CBIS'24
XX Congresso Brasileiro de Informática em Saúde
08/10 a 11/10 de 2024 - Belo Horizonte/MG - Brasil

Explicabilidade baseada em conhecimento temporal: um estudo de casos em mHealth

Temporal knowledge-based explanations for inductive reasoning: a mHealth case example

Explicaciones temporales basadas en el conocimiento para el razonamiento inductivo: un ejemplo de caso de mHealth

Isabela Nascimento[1], Clauirton Siebra[2]

1 Informatics Center, Federal University of Paraiba, Joao Pessoa (PB), Brazil.
2 LIAA, Federal University of Paraiba, Joao Pessoa (PB), Brazil.

Autor correspondente: Dr. Clauirton de Albuquerque Siebra
E-mail: clauirton@ci.ufpb.br

Resumo

Objetivo: Investigar a geração de explicações para sistemas indutivos utilizando uma ontologia unificada que representa o estado de saúde de usuários móveis. Esta ontologia serve como conhecimento a priori, facilitando a geração de explicações. *Método*: Examinamos 24 aplicativos móveis de saúde (mHealth) para desenvolvimento da ontologia, enfatizando extensões que consideram aspectos temporais. Tais aspectos costumam ser negligenciados nas representações de saúde, dada a limitação das ontologias em modelar relações temporais ternárias. Em seguida, aplicamos diferentes configurações de um algoritmo indutivo que recebe esta ontologia como entrada, gerando explicações para seus resultados indutivos. *Resultados*: Experimentos mostram que a estrutura do modelo temporal afeta a legibilidade das explicações. Além disso, os experimentos enfatizam o tradeoff entre precisão e poder de generalização. Conclusão: As extensões temporais melhoram a expressividade das explicações, uma vez que as relações e conceitos temporais são explorados para melhor contextualizar fatos temporais associados a resultados indutivos.

Descritores: mHealth; Explicabilidade; Representação do Conhecimento

Abstract

Objective: Investigate the generation of explanations for inductive systems using a unified ontology that represents the health status of mobile users. This ontology serves as a priori knowledge, facilitating the generation of explanations. *Method*: We examined 24 Mobile health (mHealth) apps to develop this ontology, emphasizing extensions that consider temporal aspects. Such aspects are usually neglected in health representations, given the limitation of ontologies in modelling ternary temporal relations. After that, we applied different configurations of an inductive algorithm that receives this ontology as input, generating explanations for their inductive outcomes. *Results*: Experiments show that the temporal model structure affects the readability of explanations. Moreover, experiments emphasize the tradeoff between accuracy and generalization power. *Conclusion*: Temporal extensions improve the expressiveness of explanations since temporal relations and concepts are explored to better contextualize temporal-based facts associated with inductive outcomes.

Keywords: mHealth; Explicability; Knowledge Representation

Resumen

Objetivo: Investigar la generación de explicaciones para sistemas inductivos utilizando una ontología unificada que represente el estado de salud de los usuarios móviles. Esta ontología sirve como conocimiento a priori, facilitando la generación de explicaciones. *Método*: Examinamos 24 aplicaciones de salud móvil (mHealth) para desarrollar esta ontología, enfatizando extensiones que consideran aspectos temporales, dada la limitación de las ontologías a la hora de modelar relaciones temporales ternarias. Posteriormente, aplicamos diferentes configuraciones de un algoritmo inductivo que recibe esta ontología como entrada, generando explicaciones para sus resultados inductivos. *Resultados*: Los experimentos muestran que la estructura del modelo temporal afecta la legibilidad de las explicaciones. Además, los experimentos enfatizan el tradeoff entre precisión y poder de generalización. *Conclusión*: Las extensiones temporales mejoran la expresividad de las explicaciones, ya que se exploran las relaciones y conceptos temporales para contextualizar mejor los hechos de base temporal asociados con resultados inductivos.

Descriptores: mHealth; Explicabilidad; Representación del Conocimiento

Introdução

The use of mobile health (mHealth) applications has grown in the last decade, and several studies are presenting proposals to use mobile devices to support continuous and real-time monitoring of their users. These proposals range, for example, from cardiac rehabilitation support programs to the real-time detection of cough events. Indeed, mobile technology is very useful for this kind of domain because it can generate adequate and customized assessments and interventions.

