13 illustrates part of this application, particularly related to the nutrition habits.

As previously discussed, motivation also affects the diet and exercise behaviors, because the more motivated a person is to follow the plan offered by the app or by a

[15] https://apps.apple.com/us/app/sworkit-fitness-workout-app/id527219710

professional, the better this person will perform and increase the *ExerciseFrequency*. Figure 14 portraits such motivational traits to keep the user as engaged as possible.

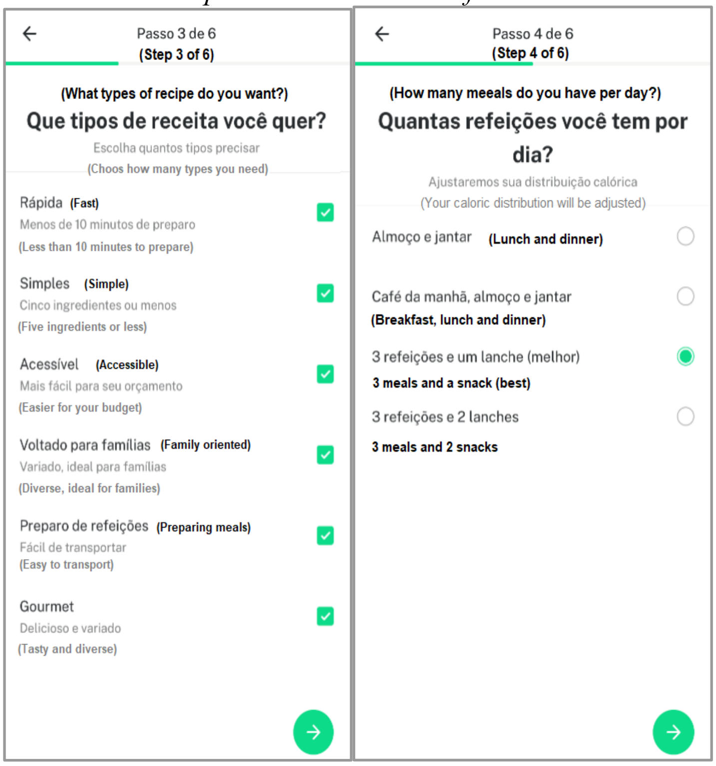
Figure 12: On the left, a screenshot from the app "30 Day Fitness Challenge". On the right, a screenshot from the app "8fit Workouts & Meal Planner".



Source: On the left, the app "30 Day Fitness Challenge". On the right, adapted to English the app "8fit Workouts & Meal Planner"[16].

[16] https://apps.apple.com/us/app/8fit-workouts-meal-planner/id866617777

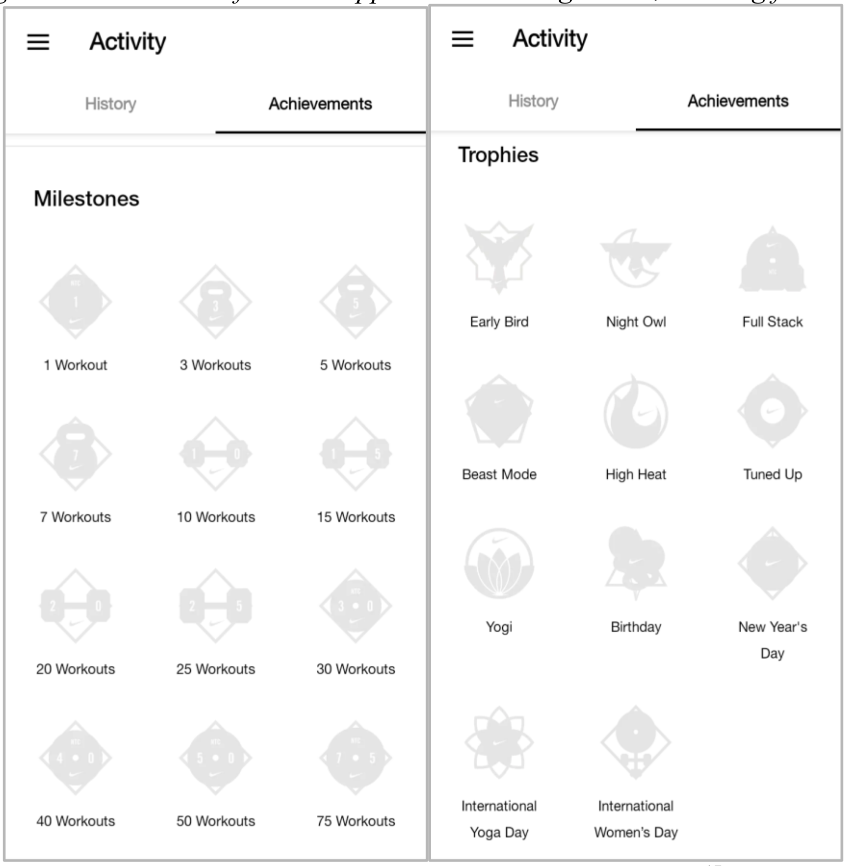
Figure 13: Preferred diet options offered by the app "8fit Workouts & Meal Planner", as a complement to workout definition.



Source: Both screenshots adapted to English from the app "8fit Workouts & Meal Planner".

The motivational features shown in Figure 14 use game components (achievements with badges, for instance) that characterize the concept of Gamification, that according to Barreto and others (2016) means the addition of game elements in non-game contexts such as productivity, finances, health care, education, sustainability and others, to improve user experience and engagement. Another aspect is how emotional-related areas, for example, can affect the diet or how a person faces an exercise routine. Factors such as stress and anxiety can lead the user to eat in an uncontrolled way or even impact the quality of sleep, a factor that is also related to *PhysicalHealth*. The latter contains applications that seek for measuring and helping with blood pressure, heartbeat, sleep quality, breathing, pregnancy, among others (see Table 7 in Appendix). Figure 15 and 16 illustrates other *PhysicalHealth* QoL related apps.

Figure 14: Screenshot from the app "Nike Training Club", showing features.



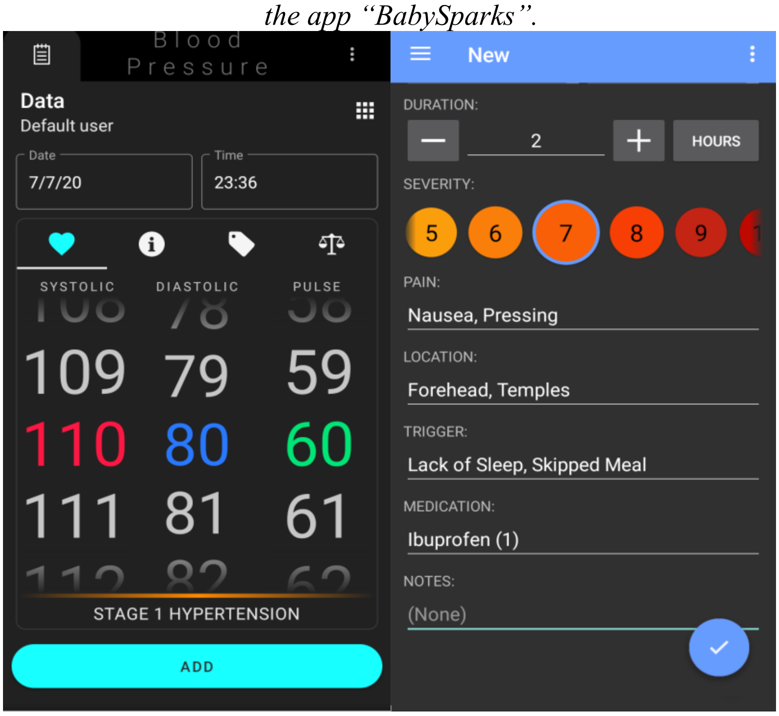
Source: Screenshot from the app "Nike Training Club"[17].

[17] https://apps.apple.com/us/app/nike-training-club/id301521403

Figure 15: On the left, a screenshot from the app "Sleepzy". On the right, a screenshot from the app "BabySparks".

Source: On the left, a screenshot from the app "Sleepzy"[18] adapted to English. On the right, a screenshot from the app "BabySparks"[19].

Figure 16: On the left, a screenshot from the app "Blood Pressure". On the right, a screenshot from the app "BabySparks".



---

[18] https://apps.apple.com/us/app/sleepzy-sleep-cycle-tracker/id1064910141

[19] https://apps.apple.com/us/app/babysparks-development-app/id794574199

Source: On the left, a screenshot from the app “Blood Pressure”[20]. On the right, a screenshot from the app “Headache Log”[21].

The discussion of this section demonstrates some association of QoL dimensions, and the importance to understand QoL aspects as a whole, also representing the context in which people are inserted when they search for a healthier life. Our goal was to exemplify the ideas that the ontology should present, the thinking behind the existence of some connections, and what inspired the names of classes, subclasses, properties and connections.

3.4 USER INTERVIEW FOR TRAINING DATA SETS

We conducted an interview with 20 applications users to obtain the data required to create instances (individuals) to our ontology. The elements of the ontology were used to represent a small daily user context of what could be considered a healthy person or not. This same data was used to train the neural network and to generate explanations. The interview questions aimed at collecting basic personal data required by the applications (as user’s *age*, *gender*, *height*, *weight* and *healthConditioning*), all part of *Person’s* class properties defined in our ontology. Moreover, we also collected user's daily exercise routine in terms of frequency, the objective of their adopted diet, the quality of their alimentation routine, the number of days a week that they are willing to eat healthy, and their idea about their current physical health. Moreover, we have the attribute *gender* (with *1.0* for women and *0.0* for men) and the classification of *health* (for our training, the *healthy* user will be represented with *1* and the *unhealthy* with *0*).

The interview questions were as follows:

A. What is your weight, height, gender and age? Moreover, choose one option that better describes your current health conditions:
1 - Irregular/bad physical condition or with health pre-conditions;
2 - Regular physical condition;
3 - Good physical condition and no pre-conditions;

B. How often could you say you exercise? Choose one option that better describes your answer:
1 - Rather exercise;
2 - Occasionally exercise;
3 - Often exercise;

C. In terms of diet and weight, what is your current objective? Choose one option that better describes your answer:
1 - To lose weight;
2 - To maintain weight;
3 - To increase weight;

D. Which alimentation quality routine do you identify yourself associated with? Choose one option that better describes your answer:
1 - Irregular/bad alimentation quality;

[20] https://play.google.com/store/search?q=Blood%20Pressure&c=apps&hl=en

[21] https://play.google.com/store/apps/details?id=arproductions.andrew.headachelog&hl=en

2 - *Regular alimentation quality;*
3 - *Good alimentation quality;*

E. *How many days a week are you committed to having a healthy diet (0, 1, 3, 5 or 7 days)?*

The complete mapping of properties and responses can be seen in Figure 17.

Figure 17: Responses mapped into a spreadsheet for each property and individual, reaching their final classification in column 'health'.

Person	age	gender	height	weight	hasExerciseFrequency	hasDietObjective	DietObjective	qualityOfKilogramsIngested	quantityOfDaysPerWeekOnDiet	healthConditioning	health
Person-001	27.0	1.0	1.68	66.0	occasionallyExercise	loseWeight	DietObjective-001	2	1	2	0
Person-002	27.0	1.0	1.68	63.0	occasionallyExercise	loseWeight	DietObjective-002	1	1	1	0
Person-003	15.0	1.0	1.63	40.0	ratherExercise	increaseWeight	DietObjective-003	2	1	3	0
Person-004	49.0	0.0	1.71	79.8	occasionallyExercise	maintainWeight	DietObjective-004	1	5	1	0
Person-005	47.0	1.0	1.61	90.1	ratherExercise	loseWeight	DietObjective-005	1	3	3	0
Person-006	26.0	0.0	1.62	90.0	occasionallyExercise	loseWeight	DietObjective-006	3	5	3	1
Person-007	30.0	1.0	1.56	70.5	occasionallyExercise	loseWeight	DietObjective-007	2	7	2	1
Person-008	61.0	0.0	1.70	80.1	oftenExercise	loseWeight	DietObjective-008	3	5	3	1
Person-009	60.0	1.0	1.54	61.0	ratherExercise	loseWeight	DietObjective-009	1	5	2	0
Person-010	34.0	1.0	1.67	75.4	oftenExercise	loseWeight	DietObjective-010	2	3	3	1
Person-011	28.0	1.0	1.57	50.1	ratherExercise	increaseWeight	DietObjective-011	2	0	3	0
Person-012	45.0	1.0	1.60	49.6	ratherExercise	maintainWeight	DietObjective-012	1	7	3	1
Person-013	50.0	0.0	1.80	96.7	ratherExercise	loseWeight	DietObjective-013	1	0	1	0
Person-014	21.0	1.0	1.58	48.2	occasionallyExercise	maintainWeight	DietObjective-014	3	0	3	1
Person-015	29.0	1.0	1.59	47.8	ratherExercise	maintainWeight	DietObjective-015	2	5	1	1
Person-016	35.0	0.0	1.75	78.2	oftenExercise	increaseWeight	DietObjective-016	3	7	3	1
Person-017	37.0	0.0	1.70	83.2	occasionallyExercise	maintainWeight	DietObjective-017	3	5	1	1
Person-018	25.0	1.0	1.80	67.4	oftenExercise	increaseWeight	DietObjective-018	2	3	3	1
Person-019	24.0	0.0	1.78	88.7	ratherExercise	maintainWeight	DietObjective-019	3	5	2	1
Person-020	22.0	1.0	1.63	59.8	occasionallyExercise	maintainWeight	DietObjective-020	3	5	2	1

An example of a healthy person is the instance Person-016, who exercises often, has a good alimentation routine, is willing to follow a diet 7 days a week, and already has a good physical conditioning (without any pre-conditions, such as cardiorespiratory health conditions or possible diabetes). This person was perceived as healthy (in other words, 'health' = 1). As a person considered unhealthy, the example of Person-002 earns health = 0, because he/she occasionally exercises, wishes to lose weight, has an unhealthy eating routine and is willing to devote only one day to a healthy diet.