Following this process, the health assessment task usually generates a significant amount of data. Thus, studies in health typically rely on machine learning (ML) techniques to obtain insides into the underlying problems. However, the main ML techniques, such as neural networks and support vector machines (SVM), execute as black-boxes since they do not directly indicate why their internal models generated a given output. Neural networks, for example, rely on the distributed nature of the information encoded in the set of their 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, the healthcare personnel must comprehend and explain how such algorithms delivered their classifications or predictions.

Considering such a limitation, studies are investigating how interpretable explanations can be automatically generated for black-box ML methods. For example, the study of Fong and Vedalvi[(1)] proposed an agnostic framework that returns explanations for any black-box strategy. This framework relies on interpretable rules that describe the relationships between input features and inductive conclusions. Similarly, other studies, such as described in the survey of Garcez et al.[(2)], are also based on rule extraction, which employs propositional sentences as target formalism for creating explanations. As this logical family is limited in expressiveness, the outcomes also remain restricted in terms of explanations that can be generated.

An alternative approach to generating explanations involves employing background knowledge, which defines concepts and their relationships in structured data. Ontologies are the primary method for representing such knowledge, particularly in the health domain[3,14], due to their benefits, including consistency, reusability, and easy expandability. However, ontologies also have limitations. In the health domain, temporal aspects are often crucial, yet modeling evolving information over time in ontologies poses a challenge. This complexity arises since ontologies support classes and properties, which are respectively unary and binary predicates formalized according to Description Logics[4].

The approach presented in this paper focused on overcoming the current limitations of ontologies regarding their temporal support for explanations. Therefore, we created a background knowledge based mainly on the information extracted from 24 Android and iPhone mHealth applications publicly available in app stores, as discussed later in this paper. The definition of this background knowledge involved experiments with extensions of temporal models, such as 4D-fluent[5] and Onto-mQoL[6], to verify the most appropriate support regarding explanation accuracy and readability. Therefore, our contributions are: (1) Design of holistic background knowledge (i.e., ontologies) for mHealth applications; (2) Demonstration of the adequacy of temporal representations concerning the generation of explanations for inductive reasoning processes; and (3) Discussion on the quality of the explanations, considering strategies (e.g., use of anonymous classes, which are classes defined by means of restrictions) to arrange the background knowledge.

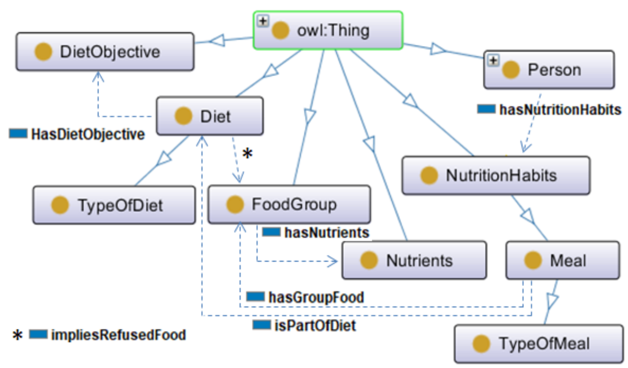
Methods

Background Knowledge

Our background knowledge was developed as a case example of ontology that integrates information about diverse health dimensions (e.g., Physical/Physiological) and enables the specification of more expressive reasoning processes to augment the accuracy and reliability of health assessments and interventions. Therefore, we relied on the three dimensions (Emotional, Physical/Physiological, and Nutritional) identified by the NESTORE European Commission program[3] as the most important for maintaining a healthy life. As our project focused on mHealth applications, we modelled these domains from the perspective of health-related apps available in the app stores. In other words, the study of these applications provided the terms (classes) and relations (properties) to the creation of our background knowledge (ontology). This perspective is an important premise of our research since the instances that compose the ontology are derived from data assessed using wearable/mobile devices.

While app stores contain hundreds of applications for each domain, we restricted our knowledge modeling to the eight applications of each domain (resulting in 24 apps) with the highest user evaluations. We only considered eight apps because, after their analysis, no new significant information was found using other apps. Only freely available apps with more than 1000 ratings were considered. The following ontology fragment (Figure 1) illustrates an example of modelling related to the nutritional dimension.

Figura 1 – Part of the ontology that represents nutritional habits of mobile users.



This part of the ontology was based on the data assessed from eight Android and iOS mobile applications related to the field of nutrition. These applications were selected based on users' evaluations (scale from 1 to 5) in their app stores and they have in common, for example, the concepts of *Diet* and *Meal*. We also used studies from academia (e.g., NESTORE [(3)] and e-NUTRI[(7)]) to adjust some ontology concepts and properties. The same idea was used for the other two dimensions (physical health and emotional state). In other words, the main available applications were analyzed to decide the main concepts that should be included in the ontology.

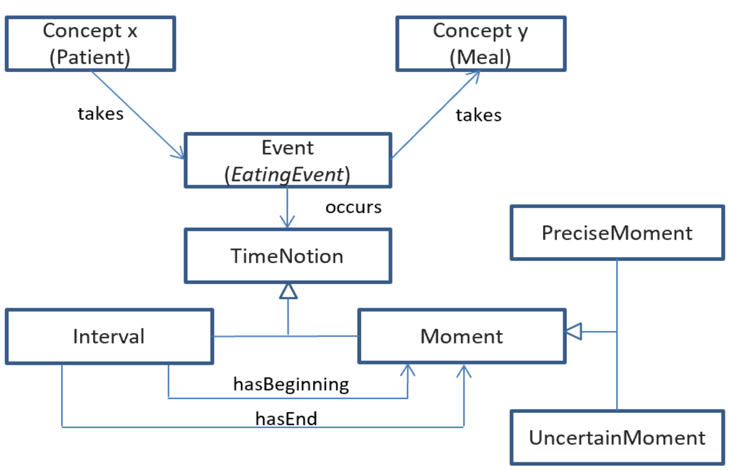
Other schemas of our representations demonstrate the inter-relation between domains. For example, a person's emotional state (emotional domain) has a motivation associated, while the activity behavior (Physical/Physiological domain) has an associated average frequency. Since the literature presents studies that associate the influence of motivation with the frequency of physical activities[8], inductive reasoning could infer this influence through a rule indicating that instances of the class *Person*, who have low motivation, tend to have an insufficient frequency regarding their activity behavior.

Ontology for Temporal Descriptions: Onto-mQoL

Onto-mQoL[6] is a proposal for ontology temporal extensions based on the n-ary relations strategy[10], which is a simplification of 4D-Fluent since it only requires a new object and two additional properties to represent 3-ary predicates. Onto-mQoL introduces some advantages for time modelling such as support for uncertainty, different time relations that involve interval and moment time notions, compatibility regarding current ontology standards, and associated knowledge method. Onto-mQoL relies on a n-ary structure to configure temporal extensions as illustrated in Figure 2.

According to the Onto-mQoL proposal, n-ary properties are represented as classes (e.g., *EatingEvent*) rather than properties, while *instances* of these classes correspond to relation instances.

Figura 2 – Onto-mQoL approach for temporal knowledge representations.



For example, the property *takes* was initially relating a subject (instance of the *Patient* class) with an object (instance of the *Meal* class). After the modification using the temporal approach (Figure 2), both concepts (*Patient* and *Meal*) are now associated with *EatingEvent*. These and other possible elements that represent events (e.g., *HydrationEvent*) are derived from the *Event* class. This class includes the *occurs* property that links such events to their notion of time (interval or moment) as indicated in Figure 2. Intervals are defined using initial and final moments and. Consequently, have the notion of duration. In contrast, moments are singular points in a time scale. For instance, the class *EatingEvent* is associated with an interval denoted as *occurs*(*EatingEvent*, *Interval*-X). This interval is specified by two moments using the properties *hasBeginning* and *hasEnd*. Moreover, any moment, and consequently intervals, can be represented as uncertain[6].

Experiments and Results

Dataset

This present study aimed at evaluating forms of background knowledge that consider temporal aspects and can support the generation of useful explanations to the inferred knowledge. The literature presents a lack of multifeatured and longitudinal studies in mHealth[11], so a proper dataset for experimentations is still unavailable. The study of the QoL Lab[12], for example, is one of the few ongoing and important efforts in such a direction. Therefore, we demonstrate this specification process from its practical perspective, using a dataset adapted from the *Heart Disease Dataset*[13]. This dataset has 303 instances (i.e., each corresponds to an individual) with 14 attributes (13 inputs and 1 outcome). We used the part of the mHealth ontology related to these attributes.

Experiment Description

Three experiments were conducted in this study. Firstly, we generated explanations for an inductive algorithm using the original static ontology as background knowledge. This means it does not present temporal concepts or relations. The aim of this experiment was to compare the accuracy of the algorithm against the baseline implementation and verify the explanations outcomes.