3.5 NEURAL NETWORKS EXPERIMENTS

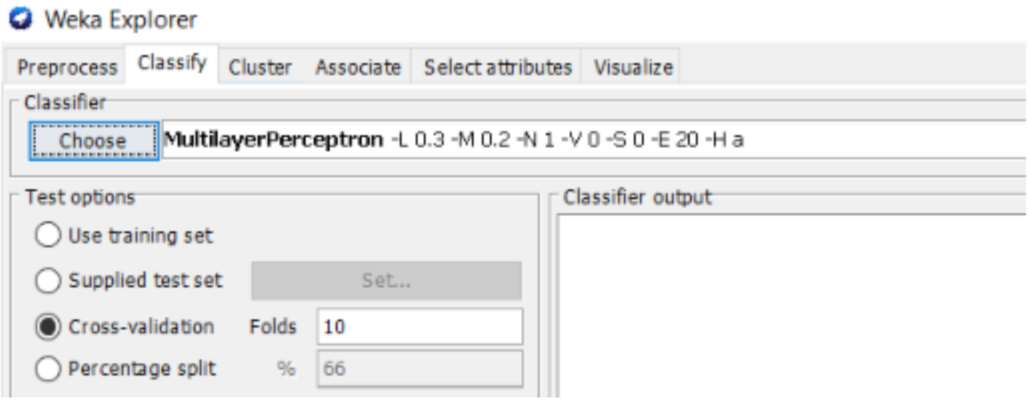
The main function of *mHealth* applications is to assess health data about their users. Our analysis showed that some of these apps try to relate this data to return simple interventions. However, more powerful conclusions are given by inductive reasoning processes, which can use this input data and indicate, for example, if mobile users have some cardiac problem or a high possibility to have such problems. Based on the parameters raised in the construction of the ontology, such as age, gender, stress level, blood pressure, among others; a case study was specified to show how an inductive reasoning process could use this data to make conclusions about mobile users. Moreover, the most important result of this stage of the research was to find an accuracy baseline that allows a comparison with the accuracies of explanations algorithms.

It is important to observe that the literature presents a lack of longitudinal and multidimensional studies in *mHealth*, Manea and Wac (2018). Thus, a proper dataset for experimentations is still not available. The study of the QoL Lab conducted by Manea et al (2019), for example, is one of the few ongoing efforts in such a direction. Therefore, the set of instances used in this experiment is only represented by the data obtained using our previous questionnaire. Apart from its small size (20 instances), it is enough to conduct some tests of concepts.

For the construction of the Neural Network, this project used an open-source software called Weka[22], developed at the University of Waikato (New Zealand). According to Hornik and others (2008) this software brings a comprehensive collection of machine-learning algorithms for data mining tasks written in Java and released under the GPL (General Public License). The experiment used an *arff* file, which stands for Attribute-Relation File Format (Weka input format). This file contains a list of instances representing the results of the questionnaires. The correspondent ontology representation of this file can be seen in Section 4, where the results are exposed.

The experiment used the *Multilayer Perceptron* function, which is a traditional function used for Neural Network problems. The idea was to perform several tests to tune the network and obtain an "as best as possible" accuracy. Figures 18 and 19 show the parameters used in this experiment.

Figure 18: Multilayer Perceptron function and its parameters.



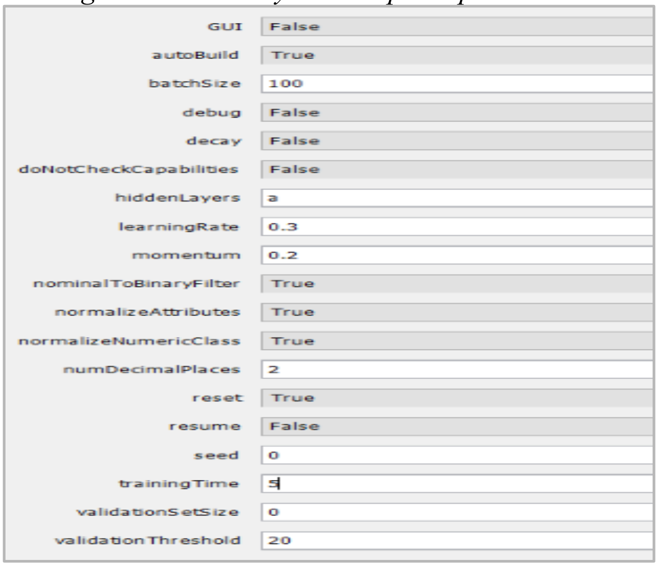
During the experiment, the parameter *trainingTime* was set with different values, since this parameter defines how much time the algorithm will spend training. The longer is the time, the better the result tends to be. However, there is a saturation point and, after that point, longer times do not affect the accuracy. In addition to the *trainingTime* parameter, we also tested the algorithm by modifying the *hiddenLayers* parameter, which initially had its value set to *'a'*. The value *'a'* means the algorithm generates a number of hidden layers, defined by adding the number of attributes plus the number of classes, and then dividing by two, as described below:

$$hidden\_layers\_amount = (attribute\ count + class\ count)/2$$

Finally, the *learningRate* and *momentum* parameters were also modified in order to try better accuracies.

[22] https://www.cs.waikato.ac.nz/ml/weka

Figure 19: Multilayer Perceptron parameters.



GUI	False
autoBuild	True
batchSize	100
debug	False
decay	False
doNotCheckCapabilities	False
hiddenLayers	a
learningRate	0.3
momentum	0.2
nominalToBinaryFilter	True
normalizeAttributes	True
normalizeNumericClass	True
numDecimalPlaces	2
reset	True
resume	False
seed	0
trainingTime	5
validationSetSize	0
validationThreshold	20

Figure 20 shows the final arff file that represents the data collected in the interview (Section 3.4):

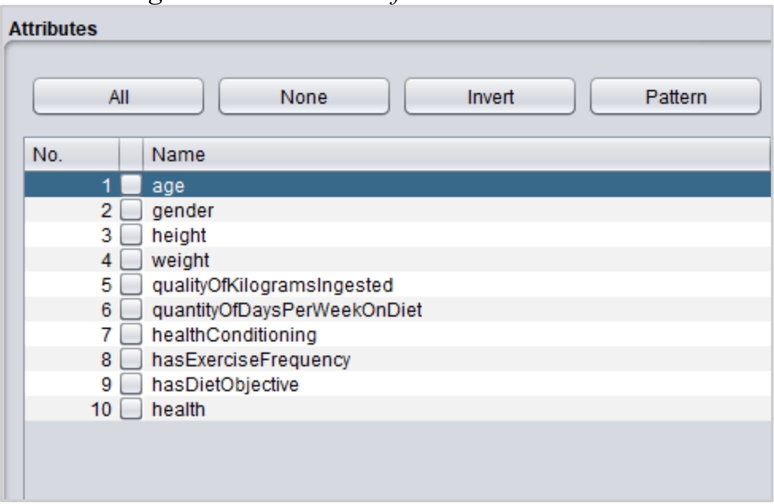
Figure 20: Final arff file.

```
@RELATION is_healthy
@ATTRIBUTE age                NUMERIC
@ATTRIBUTE gender             NUMERIC
@ATTRIBUTE stressLevel        NUMERIC
@ATTRIBUTE bloodPressure      NUMERIC
@ATTRIBUTE cholesterolamount  NUMERIC
@ATTRIBUTE sugaramount        NUMERIC
@ATTRIBUTE restingquality     NUMERIC
@ATTRIBUTE maxheartrate       NUMERIC
@ATTRIBUTE stepsaverageperday NUMERIC
@ATTRIBUTE heartbehaviourfromrestingtoaction NUMERIC
@ATTRIBUTE heartbehaviour     NUMERIC
@ATTRIBUTE minimumdaysperweek NUMERIC
@ATTRIBUTE meanbloodglucose   NUMERIC
@ATTRIBUTE health             {0, 1, 2, 3, 4}

@DATA
63.0,1.0,1.0,145.0,233.0,1.0,2.0,150.0,0.0,2.3,3.0,0.0,6.0,0
67.0,1.0,4.0,160.0,286.0,0.0,2.0,108.0,1.0,1.5,2.0,3.0,3.0,2
67.0,1.0,4.0,120.0,229.0,0.0,2.0,129.0,1.0,2.6,2.0,2.0,7.0,1
37.0,1.0,3.0,130.0,250.0,0.0,0.0,187.0,0.0,3.5,3.0,0.0,3.0,0
41.0,0.0,2.0,130.0,204.0,0.0,2.0,172.0,0.0,1.4,1.0,0.0,3.0,0
56.0,1.0,2.0,120.0,236.0,0.0,0.0,178.0,0.0,0.8,1.0,0.0,3.0,0
62.0,0.0,4.0,140.0,268.0,0.0,2.0,160.0,0.0,3.6,3.0,2.0,3.0,3
```

Figure 21 shows all of the attributes used. In this case, we have nine input attributes and one output (health) with two possible values. This strategy is very poor in terms of data analysis, since it is a black box method, as earlier introduced in Section 2.5. This means that it is not possible to have an easy interpretation of these values, which, by the way, is one of the motivators of our research.

Figure 21: Attributes of the Neural Network.



This experiment was important to exemplify a black-box technique and give a notion about the accuracy that an inductive reasoning process could obtain, considering this particular dataset. After this step, our next activities employed algorithms to generate explanations aimed at obtaining precisions as close as possible to the results of this case study (Section 4.4).

3.6 GENERATING EXPLANATIONS WITH ONTOLOGY DATA

The use of inductive reasoning is very common in the health area and the literature presents several approaches that use datasets to create classification and regression models. However, these models only return the final results (e.g., classification for new instances) so we need to specify additional methods that can explain how these results were generated. Our approach uses the complex structure of available background knowledge over the process of learning hypotheses. As this knowledge is represented in the form of ontologies, this explanation will be given as a logical formula – class expression in DL – where all (or as many as possible) positive examples are instances of this expression, while none (or as few as possible) negative examples are not instances of such an expression. To generate this type of explanation, we are employing the DL-Learner framework, which, conforming to Bühmann and others (2016), supports inference reasoning in the form of supervised machine learning, using OWL representations (ontologies) as background knowledge.

In our study, the positive and negative examples are represented by the same data instances of the previous Weka experiment: *healthy* (as a positive example) and *unhealthy* (as a negative example). As this data is originally saved as an Excel datasheet, we employed a Protégé's embedded plugin, called Cellfie[23], which automatically generates the instances of our ontology according to the data of this datasheet. The rules that control this import process are illustrated in Figure 22.

[23] https://github.com/protegeproject/cellfie-plugin

Figure 22: Plugin to import data from spreadsheet into the ontology by setting rules to columns.

Sheet Name	Start Column	End Column	Start Row	End Row	Rule
Plan1	H	K	2	+	Individual: @I* Types: DietObjective Facts: hasDietObjective @H*, qualityOfKilogramsIngested @J*(xsd:decimal), quantityOfDaysPerWeekOnDiet @K*(xsd:decimal)
Plan1	B	B	2	+	Individual: @B* Types: Person Facts: age @C*(xsd:decimal), gender @D*(xsd:decimal), height @E*(xsd:decimal), weight @F*(xsd:decimal), healthConditioning @L*(xsd:decimal), hasExerciseFrequency @G*, hasDietObjective @H*

The DL-Learner configuration file contains as main parameters: the ontology name, negative instances, positive instances, algorithm to be used, *noisePercentage*, and *maxExecutionTimeInSeconds*. Part of this configuration file can be seen in Figure 23.

Figure 23: Part of DL-Learner's customized configuration file, used to run the experiment.

```
lp.positiveExamples = {"isa:Person-006", "isa:Person-007",

lp.negativeExamples = {"isa:Person-001", "isa:Person-002",

// create learning algorithm to run
//(ELTL) (ISLE)
alg.type = "celoe"
alg.maxExecutionTimeInSeconds = 120
alg.noisePercentage = 20.0
alg.writeSearchTree = false
alg.useMinimizer = false
```

To conduct the experiments regarding explanations, we first generated explanations using the original ontology as background knowledge. Secondly, we modified the ontology, including some unnamed classes. These changes and the ontologies will be discussed in details in Section 4.2. The concept of unnamed classes is very useful in ontologies because they can automatically classify instances in clusters that respect one or more constraints. For example, the unnamed class *Adult* represents a subset of persons that are in a specific age interval (>= 18 and <= 59).

The explanations were evaluated using qualitative (readability) and quantitative (accuracy and F-measure) metrics. The DL-Learner algorithms used to generate explanations in this study were the Class Expression Learning for ontology Engineering (CELOE) and EL Tree Learner (ELTL). The CELOE general idea as reported by Bühmann and others (2016) is to build a search tree based on a refinement operator and make use of a heuristic to find good candidates to look at. Meanwhile, ELTL is an algorithm optimized for learning trees using the idea of refinement operator.

Figure 24 shows an example of DL-Learner output using as configuration values: *alg.type* = "*celoe*", a*lg.maxExecutionTimeInSeconds* = *120* and *alg.noisePercentage* = *20.0*.

Figure 24: First experiment execution, with the parameters alg.type = "celoe", alg.maxExecutionTimeInSeconds = 120 and alg.noisePercentage = 20.0.

```
Solutions:
01: (height some decimal[>= 1.57]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 81,48%)
02: (height some deicmal[>= 1.56]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 82,76%)
03: (age some decimal[>= 48.0]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 81,48%)
04: PhysicalHealth or ((height some decimal[>= 1.57]) and (weight some decimal[<= 90.0])) (pred. acc.: 75,00%, F-measure: 81,48%)
05: PhysicalHealth or ((height some decimal[>= 1.56]) and (weight some decimal[<= 90.0])) (pred. acc.: 75,00%, F-measure: 82,76%)
06: PhysicalHealth or ((height some decimal[>= 48.0]) and (weight some decimal[<= 90.0])) (pred. acc.: 75,00%, F-measure: 81,48%)
07: PhysicalExercise or ((height some decimal[>= 1.56]) and (weight some decimal[<= 90.0])) (pred. acc.: 75,00%, F-measure: 82,76%)
08: Person and (height some decimal[>= 1.57]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 81,48%)
09: Person and (height some decimal[>= 1.56]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 82,76%)
10: Person and (age some decimal[<= 48.0]) and (weight some decimal[<= 90.0]) (pred. acc.: 75,00%, F-measure: 81,48%)
```

Figure 25 describes the parameters used as *noisePercentage*, which were used with both the original and modified ontologies.

Figure 25: noisePercentage value changes during the experiment's evolution.

Noise Percentage Parameter
value = 0.00
value = 20.0
value = 50.0

As previously discussed, the dataset used in this study consisted of 20 instances with 10 attributes (9 inputs and 1 outcome), which are described as follows. This description already indicates the unnamed classes used to extend the original ontology:

- **A1** - age range, which is divided into three unnamed classes: *Young* (<18), *Adult* (18-60), and *Older* (>60);
- **A2** - gender, which is divided into *Male* and *Female* unnamed classes;
- **A3** - height, indicates the individuals' height and can be divided into three unnamed classes (*low height* (<1.65), *medium height* (1.65 - 1.75) and *greater height* (>1.75));
- **A4** - weight, where it's considered three unnamed classes (*overweight*, *ideal weight* and *underweight*), this value is also calculated by some apps considering the height of the user;
- **A5 -** quality of kilograms ingested, which is divided into three unnamed classes (*bad quality*, *regular quality* and *good quality)*;
- **A6** - related to the number of days a person spends on a healthy diet, divided into three unnamed classes (*few days*, *weekdays* and *whole week*);
- **A7** - indicates if individuals may have some health problem. We could have here some classes related to preconditions, like diabetes, pregnancy, high sugar levels and high blood pressure. For example, if the pressure property level is higher than 140 mm Hg value, then the instance can be considered as part of the *HighBloodPressure* class.
- **A8** - linked to the number of days that a physical activity is currently performed in a week, connected to a data property with three potential values: *rather exercise*, *occasionally exercise* and *often exercise*;
- **A9** - is related to a data property with three possible values: *lose weight*, *increase weight*

and *maintain weight*;

- **A10** - prediction attribute that indicates if individuals may have some health problem (*true* or *false*).

  More details are presented in Section 4.

4 RESULTS

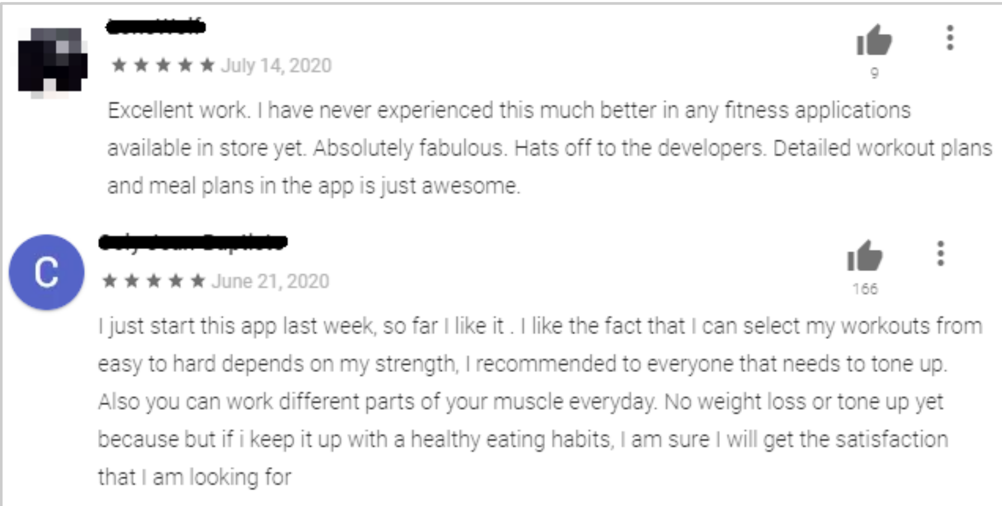
This chapter discusses the results of our study, presenting the questions' results (Section 4.1); the final ontology (Section 4.2); the inductive reasoning process with Neural Networks using Weka (Section 4.3); the experiment with the generation of explanations based on ontology data results and DL-Learner (Section 4.4), and the analysis of experiment metrics in terms of qualitative readability and quantitative accuracy (Section 4.5).

4.1 QUESTION RESULTS

A. What are the advantages and disadvantages of applications that aid a user's life, according to user ratings and related studies?

As people currently spend long times using or close to their smartphones, continuous actions of monitoring and assessment are now possible and can be conducted by mobile applications. In order, the mobile technology is able to generate proper just in time (right support at the right moment/context - location, and in the right amount) and customized (based on an individual's own performance and goal) assessments and interventions. Figure 26 exemplifies such examples of good assessments:

Figure 26: Users evaluation for the app "30 Day Fitness Challenge", from Physical Health group.

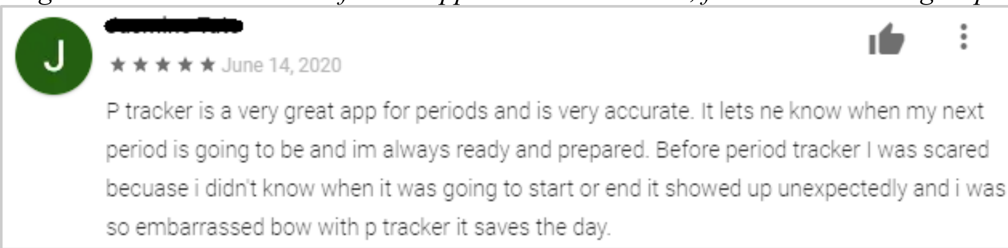


Source: App "30 Day Fitness Challenge".

According to the authors Ciman, Matteo, and Wac (2016), another advantage of *mHealth* applications is the pervasive and unobtrusive nature, compared to wearables, making it easier for user acceptance. It's possible to monitor an individual's parameters without wearables, which may cause users to feel uncomfortable or fear judgement from people. For example, the authors argue that when individuals are angry or stressed, they use smartphones differently and these differences can be used to computationally model and evaluate their stress state. Similarly to stress, several other QoL states can be unobtrusively obtained to monitor health parameters, while users carry out their daily activities.

Improving QoL aspects with mobile health applications also have the potential to improve the well-being feelings, such as described by the next user evaluation (Figure 27), where she states that she used to be afraid, but after using the app related to period tracking, she feels prepared:

Figure 27: User evaluation for the app "Period Tracker", from Health Care group.

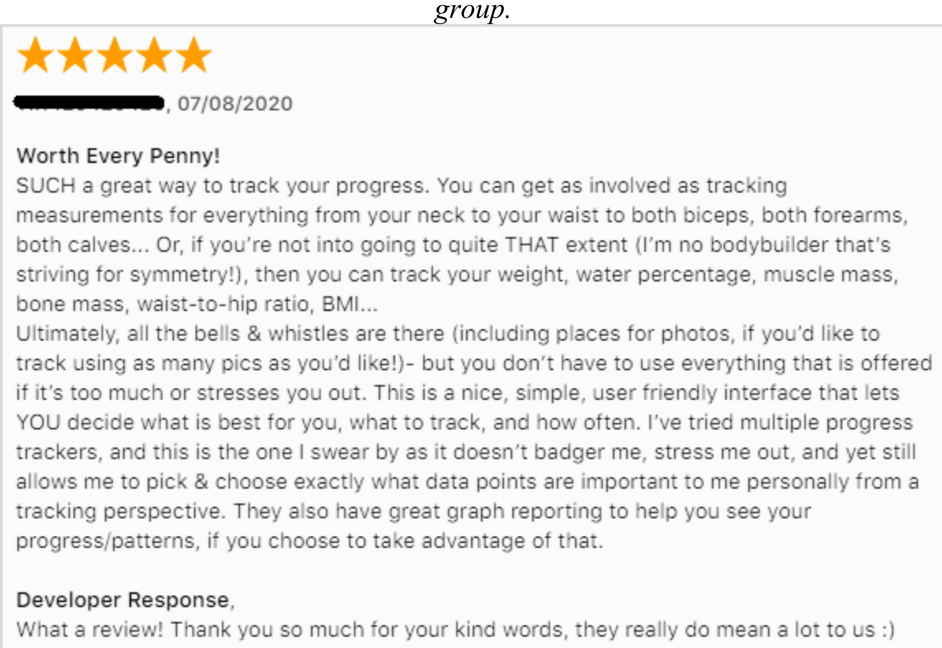

J
[redacted]
★★★★★ June 14, 2020
P tracker is a very great app for periods and is very accurate. It lets ne know when my next period is going to be and im always ready and prepared. Before period tracker I was scared becuase i didn't know when it was going to start or end it showed up unexpectedly and i was so embarrassed bow with p tracker it saves the day.

Source: App "Period Tracker"[24].

There's also the next review where the user emphasizes the apps' simplicity and user-friendly manner, and really enjoyed being in control of the actions that the customer thinks it's best for him, deciding what to track and how often. Good emotions were also a result from this *mHealth* solution.

On the other hand, there are also disadvantages. For example, the users' privacy is becoming more exposed, and big data centers like Google are getting more information about their users. This is a critical issue when health data is considered. The own question of exposure can cause harm to the user's health itself.

Figure 28: User evaluation for the app "Progress Body Tracker & Health", from Diet application group.


★★★★★
[redacted], 07/08/2020

Worth Every Penny!
SUCH a great way to track your progress. You can get as involved as tracking measurements for everything from your neck to your waist to both biceps, both forearms, both calves... Or, if you're not into going to quite THAT extent (I'm no bodybuilder that's striving for symmetry!), then you can track your weight, water percentage, muscle mass, bone mass, waist-to-hip ratio, BMI...
Ultimately, all the bells & whistles are there (including places for photos, if you'd like to track using as many pics as you'd like!)- but you don't have to use everything that is offered if it's too much or stresses you out. This is a nice, simple, user friendly interface that lets YOU decide what is best for you, what to track, and how often. I've tried multiple progress trackers, and this is the one I swear by as it doesn't badger me, stress me out, and yet still allows me to pick & choose exactly what data points are important to me personally from a tracking perspective. They also have great graph reporting to help you see your progress/patterns, if you choose to take advantage of that.