Secondly, we modified the ontology, including some temporal aspects. This modification was conducted using the 4D-Fluent and Onto-mQoL approaches. We did not consider other approaches such as Temporal DLs since they require extensions in the tools used in the experiments (e.g., Protégé). The aim of this experiment was to verify if the structure of temporal extensions affects the explanations accuracy and readability.

Finally, explanations were evaluated using qualitative (readability) and quantitative (accuracy and F-measure) metrics. We used the DL-Learner framework to generate explanations, configured with the the *Class Expression Learning for Ontology Engineering* (CELOE)[9] algorithm. Its fundamental concept is to construct a search tree using a refinement operator, employing a heuristic to identify promising candidates for exploration. The algorithm's default execution time was progressively extended in pursuit of improved results, continuing until reaching a saturation point where further increases no longer yield significant gains.

Results – Static Domain

We used the Multilayer Perceptron algorithm (neural network) as the inductive reasoner to classify individuals as healthy or unhealthy. The accuracy provided by this model (80.86%) gives a basis for analyzing explanations concerning their accuracy. The next step was to create explanations indicating why individuals were classified as unhealthy. Therefore, the ontology and instances were modeled in the Protégé tool, creating the Tbox (terminological component) and Abox (assertion components) of the knowledge base.

Table 2 shows the results of six different configurations (Ci) for the experiments. The *Anonymous classes* attribute indicates if such a type of class was used in the ontology, the *Noise* attribute indicates the approximated percentage of data that can be considered noise, and the *saturation time* indicates the moment where improvements were no longer significative. Table 1 also shows that the use of the dataset with 0% noise (C1 and C2) generates explanations with low accuracy. However, it is widely acknowledged that medical datasets frequently suffer from noise and incompleteness, stemming from challenges in data collection and integration[28]. Therefore, we experimented with allowing 10% and 20% noise within the dataset. While this approach did enhance accuracy, the achieved accuracy still fell short of the baseline result (80.86%).

Table 1 – Quantitative Results for different configurations

Ci	*Anonymous classes*	*Noise*	*Highest accuracy*	*Highest F-measure*	*Satura-tion time*
C1	Yes	0%	49.17%	64.35%	00m21s
C2	No	0%	49.50%	64.50%	01m31s
C3	Yes	10%	63.35%	70.91%	00m57s
C4	No	10%	71.29%	74.78%	01m46s
C5	Yes	20%	76.57%	76.57%	01m03s
C6	No	20%	73.27%	74.92%	02m26s

Regarding the anonymous classes, we could observe differences between C1/C2, C3/C4 and C5/C6 in the metrics, such as the shorter time to obtain the explanations when the anonymous classes are used. Differently, we cannot affirm that the use of such a type of class improves the accuracy of the explanations. We also expected that using anonymous classes could enrich the background knowledge, creating more possibilities for the specification of rules that cover the results of inductive reasoning. However, this phenomenon was not consistently observed. Apart from C1 and C2, which had low accuracy, the other generated explanations can be read as:

C3: *Male or (hasPhysicalHealth only (PreDiabetic and (not (IdealBloodPressure))))*

This means 63.35% of unhealthy persons are male or persons that are pre-diabetic and do not have ideal blood pressure.

C4: *(hasPhysicalHealth only (meanBloodGlucose some decimal[>= 4.5])) or (age some decimal[>= 55.5])*

This means 71.29% of unhealthy persons have blood glucose average level higher than 4.5 and age higher than 55.5 years.

C5: *(Male and Older) or (hasEmotionalState only HighStressLevel)*

This means 76.57% of unhealthy persons are male and older, or persons associated with a high level of stress.