Developer Response,
What a review! Thank you so much for your kind words, they really do mean a lot to us :)

Source: App "Progress Body Tracker & Health"[25].

For some studies, it was necessary to fill out forms on a daily basis and continuously feed the databases with personal information. The invasion of privacy was also something that worried users, although the data is confidential. For example, the authors Ciman, Matteo, and

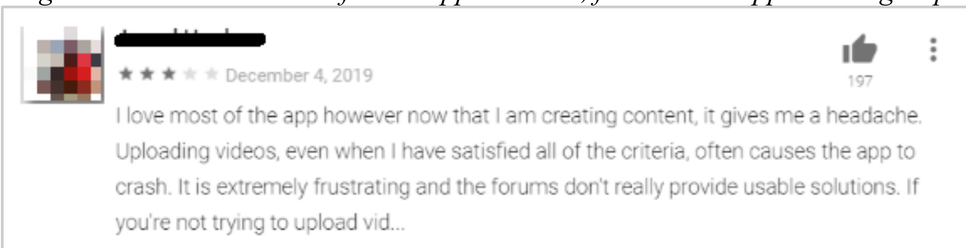
24 https://play.google.com/store/apps/details?id=com.period.tracker.lite

25 https://apps.apple.com/us/app/progress-body-tracker-health/id583840813

Wac (2016) collected data on screen ON/OFF, SMSs (number, length, the number of receivers, etc.), calls (receivers, duration, etc.) and the user location. They also collected information about audio, physical activity and communication data gathered during a working day, and heart rate variability. In this research, authors could conclude that users are not so engaged when they need to respond to questionnaires.

Based on the user rating from Figure 29, feelings of frustration may arise from the use of apps, mainly if the app is difficult to use or crashes frequently after a lot of effort has been given. This is important to show that bad applications can negatively affect the emotional state of their users rather than support them.

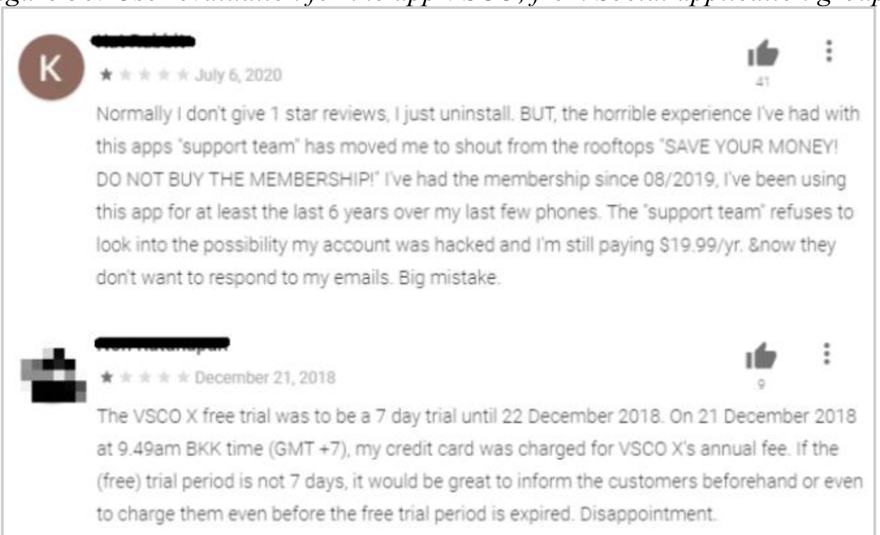
Figure 29: User evaluation for the app Linkedin, from Social application group.



Source: App "LinkedIn"[26].

Another problem that negatively affects the emotional state of users is the lack of a proper customer support, which leaves people feeling injured, disappointed and stressed, as exemplified in Figure 30.

Figure 30: User evaluation for the app VSCO, from Social application group.



Source: App "VSCO"[27].

B. Are there applications that connect more than one QoL dimension?

Most of the applications evaluate attributes from different QoL dimensions. For example, some applications link the *Emotional* dimension to Physical Health, aiming to

[26] https://play.google.com/store/search?q=linkedin&c=apps

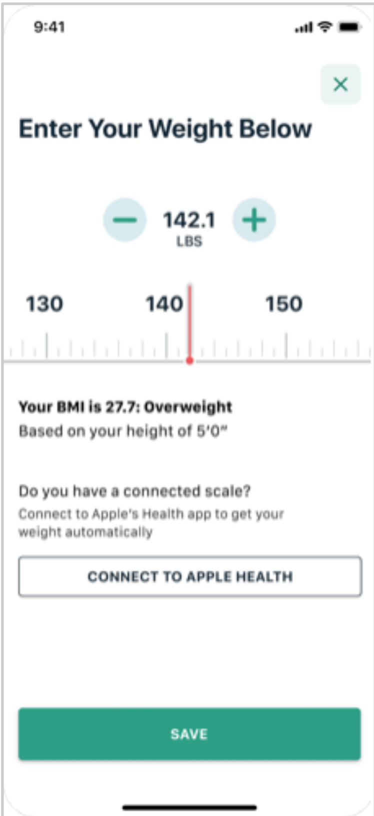
[27] https://play.google.com/store/apps/details?id=com.vsco.cam

understand what activities are practiced (or not) and may reduce factors such as stress, anxiety, among others. Another case is that all applications related to *Diet* are also connected with Physical Health, just as the opposite also happens. This means that many applications from Physical Health are related to parameters that make reference to *Diet*. Moreover, there are many cases where these applications provide their users a kind of Social Network with awards and possible interactions with other users (who are not professionals in the area). Thus, people with some health objectives can meet others with similar goals and, together, they can seek to reach their purposes. Examples of such apps and interconnectivity were shown in Section 3.1. Based on our study, about 90% of applications consider more than one QoL facet, which motivates the specification of unified representations that can in fact support holistic processes of reasoning.

C. What are the attributes most used by applications that link more than one group? Are there inputs that may be considered as more relevant?

Following the QoL dimensions described in Figure 1 (Section 2.2), the most successful types of downloaded applications, according to the analysis of rated stars and evaluations, are the ones belonging to the Physical Health group. Furthermore, the inputs common to several apps were basic information related to the user, such as age, gender, weight and height, as they are very important to determine health conditions in medicine. For example, Figure 31 illustrates attributes being requested for an app committed with heart conditioning.

Figure 31: Attributes weight and height often requested by Physical Health related apps.



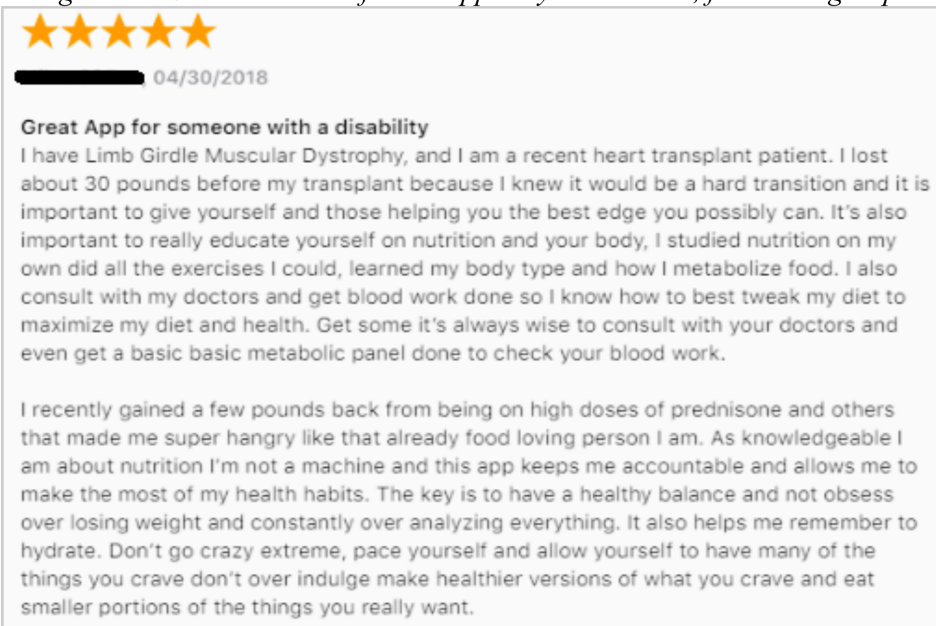
Source: App “Kardia”[28].

By the frequency that such pieces of information are requested, they can be considered the most important attributes. However, depending on the application, other attributes can also be considered as essential for the application's success. For example, an important input within *Diet* applications was to know how the user's daily diet was. Differently, the frequency of exercises could be more relevant for physical activity apps.

D. Is there a real impact reported by the use of these applications in daily life?

There is a real impact on the use of applications in people’s daily lives, as the use of smartphones is changing the way they behave, what they choose to do, how to do, and to show what they are doing. The amount of time that users spend on mobile phones only increases, and their impact on their real life is undeniable. The research has shown that domains with higher impact in users’ lives are: *Social*, *Entertainment* and *Physical Health*. These applications support the operationalization of new and current activities that are intrinsic to the user's life, as depicted in Figures 32, 33 and 34.

Figure 32: User evaluation for the app “MyFitnessPal”, from Diet group.


[redacted] 04/30/2018

Great App for someone with a disability

I have Limb Girdle Muscular Dystrophy, and I am a recent heart transplant patient. I lost about 30 pounds before my transplant because I knew it would be a hard transition and it is important to give yourself and those helping you the best edge you possibly can. It's also important to really educate yourself on nutrition and your body, I studied nutrition on my own did all the exercises I could, learned my body type and how I metabolize food. I also consult with my doctors and get blood work done so I know how to best tweak my diet to maximize my diet and health. Get some it's always wise to consult with your doctors and even get a basic basic metabolic panel done to check your blood work.

I recently gained a few pounds back from being on high doses of prednisone and others that made me super hangry like that already food loving person I am. As knowledgeable I am about nutrition I'm not a machine and this app keeps me accountable and allows me to make the most of my health habits. The key is to have a healthy balance and not obsess over losing weight and constantly over analyzing everything. It also helps me remember to hydrate. Don't go crazy extreme, pace yourself and allow yourself to have many of the things you crave don't over indulge make healthier versions of what you crave and eat smaller portions of the things you really want.

Source: App “MyFitnessPal”[29].

The image above describes an impact over nutrition and healthy habits, helping users to develop their diet plans and track calories for better health conditions. Figure 33 exemplifies another evidence of great impact in daily life activities, such as training. The user emphasizes the perfection of the app focused on workout, as he states that it pushes and motivates to accomplish goals.

[28] https://apps.apple.com/us/app/kardia/id579769143
[29] https://apps.apple.com/us/app/myfitnesspal/id341232718

Figure 33: User evaluation for the app "Nike Training Club", from Physical Health group.

★★★★★

[redacted], 12/31/2017

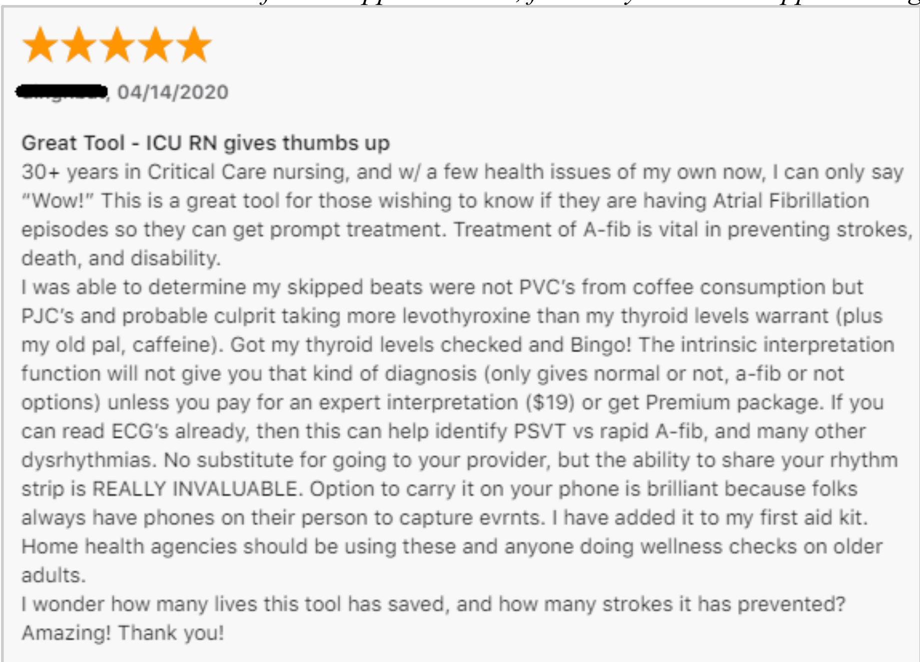
best workout app out there

i have tried many many MANY workout apps, and believe me i have gotten very agitated. i've googled and searched and downloaded so so many apps and none of them were right. they were either not free of just not a good overall app. this app is PERFECT. it shows you exactly what to do and adjusts your plan to your lifestyle. it helps you keep your workout goal as well as push yourself. it speaks aloud so you don't have to constantly look down at your phone when all you want to do is focus. there are such a variety of exercises and workout routines so you can pick the one that's right for you. this is the best workout app i have ever found and i will definitely be using it a lot his year. i've had it for only 2 days and i am ABSOLUTELY IN LOVE with it. it motivates me to workout and gives me checking points on my personal workout. this is just what i needed to turn my workout goals into a habit. i have never been more motivated to reach my workout goal. this app was life changing.

Source: App "Nike Training Club".

Another impact exemplified by Figure 34 is a report from a very satisfied user regarding the prevention and accessory of heart issues, where the user states that the app provides invaluable features and wonders how many lives have been saved from its use.

Figure 34: User evaluation for the app "Kardia", from Physical Care application group.



★★★★★

[redacted], 04/14/2020

Great Tool - ICU RN gives thumbs up

30+ years in Critical Care nursing, and w/ a few health issues of my own now, I can only say "Wow!" This is a great tool for those wishing to know if they are having Atrial Fibrillation episodes so they can get prompt treatment. Treatment of A-fib is vital in preventing strokes, death, and disability.

I was able to determine my skipped beats were not PVC's from coffee consumption but PJC's and probable culprit taking more levothyroxine than my thyroid levels warrant (plus my old pal, caffeine). Got my thyroid levels checked and Bingo! The intrinsic interpretation function will not give you that kind of diagnosis (only gives normal or not, a-fib or not options) unless you pay for an expert interpretation ($19) or get Premium package. If you can read ECG's already, then this can help identify PSVT vs rapid A-fib, and many other dysrhythmias. No substitute for going to your provider, but the ability to share your rhythm strip is REALLY INVALUABLE. Option to carry it on your phone is brilliant because folks always have phones on their person to capture evrnts. I have added it to my first aid kit. Home health agencies should be using these and anyone doing wellness checks on older adults.

I wonder how many lives this tool has saved, and how many strokes it has prevented? Amazing! Thank you!

Source: App "Kardia".

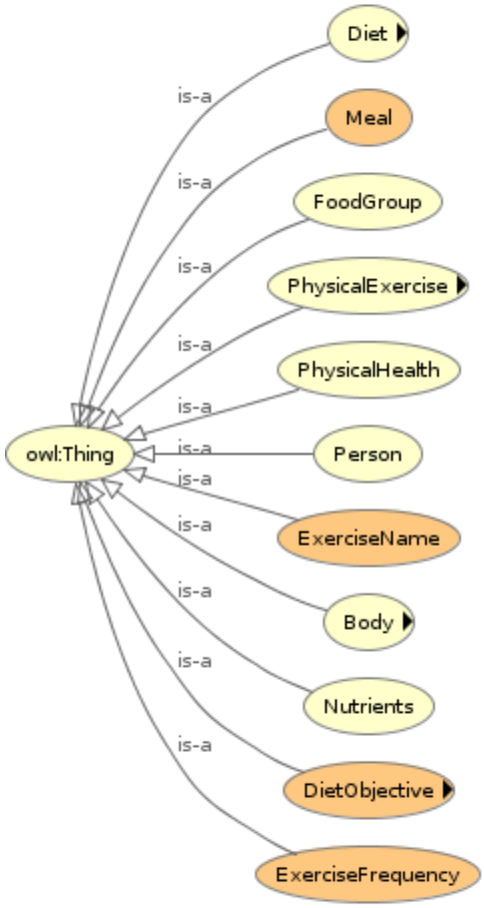
We can conclude that *mHealth* applications have positive impacts in specific health domains. However, its use for QoL as a whole is still challenging since such applications should involve the assessment of several attributes and consider all of them over holistic reasoning processes.

4.2 ONTOLOGY RESULTS

While applications generate data, the health area is also proposing strategies to properly organize such data. The NESTORE project (2019), for example, defined an ontology model focusing on models for Healthy Older People, with the domains of *Physiological Status*, *Physical Activity Behavior*, *Nutrition*, *Social Behavior*, *Cognitive* and *Mental Status*. The intention is to provide customized coaching to support healthy ageing, creating the motivation to take care of health. Although these models are valid to the characterization of individuals regarding the QoL dimensions, their support for holistic reasoning is not clear since they do not present relations between concepts of different dimensions that should be considered together when inferences are conducted. The study conducted by Pramono and others (2013) is another example that proposes a recommendation system based on ontology, which provides physical activity/exercise recommendations for diabetic patients. The domains considered on the model were related to exercise types of different intensity, frequency, and duration in accordance with the patient’s condition (age, complication, body mass index, calorie consumption, type of diabetes).

The ontology presented below (Figure 35) is the simplified version of our proposal, acquired from the studies of QoL facets, related *mHealth* apps, literature and iterative improvements.

Figure 35: Simplified ontology.



This ontology is simplified because it is only showing the main classes. Its complete version has subclasses and instances that would not fit well for visualization. The full ontology version is available in this link[30], but a preview is also available in Appendix. Using Protégé, the object properties below (Figure 36) were created to form relations between concepts:

Figure 36: Object properties defining relationship between concepts.

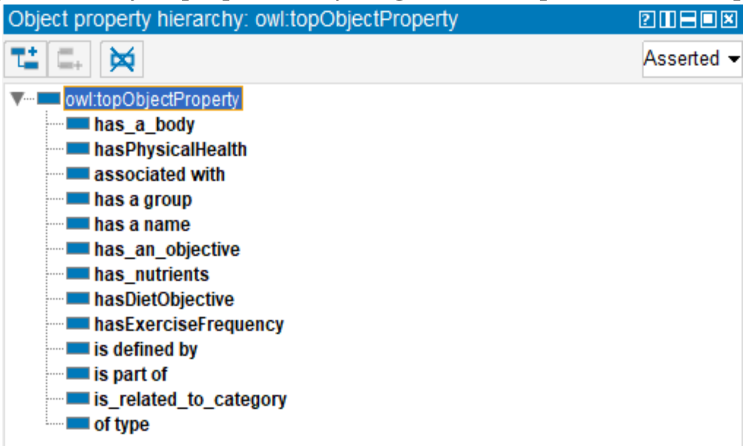
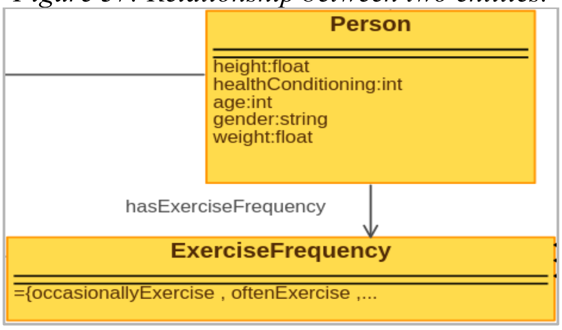


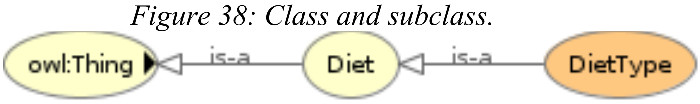
Figure 37 exemplifies a relation between the *Person* and *ExerciseFrequency* concepts, which indicates that a *Person* has an exercise frequency of one of the options defined as an enumeration by the *ExerciseFrequency* class:

Figure 37: Relationship between two entities.



Throughout the ontology, other entities have an implicit "*is a*" relationship, whenever a class is a subclass of another, such as *Diet* and *DietObjective* (subclass) classes (Figure 38).

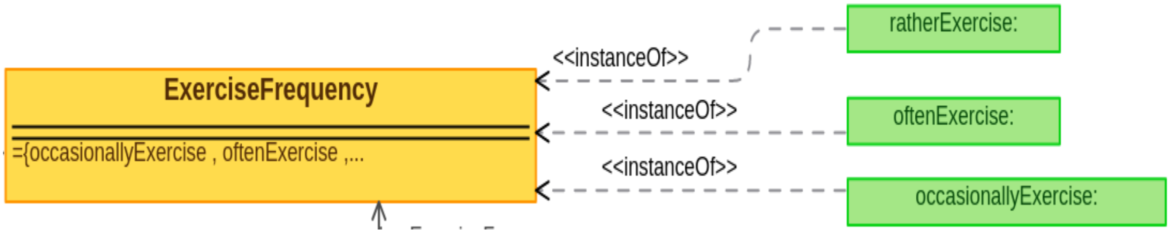
Figure 38: Class and subclass.



In the complete ontology model, there are some differences in the visualization, as some classes (enumerated classes) have explicit instances in green, which are used as the unique possible instances of this class (Figure 39). Other examples of enumerated classes are: *ExerciseCategory*, *ExerciseType*, *ExerciseName*, *DietType* and *DietObjectiveChosen*.

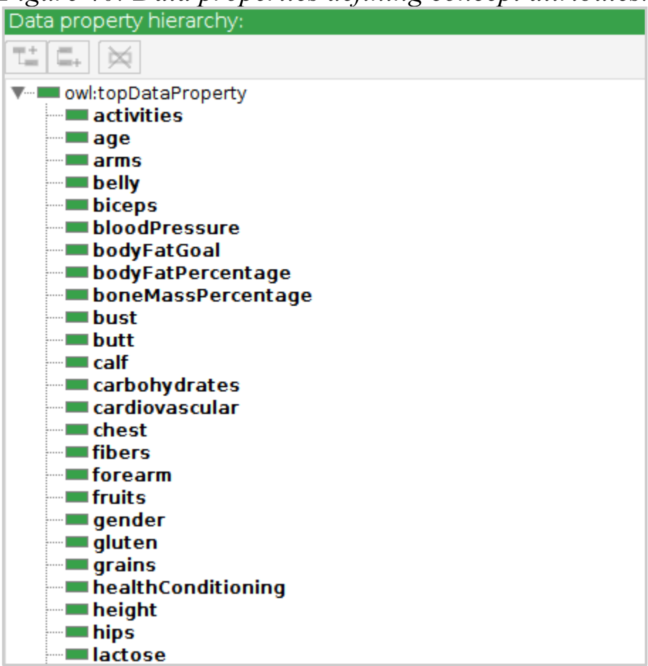
[30]http://owlgred.lumii.lv/online_visualization/my45

Figure 39: Individual instances of ExerciseFrequency Enum.



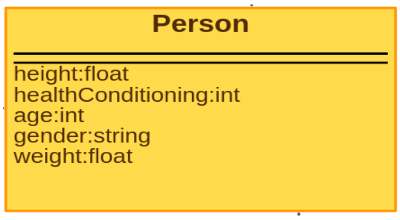
Regarding the properties, some of them are defined as functional. For example, the *hasExerciseFrequency* is a functional property, which means that only one instance of this relation is possible between instances of the involved concepts (relation 1:1). Other properties allow multiple relations (1:n). For example, *hasGroupFood* between *Meal* and *FoodGroup*. This means, a *Meal* is composed of one or more *FoodGroups*. Similarly, a *FoodGroup* has one or more relations with the *Nutrients* class. We also included data properties to the entities, representing the attributes identified during our previous literature and app analysis. The data properties are listed in Figure 40.

Figure 40: Data properties defining concept attributes.



Such data properties are linked to the classes they are related to. For example, the class *Person* (Figure 41) has the properties age, gender, height and others.

Figure 41: Class Person and its attributes.



We see this ontology provides an overview of different areas that impact each other and their relations. Considering this final model, it is possible to organize our knowledge base and specify the support for applications that intend, for example, to identify whether a person can be considered healthy or not. This process can be naturally conducted with the use of an inductive process.

Some metrics generated by Protégé regarding the number of axioms and related data are presented in Figure 42.

Figure 42: Ontology metrics generated by Protégé.

Metrics	
Axiom	**571**
Logical axiom count	**411**
Declaration axioms count	**154**
Class count	**17**
Object property count	**14**
Data property count	**52**
Individual count	**72**
Annotation Property count	**1**

4.3 NEURAL NETWORKS EXPERIMENTS WITH ONTOLOGY DATA

Section 3.5 described the steps of this experiment, which used a dataset mapped to an *arff* file loaded by Weka. The Neural Network Multilayer Perceptron algorithm was chosen and then modified through their respective parameters. The parameters modifications and results obtained in the tests will be better explained in this section.

As explained above, Multilayer Perceptron was employed as the inductive reasoner to create a model to classify persons as healthy or unhealthy. The accuracy provided by this model also gave a basis for the analysis of the explanations concerning their accuracies. Figure 43 shows the parameters configured in Weka and its final accuracy result of 75.00%.

Figure 43: Experiment parameters in Weka for 75% accuracy.