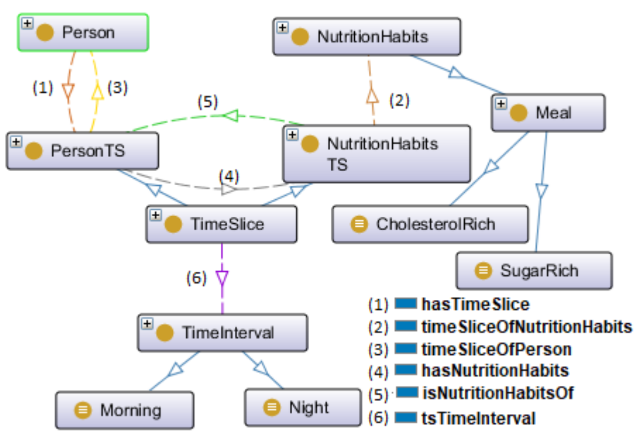
C6: *(hasPhysicalHealth only (meanBloodGlucose some decimal[>= 4.5])) or (age some decimal[>= 59.5])*

This means 73.27% of unhealthy people have an average level of glucose in the blood higher than 4.5 and age higher than 59.5 years. These explanations show that the use of anonymous classes (C3 and C5) improves the readability of the sentences since the semantics of the values are used over the inductive reasoning rather than only qualitative values. Moreover, in many cases, the qualification of the value (e.g., young, adult, or older) is more important than a simple value such as 55.5.

Results – 4D-Fluent Approach

The ontology in Figure 1 was modified according to the 4D-Fluent approach, so we could evaluate the explanations when temporal aspects are part of the model. First, we have identified an object property of the ontology that requires this temporal aspect. This property was *hasNutritionHabits* (and its inverse property *isNutritionHabitsOf*) since a person can have different nutrition habits over his/her life. Then, the following model was created to this part of the ontology (Figure 3) according to the 4D-Fluent approach (the anonymous classes *Morning* and *Night* were created just to simplify the experiment). However, the experiments with this representation did not return useful recommendations and highlighted a problem concerning their generation.

Figura 3 – The hasNutritionHabits as a 4D-Fluent element.



Consider, for example, a dataset where all people that eat meals rich in sugar and cholesterol at night are unhealthy. In such a case, a neural network generates a model with accuracy of 100%. Then, we could expect an explanation like:

Explanation expected. *hasTimeSlice some ((hasNutritionHabits some ((tsTimeInterval some Night) and (timeSliceOfNutritionHabits some (CholesterolRich and SugarRich)))) and (PersonTS and (tsTimeInterval some Night)))*

However, the resultant explanation was:

Explanation returned. *(hasTimeSlice some (tsTimeInterval only Night)) and (age some decimal[>= 40.5])*

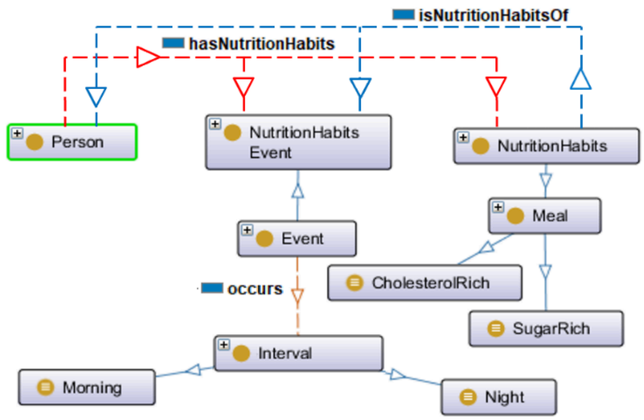
While the former explanation is complex and presents low readability, the former is simpler but does not bring useful information because it does not mainly include the *hasNutritionHabits* property. Thus, the time interval concept lacks semantics. This situation tends to be common because the class that distinguishes the *NutritionHabitsTS* instances, i.e. *TimeInterval*, as part of the explanation is the same for *PersonTS* (total correlation). Thus, independently of the time used to generate the explanations, the algorithm returns the same result.

Results: Onto-mQoL

Once more, the original ontology was modified, but this time according to the Onto-mQoL approach, also considering the *hasNutritionHabits* (and its inverse *isNutritionHabitsOf*) as the property converted to temporal (Figure. 4). This model shows the inclusion of the *NutritionHabitsEvent* class, which extends the *Event* class and has the object property *occurs* that accounts for relating an event to an *Interval* (subclass of *TimeNotion*). An important detail is the domain and range modifications of *hasNutritionHabits* and *isNutritionHabitsOf*, which are now represented as:

- hasNutritionHabits - Domain: *Person* OR *NutritionHabitsEvent*; Range: *NutritionHabitsEvent* OR *NutritionHabits;*
- isNutritionHabitsOf - Domain: *NutritionHabits* OR *NutritionHabitsEvent;* Range: *NutritionHabitsEvent* OR *Person*.