Parameter	*Value*
Number of layers	3
Neurons by layer	[13, 7, 2]
Number of epochs	500
Learning rate	0.3
Momentum	0.2
Instances	20
Class division	11 healthy, 9 unhealthy
Validation strategy	4-fold Cross-validation
Correctly classified instances	15 (75.00%)
Incorrectly classified instances	5 (25%)

Some parameters were modified trying to find a better accuracy. For example, the values 1, 10, 20, 100, 500, 800, 1000, 10.000 and 30.000 were used as the number of epochs (times that all dataset is used to training the network). The best value was defined as 500 (Figure 44) since, after this value, the training was not able to improve the accuracy.

Figure 44: Results for the MultilayerPerceptron function with trainingTime of 500.

```
=== Summary ===

Correctly Classified Instances          15               75      %
Incorrectly Classified Instances         5               25      %
Kappa statistic                          0.4898
Mean absolute error                      0.263
Root mean squared error                  0.4887
Relative absolute error                 54.5257 %
Root relative squared error             99.7334 %
Total Number of Instances               20

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,750    0,250    0,667      0,750    0,706      0,492    0,885     0,851     0
                 0,750    0,250    0,818      0,750    0,783      0,492    0,885     0,938     1
Weighted Avg.    0,750    0,250    0,758      0,750    0,752      0,492    0,885     0,903

=== Confusion Matrix ===

 a b   <-- classified as
 6 2 | a = 0
 3 9 | b = 1
```

Figure 45 illustrates the neural network architecture, which has thirteen, seven and two nodes [13,7,2] to respectively compose their three layers. Other architectures were also tested, such as the example in Figure 46.

Figure 45: MultilayerPerceptron function with GUI parameter set to 'True'.

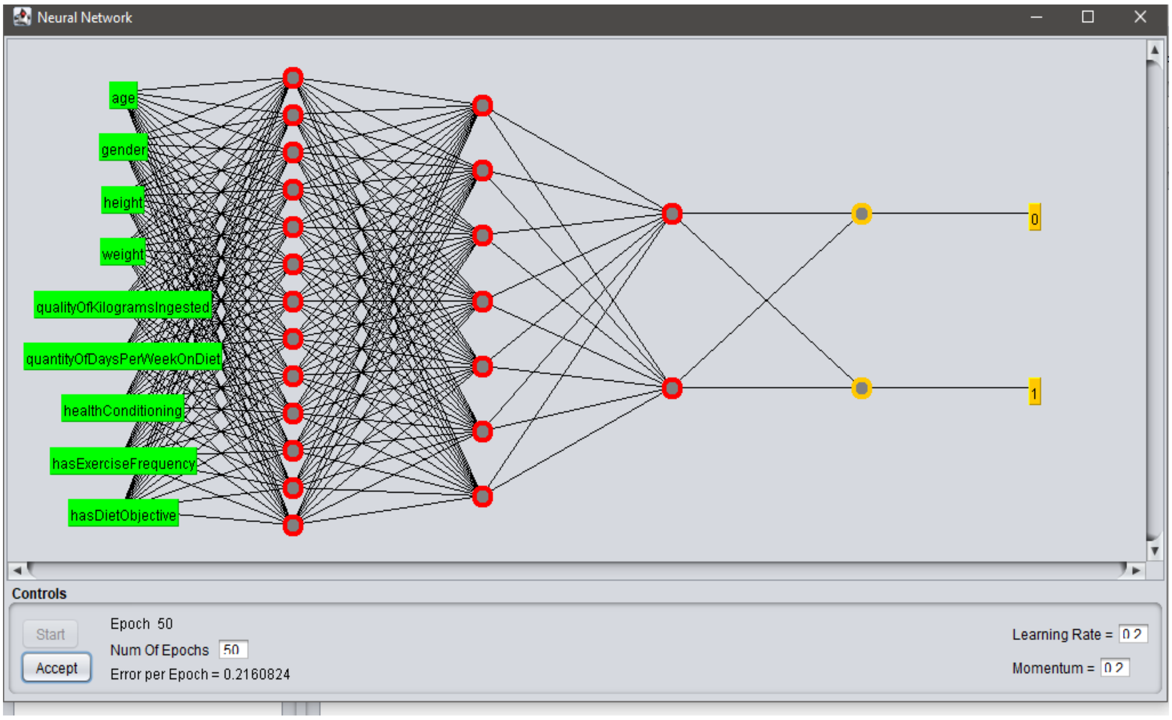
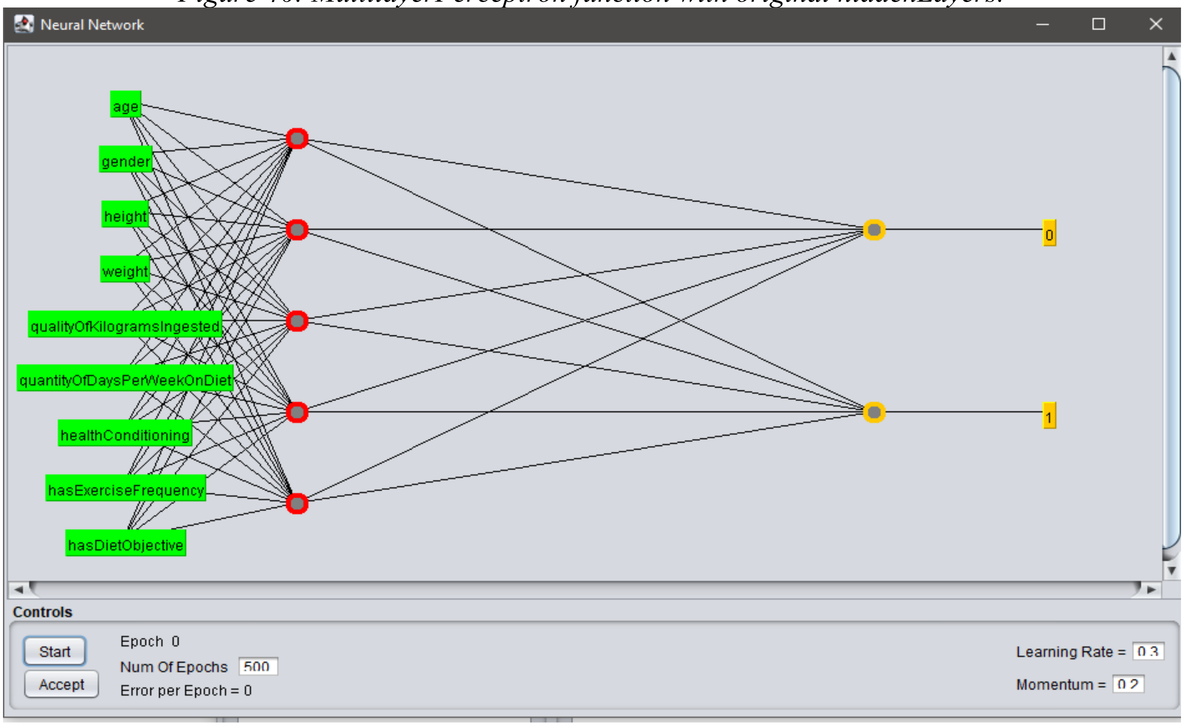


Figure 46: MultilayerPerceptron function with original hiddenLayers.



The accuracy provided by this model is important to give a basis for the analysis of the explanations. In order, algorithms used to generate the explanations should balance precision and readability. Then, good explanations must present accuracies closer to the accuracy of the neural network, trying to maintain the readability of such explanations as clearly as possible.

4.4 GENERATING EXPLANATIONS WITH ONTOLOGY DATA

For the generation of explanations in our study, we conducted two experiments through DL-Learner. The first experiment used the CELOE algorithm with 3 variations of *noisePercentage* parameter (*0.0*, *20.0* and *50.0*). This attribute indicates the approximated percentage of instances that can be considered as noise and discarded from the process. The learning problem parameter was set with *posNegStandard* value (positives and negatives examples are used over the process of learning). The *maxExecutionTimeInSeconds* parameter was set to *120s* for all cases. The second experiment used the same configurations, however employing the ELTL algorithm.

A premise that we also intend to evaluate is if the explanations readability improves when more refined ontologies are used as background knowledge. For example, ontologies that use unnamed classes. Figure 47 (with CELOE algorithm) and Figure 48 (with ELTL algorithm) show the results for all these combinations. The *Saturation time* in such tables means the moment when new increases do not present significant accuracy gains.

Figure 47: Experiment results with CELOE algorithm.

Ci	Unnamed classes	Noise	*Best accuracy*	*Best F-measure*	*Satur. time*
C1	Yes	0%	80.00%	85.71%	02m00s
C2	No	0%	75.00%	85.71%	02m00s
C3	Yes	20%	80.00%	85.71%	02m00s
C4	No	20%	75.00%	82.76%	02m00s
C5	Yes	50%	80.00%	85.71%	02m00s
C6	No	50%	75.00%	82.76%	02m00s

Figure 48: Experiment results with ELTL algorithm.

Ci	Unnamed classes	Noise	*Best accuracy*	*Best F-measure*	*Satur. time*
C7	Yes	0%	65.00%	77.42%	02m04s
C8	No	0%	75.00%	82.76%	02m00s
C9	Yes	20%	65.00%	77.42%	02m00s
C10	No	20%	75.00%	82.76%	02m00s
C11	Yes	50%	65.00%	77.42%	02m01s
C12	No	50%	75.00%	82.76%	02m00s

It is well-recognized that medical datasets are often noisy and incomplete due to the difficulties in data collection and integration according to Sáez and others (2016). For this reason, we used values of 20% and 50% as the percentage of noise allowed within the dataset. However, such variations did not bring differences for our experiments due to the small number of instances. Regarding the *Unnamed classes*, we could observe small differences in the metrics between C1/C2, C3/C4, C5/C6 (on CELOE experiment - Figure 47) and between C7/C8, C9/C10, C11/C12 (on ELTL side - Figure 48). Even using the dataset with 0% noise (C1, C2, C7 and C8), there was only a small difference of 5% in terms of percentage regarding the accuracy. In terms of execution time (*Saturation time*), none of the cases deflected significantly from what was defined, with an average time of 2 minutes (*maxExecutionTimeInSeconds* = 120s) per execution. For the CELOE experiment, our best *F-measure* was 85.7%, whereas for ELTL it was 82.76%, which is also a small difference.

It should be noticed that most data in our training dataset concerning the age is between 18 and 60. Thus, the majority of explanations was directed to this group. Another point is that our collected data sample was very simple, making some explanations of the experiments very similar, so we brought only one of each case (C_i), for a general and more pertinent view of the data and explanations.

Apart from C4, C6, C8, C10, C11 and C12 which have an equivalent explanation as the ones that are exposed, the other generated explanations can be read as:

C1: *(Adult or MediumHeight) and (weigh some decimal [<= 90]):*

> This means that 80.00% of healthy persons are adults or medium height (which is between 1.65 - 1.75) and have less than (or exactly) 90 pounds.

C2: *(height some decimal >= 1.56) and (weigh some decimal [<= 90]):*

> This means that 75.00% of healthy persons height is higher (or equal) to 1.56 and have less than (or exactly) 90 pounds.

C3: *(Adult and (not (MediumHeight))) or (weight some decimal [>= 72.9])*

> This means that 80.00% of healthy persons are adults and not medium height or weight higher than (or exactly) 72.9 pounds.

C5: *(Adult and (not (LowHeight))) or (weight some decimal [>= 66.7])*

> This means that 80.00% of healthy persons are adults and don't have low height (which means having less than 1.65 in our study) or have weight higher than (or exactly) 66.7 pounds.

C7: *Adult and (hasExerciseFrequency some ExerciseFrequency)*

> This means that 65.00% of healthy persons are adults and practice with some frequency physical exercise.

C9: *Adult and (hasDietObjective some DietObjective)*

> This means that 75.00% of healthy persons are adults and have a diet objective.

In general, this means that more than 80.00% of persons are classified as *Adults,* and from this group, approximately 61.12% are considered healthy. Also, regarding to height, the explanation results showed that heights between 1.65 and 1.75 are more likely to be healthy. Regarding weight, the results revealed that this property can be associated with height in order for the individual to be considered healthy or not. For example, a person with low height generally needs to have about 66.7 pounds to be recognized as healthy. Whereas a person with medium height would need more than 72.9 pounds and less than 90.0 pounds.

These explanations show that the use of unnamed classes (C1, C3, C5, C7, C9 and C11) improves the readability of the sentences, because rather than raw values, their semantics were used over the inductive reasoning. Moreover, in many cases, the qualification of the value (e.g., young, adult, or older) is more important than a simple value such as 25.5.

Next subsection discusses the accuracy of experiments, and also the quality of explanations generated as a result of our study.

4.5 DISCUSSION

After performing three experiments (one with Neural Networks and two with explanation algorithms), we found that the use of a neural network was important to generate an accuracy baseline (75%) for comparisons with the results of explanations algorithms. Therefore, we configured the DL-Learner to use the same database than Weka, and demonstrated that we could provide outputs closer to the accuracy baseline. For example, using CELOE and ELTL algorithms, the results reached an accuracy of 80.00%, 75.00% and 65.00%, which are around the baseline of 75.00%. In practical terms, the ontology design must allow these results and this may be a new form to evaluate ontologies.

About the explanations, the use of unnamed classes is another structural feature that affects the readability and the accuracy of explanations. The suggestion is to create unnamed classes that represent important domain concepts. For example, *SugarRich* and *CholesterolRich* could be concepts that aggregate value to our explanations. Therefore, several other classes could be created to cluster instances of the domain and emphasize their features.

Algorithms to generate explanations, such as the examples discussed by Bühmann and others (2016), are also based on an inductive reasoning system inspired by inductive logic programming. Thus, their execution time is another factor that affects the explanations quality. Our experiments, for example, obtained explanations with maximum generation time of 120s. More details and accuracy could be obtained with further tests using different maximum generation time. However, we need more complete datasets to actually identify such improvements.

The experiments also showed that the algorithms can generate results that only present restricted explanations. For example, one of the explanations generated was: *Adult and (hasExerciseFrequency some ExerciseFrequency)*. As *ExerciseFrequency* has three disjoint classes (*occasionallyExercise*, *oftenExercise* and *ratherExercise*), this explanation indicates that persons can be a member of any other class, rather than an exact class. Moreover, the algorithms also present a high number of configurations (CELOE for example, has about 30 parameters) so their efficiency may also depend on the values of such parameters.

5. CONCLUSION

The main contribution of this study is the proposal of a strategy to create background knowledge for *mHealth* applications that support holistic reasoning and explanations regarding the results obtained by means of inductive reasoning. For that end, we provided an ontology that integrates areas that are related and connected. We also discuss important elements that can affect the readability and accuracy of explanations, such as the use of unnamed classes and configuration of the explanation algorithms.

However, we recognize that for a better refinement of the explanations quality and prediction cases, there are some issues that can be improved. The number of people interviewed for the experiments database, together with a high variation of data (e.g., age variations), would make the outputs more robust and reliable. Thus, the dataset would be filled with balanced information and be in fact representative. Secondly, with a larger number of questions in the interview, and an extension of areas related to health (such as the areas removed from this research, such as *Emotional* classes), it would be possible to have a greater view of user's full context, their relations, and mutual influences. This extension could directly impact the explanations about some ideas that were not explored until now, such as habits, motivation and body and mind's health (all together). Finally, several other experiments and their combinations could identify opportunities to extend the explanation algorithms.

Extensions of this study intend to apply this research method in other health domains to verify the generality of the approach and investigate ways to link the concepts of the *mHealth* domain with traditional health ontologies. Thus, the background knowledge could be augmented, generating explanations with richer content. Moreover, it is intended to demonstrate the adequacy of the n-ary approach for temporal representation and their related explanations. Thus, complex time aspects (e.g., uncertain time and its relations), which are not naturally supported by ontologies, could be also part of the explanations.

REFERENCES

Pew Research Center. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally, Available in

https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing rapidly-around-the-world-but-not-always-equally (2019). Access at: December 15th, 2019.

Wac, K., Fiordelli, M., Gustarini, M. and Rivas, H., "Quality of Life Technologies: Experiences from the Field and Key Challenges," in IEEE Internet Computing, vol. 19, no. 4, pp. 28-35, July-Aug. 2015.

Wac, K., "From quantified self to quality of life." In Digital Health, pp. 83-108. Springer, Cham, 2018.

G. Chiarini, P. Ray, S. Akter, C. Masella and A. Ganz, "mHealth Technologies for Chronic Diseases and Elders: A Systematic Review," in IEEE Journal on Selected Areas in Communications, vol. 31, no. 9, pp. 6-18, September 2013.

Ciman, Matteo, and Wac, K., "Individuals' stress assessment using human smartphone interaction analysis." in IEEE Transactions on Affective Computing 9, no. 1, pp. 51-65, 2016.

Lane, N. D., and Georgiev, P., "Can deep learning revolutionize mobile sensing?." in Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications, ACM, 2015.

Jiang, Wenchao, and Zhaozheng Yin, "Human activity recognition using wearable sensors by deep convolutional neural networks." in Proceedings of the 23rd ACM international conference on Multimedia, ACM, 2015.

Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., and Kawsar, F., "An early resource characterization of deep learning on wearables, smartphones and internet-of-things devices." In Proceedings of the 2015 international workshop on internet of things towards applications, pp. 7-12, 2015.

Miller, W. C., Koceja, D. M. and Hamilton, E. J., "A meta-analysis of the past 25 years of weight loss research using diet, exercise or diet plus exercise intervention." in International journal of obesity, vol. 21, no. 10, pp. 941-947, 1997.

Caudwell, P., Hopkins, M., King, N. A., Stubbs, R. J., & Blundell, J. E, "Exercise alone is not enough: weight loss also needs a healthy (Mediterranean) diet?" in Public health nutrition, vol. 12, no. 9A, pp. 1663-1666, 2009.

Snae, Chakkrit, and Brüeckner, M., "Ontology-driven e-learning system based on roles and activities for Thai learning environment." in Interdisciplinary Journal of E Learning and Learning Objects 3, no. 1, pp. 1-17, 2007.

Sungmee, P. and Jayaraman, S., "Enhancing the quality of life through wearable technology," in IEEE Engineering in Medicine and Biology Magazine, vol. 22, no. 3, pp. 41-48, May-June 2003.

Gruber, T. R., "A translation approach to portable ontology specifications". Knowledge acquisition, vol. 5, no. 2, pp. 199-221, 1993.

Noy, N. F. and McGuinness, D. L., “Ontology development 101: A guide to creating your first ontology”, 2001.

Automotive Ontology Community Group Wiki. The World Wide Web Consortium (W3C), 2016. Available in https://www.w3.org/community/gao/wiki/Main_Page. Access at: July 13th, 2020.

WHOQOL: Measuring Quality of Life. World Health Organization, 2020. Available in www.who.int/healthinfo/survey/whoqol-qualityoflife/en/index1.html. Access at: June 27th, 2020.

Peart, Dan & Balsalobre-Fernández, Carlos & Shaw, Matthew. “The Use of Mobile Applications to Collect Data in Sport, Health and Exercise Science: A Narrative Review”, Journal of Strength and Conditioning Research, vol. 33, no. 1, 2017.

Fallaize, R. et al, “Popular Nutrition-Related Mobile Apps: An Agreement Assessment Against a UK Reference Method.”, JMIR mHealth and uHealth vol. 7,2 e9838, 20 Feb, 2019.

Chen, Yuqing, and Yang Xue, "A deep learning approach to human activity recognition based on single accelerometer.", 2015 IEEE International Conference on Systems, Man, and Cybernetics. IEEE, 2015.

Michalski, R. S., "A theory and methodology of inductive learning.", Machine learning, Springer, Berlin, Heidelberg, pp. 83-134, 1983.

Mantaras, R., & Armengol, E., “Machine learning from examples: Inductive and Lazy methods”, Data & Knowledge Engineering, vol. 25, no. 1-2, pp. 99–123, 1998.

Bhavsar, H., & Ganatra, A., “A comparative study of training algorithms for supervised machine learning”, International Journal of Soft Computing and Engineering (IJSCE), 2(4), pp. 2231-2307, 2012.

Mohammad, A. L., Chaoji, V., Salem, S., and Zaki, M., "Link prediction using supervised learning.", In SDM06: workshop on link analysis, counter-terrorism and security, vol. 30, pp. 798-805, 2006.

Sathya, R., & Abraham, A., “Comparison of supervised and unsupervised learning algorithms for pattern classification”, in International Journal of Advanced Research in Artificial Intelligence, vol. 2, no. 2, pp. 34-38, 2013.

Manea, V. and Wac, K., “mQoL: mobile quality of life lab: from behavior change to QoL”, In Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, pp. 642-647, 2018.

H2020 NESTORE project website, Available in https://nestore-coach.eu/home Project ID 769643, page visited on 20th Aug 2019.

Sivilai, S., Chakkrit S., and Brueckner, M., "Ontology-driven personalized food and nutrition planning system for the elderly.", In The 2nd international conference in business management and information sciences, 2012.

Pramono, D., Nanang, Y. S., Riyanarto, S., and Mohamad, S., "Physical activity recommendation for diabetic patients based on ontology.", In 7th International Conference on Information & Communication Technology and Systems, pp. 27-32, 2013.

Nicoletti, M. C., “Ampliando Os Limites Do Aprendizado Indutivo De Máquina Através Das Abordagens Construtiva E Relacional”, pp. 22-32, 1994.

Hornik, K., Buchta, C., & Zeileis, A., “*Open-source machine learning: R meets Weka”, Computational Statistics, vol. 24, no.2, pp. 225–232*, 2008.

Wang, S. C., “*Artificial Neural Network. Interdisciplinary Computing in Java Programming”, pp. 81–100,* 2003.

Barreto, L., Cavaco, I., Monteiro, A., Rousy, D., "Gamification Aspects in Detail: Collectanea of Studies to Renew Traditional Education." Revista Eletrônica Argentina Brasil de Tecnologias da Informação e da Comunicação 1.4 (2016).

Kulkarni, P., Joshi, P., “ARTIFICIAL INTELLIGENCE: Building Intelligent Systems”, PHI Learning, pp. 9, 2015.

Montavon, G. et al., “Explaining nonlinear classification decisions with deep Taylor decomposition”, Pattern Recognition, vol. 65, pp. 211–222, 2017.

Manea, V., et al. “LDC ’19: international workshop on longitudinal data collection in human subject studies”, In Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, Association for Computing Machinery, pp. 878–881, 2019.

L. Bühmann, J. Lehmann, P. Westphal, “DL-Learner-A framework for inductive learning on the Semantic Web”, Journal of Web Semantics , vol. 39, pp. 15- 24, 2016.

Maxat, K. et al. “Semantic similarity and machine learning with ontologies”, Briefings in Bioinformatics, 2020.

I. Sim, “Mobile devices and health,” in New England Journal of Medicine, vol. 381, no. 10, pp. 956-968, 2019.

M. B. D. Rosario et al., "Evaluation of an mHealth-Based Adjunct to Outpatient Cardiac Rehabilitation," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 6, pp. 1938-1948, Nov. 2018.

C. Hoyos-Barceló, J. Monge-Álvarez, M. Zeeshan Shakir, J. AlcarazCalero and P. Casaseca-de-la-Higuera, "Efficient k-NN Implementation for Real-Time Detection of Cough Events in Smartphones," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 5, pp. 1662-1671, Sept. 2018.

R. S. Istepanian, T. Al-Anzi, “m-Health 2.0: new perspectives on mobile health, machine learning and big data analytics,” in Methods, vol. 151, pp. 34-40, 2018.

D. A. Melis, T. Jaakkola, "Towards robust interpretability with self explaining neural networks," Advances in Neural Information Processing Systems, 2018, pp. 7775-7784.

R. C. Fong, A. Vedaldi, "Interpretable Explanations of Black Boxes by Meaningful Perturbation," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017, pp. 3429-3 7.

J. Lehmann, S. Bader, P. Hitzler, “Extracting reduced logic programs from artificial neural networks,” Applied Intelligence, vol. 32, no. 3, pp. 249-266, 2010.

R. Cornet, N. Keizer, "Forty years of SNOMED: a literature review," BMC Medical Informatics and Decision Making, vol. 8, no. 1, pp. 1-6, 2008.

A. Mastropietro, C. Roecke, S. Porcelli, J. del Bas, N. Boquè, L. Maldonado, G. Rizzo, “Multi-domain Model of Healthy Ageing: The Experience of the H2020 NESTORE Project,” In Italian Forum of Ambient Assisted Living, 2018, pp. 13-21.

J. Sáez, B. Krawczyk, M. Woźniak, “On the Influence of Class Noise in Medical Data Classification: Treatment Using Noise Filtering Methods," Applied Artificial Intelligence, vol. 30, no. 6, pp. 590-609, 2016.

APPENDIX

Table 5: Physical Health application group.

IOS	Android
Nike Training Club - 5 starts	Stronglifts 5x5 - Weight Lifting & Gym Workout Log - 4.9 stars
8fit Workouts & Meal Planner - 4.7 stars	Step Counter - 4.7 stars
Seven, 7 Minute Workout - 4.8 stars	Legs & Butt Workout - 4.6 stars
Freeletics - Workout & Fitness - 4.6 stars	30 Day Fitness Challenge - 4.8 stars
Sworkit Fitness & Workout - 4.7 stars	Daily Workouts - 4.7 stars

Table 6: Health care application group.

IOS	Android
Sleepzy - 4.5 stars	Blood Pressure - 4.2 stars
Kardia - 4.9 stars	PsicoTests - 4.6 stars
Dental Drugs - 4.8 stars	Headache Log - 4.6 stars
Pregnancy + - 4.8 stars	Period Tracker - 4.9 stars
BabySparks - 4.7 stars	My Pregnancy - 4.4 stars

Table 7: Spiritual application group.

IOS	Android
Zen - 4.9 stars	Headspace: Meditation & Sleep - 4.3 stars
Calm - 4.7 stars	Law of Attraction Space - 4.7 stars
Colorfy - 4.6 stars	5’ Minutes, I meditate - 4.4 stars
Daily Spiritual Quotes - 5 stars	Spiritual Me: Masters Edition - 4.6 stars
Headspace: Meditation & Sleep - 4.9 stars	Spiritual Transformation Daily - 4.5 stars

Table 8: Educational application group.

IOS	Android
Duolingo - 4.8 stars	Daily Art - 4.8 stars

Linguee - 4.9 stars	Neuronational Memory Trainer - 4.4 stars
HandTalk - 4.3 stars	GeoExpert - 4.4 stars
HOMER Reading: Learn to Read - 4.5 stars	Khan Academy - 4.5 stars
Kids Academy Talented & Gifted - 4.5 stars	Udemy - 4.6 stars

Table 9: Diet application group.

IOS	Android
MyFitnessPal - 5 stars	Diet diary - 4.6 stars
Lifesum - 4.8 stars	Dietbox - 4.7 stars
WebDiet - 5 stars	Health Diet Foods Fitness Help - 4.6 stars
Diet & Training by Ann - 4.9 stars	Calorie Counter - 4.5 stars
Progress Body Tracker & Health - 4.6 stars	Diet and Health - Lose Weight - 4.4 stars

Table 10: Entertainment application group.

IOS	Android
Spotify - 4.8 stars	Prime Video - 4.4 stars
Youtube - 4.7 stars	Cine Plus - 4.8 stars
Shazam - 4.9 stars	9gag Funny Gifs - 4.6 stars
Candy Crush Saga - 4.8 stars	Netflix - 4.5 stars
Traffic Rider Moto Game - 4.6 stars	Tiktok - 4.5 stars

Table 11: Social application group.

IOS	Android
Instagram - 4.8 stars	Tumblr - 4.6 stars
Pinterest - 4.8 stars	Telegram - 4.4 stars
Twitter - 4.8 stars	Linkedin - 4.2 stars
Whatsapp - 4.7 stars	Tinder - 4 stars
Snapchat - 4 stars	VSCO - 4.4 stars

Table 12: Business application group.

IOS	Android

Whatsapp Business - 4.8 stars	Tiny Scanner - 4.8 stars
Google My Business - 4.7 stars	Business Card Maker - 4.5 stars
The Economic Times - 4.5 stars	Box - 4.7 stars
Logo Maker Shop - 4.7 stars	Square Point of Sales - 4.5 stars
Mint Personal Finance & Money - 4.7 stars	Entrepreneur Business Ideas - Tools & Tutorials - 4.5 stars

Table 13: Recreation application group.

IOS	Android
TripAdvisor - 4.8 stars	Booking - 4.8 stars
Airbnb - 4.8 stars	Hoteis.com - 4.6 stars
App in the Air - 4.6 stars	KAYAK - 4.5 stars
Trivago - 4.8 stars	Packpoint List - 4.7 stars
Regal: Movie Tickets & Times - 4.7 stars	BlaBlaCar - 4.6 stars

Table 14: Productivity application group.

IOS	Android
Trello - 4.8 stars	Word Office - 4.5 stars
Google Drive - 4.7 stars	Loop Habit Tracker - 4.7 stars
Focus - Time Management - 4.6 stars	Social Media Post Maker: Planner & Graphic Design - 4.6 stars
Google Calendar - 4.6 stars	Forest: Stay Focused - 4.5 stars
Ebook Downloader - 4.6 stars	Dreamie Planner - 4.4 stars

Table 15: Emotional application group.

IOS	Android
Zenklub - 4.8 stars	Youper emotional health - 4.7 stars
Ponder - 5 stars	Diccionario Bio-Emocional - 4.6 stars
Breath2Relax - 5 stars	Emotional Intelligence - Best Education App - 4.5 stars
Evolue App - 5 stars	Mitra - 4.1 stars

Cíngulo - 4.9 stars	Discovering Emotions with Zeely - 4 stars

Table 16: Tools application group.

IOS	Android
Adobe Acrobat - 4.7 stars	ColorNote - 4.8 stars
Outlook - 4.7 stars	Dropbox - 4.3 stars
Cam Scanner - 4.9 stars	AppBlock - 4.4 stars
Microsoft Word - 4.8 stars	Night Clock - 4.2 stars
C6 Bank - 4.6 stars	Mobi Calculator - 4.8 stars

Figure 49: Ontology - Part 1, rotated to the left.

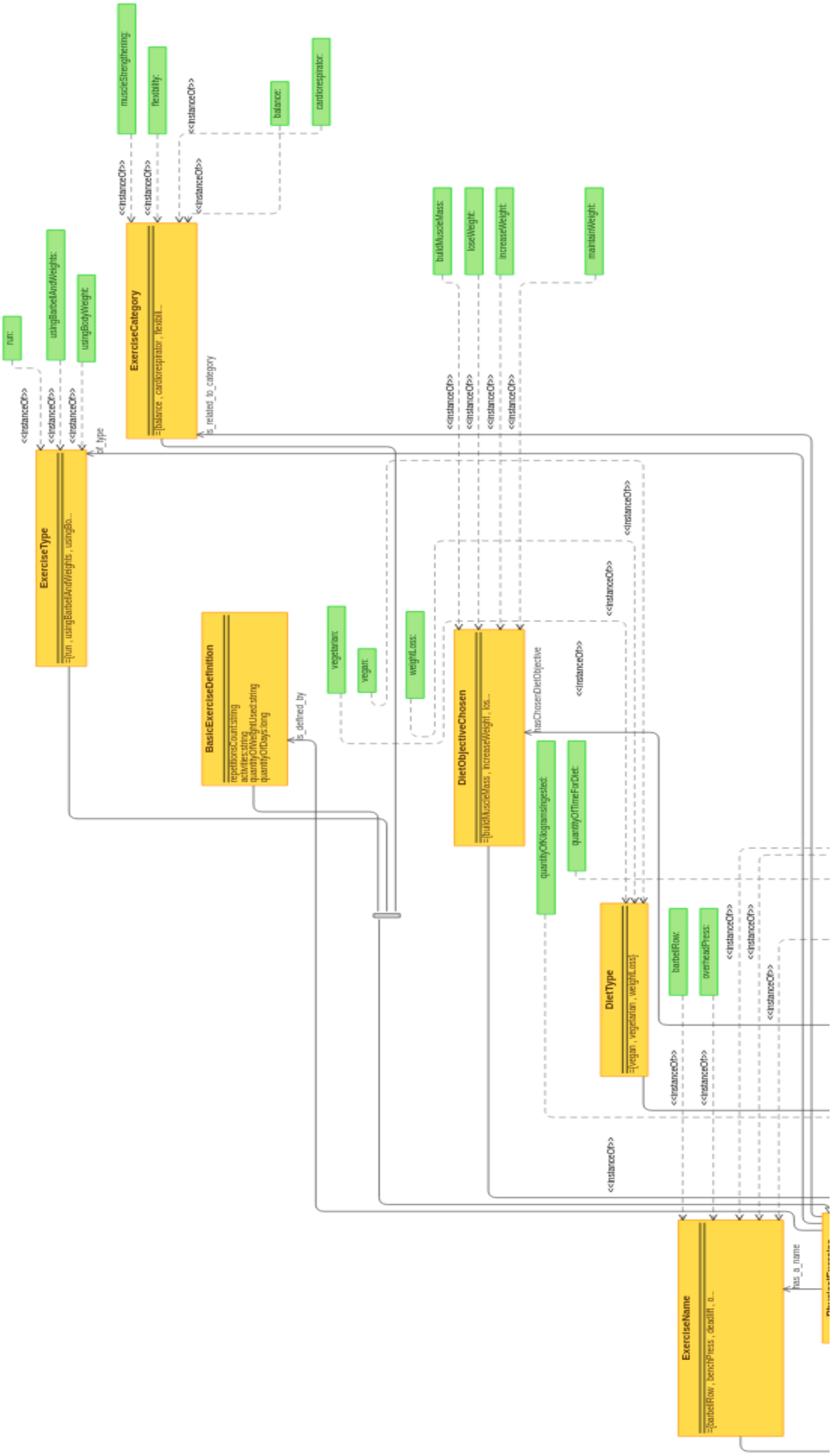


Figure 50: Ontology - Part 2, rotated to the left.

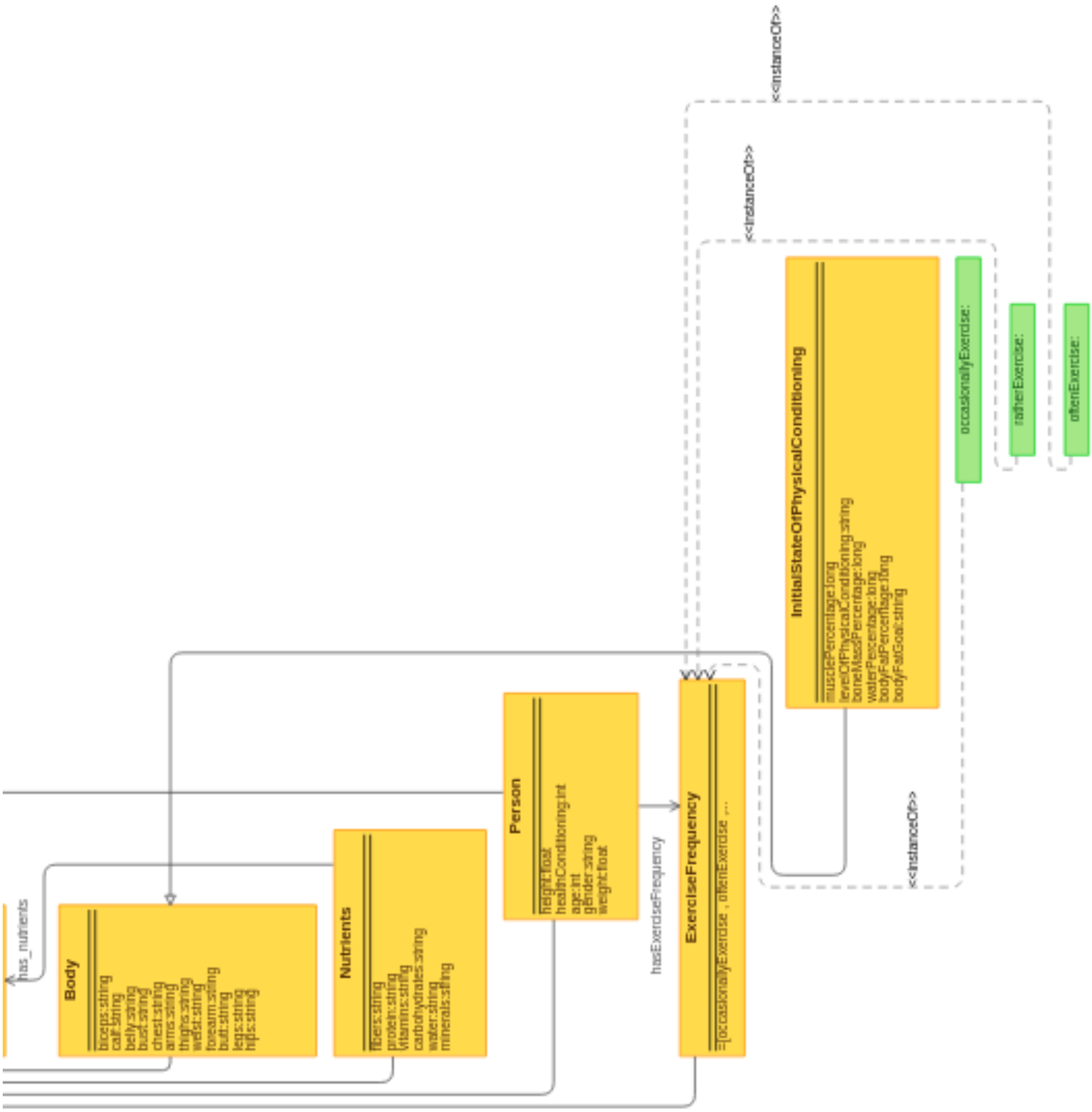


Figure 51: Ontology - Part 3, rotated to the left.