Figure 4. The hasNutritionHabits as an onto-mQoL element.



The same previous dataset was used with this modeling, so we could expect an explanation indicating that all people that have meals rich in sugar and cholesterol at night are unhealthy. The first experiments using the maximum execution time of 120 seconds (time used in the previous experiments with 4D-Fluent approach to obtain the learning convergence) returned the following explanation, which was obtained after 107 seconds:

Explanation (107s) *hasNutritionHabits only ((occurs some Night) and (hasNutritionHabits only SugarRich)) (pred. acc.: 99,34%, F-measure: 95,24%)*

This explanation had very high accuracy (99.34%) and readability. It is easy to observe that "unhealthy persons are those that only have, during the night, nutrition habits rich in sugar". However, we were expecting an explanation with 100% accuracy. As the *CholesterolRich* was not used, we decided to increase the maximum execution time (300s) to verify if the explanation algorithm could reach 100% accuracy. The algorithm obtained this result after 213 seconds:

Explanation (213s) *hasNutritionHabits only ((occurs some Night) and (hasNutritionHabits only (CholesterolRich and SugarRich))) (pred. acc.: 100,00%, F-measure: 100,00%)*

Besides the maximum accuracy, the explanation maintained very good readability since we could easily translate the sentence to "unhealthy persons are those that only have, during the night, nutrition habits rich in both cholesterol and sugar".

Discussion

The static scenario shows that the accuracy of the explainable algorithm is always lower than the accuracy of the neural network. This result still holds even when the *noise* hyperparameter increases. This fact corroborates the idea that the explanation accuracy is inversely proportional to its generalization power. This trading-off is discussed in papers such as the study of van der Veer et al.[15]. The experiments also showed that the temporal model structure affects the readability of explanations. While the 4D-Fluent approach could not generate useful explanations, the Onto-mQoL approach generated high-quality explanations regarding readability and accuracy. Some studies show that the n-ary approach, which composes the Onto-mQoL core, experiences redundancy in data storage when dealing with inverse and symmetric properties[5]. For example, the inverse of a relation is explicitly added twice instead of once as in 4D-fluents. Indeed, redundancy is a common problem of all approaches that try extending static models to temporal versions. The use of anonymous classes is another structural feature that affects the readability and accuracy of explanations. The suggestion is to create anonymous classes that represent important concepts of the domain. For example, *SugarRic*h and *CholesterolRich* are concepts that aggregate value to our explanations. However, several other classes could be created to cluster domain instances and emphasize their features.

Algorithms to generate explanations are also based on an inductive reasoning system inspired by inductive logic programming. Thus, their execution time is another factor that affects the quality of the explanations. Our experiments, for example, showed that we obtained explanations with more details and accuracy when we increased the maximum generation time from 120 to 300 seconds.

The experiments also demonstrated that the algorithms could generate results that only present restricted explanations. For example, one of the explanations generated was: *Male or (hasPhysicalHealth only (PreDiabetic and (not (IdealBloodPressure))))*. As blood pressure has four disjoint classes (low, ideal, high, and very high), this explanation indicates that persons can be a member of any other class apart from *IdealBloodPressure*, rather than an exact class. Moreover, the algorithms also present a high number of configurations (CELOE has about 30 parameters), so their efficiency may also depend on tuning the values of such parameters.

Conclusions

This study's primary contributions include: (1) proposing a strategy to develop background knowledge for mHealth applications, enabling comprehensive reasoning and explanations for results obtained through inductive reasoning; (2) demonstrating the suitability of the Onto-mQoL approach for temporal representation and related explanations; and (3) discussing crucial elements that can impact the readability and accuracy of explainability, such as the utilization of anonymous classes and the configuration of explanation algorithms. Future extensions of this work aim to apply the research method in other health domains to assess the approach's generality and explore methods to connect mHealth domain concepts with traditional health ontologies. This expansion would enrich the background knowledge and facilitate the generation of more detailed explanations.

The main limitation of this project is the lack of longitudinal and multidimensional data that could be used in the learning process. This lack avoids, for example, the development of models that can classify or predict specific health conditions. Despite the potential benefits of mHealth applications, their application has been limited, as most users stop employing them just after a few times of use. Under these circumstances, achieving data with quality from mHealth apps is less likely.

Referências

1. Fong RC, Vedaldi A. Interpretable Explanations of Black Boxes by Meaningful Perturbation. Proceedings of the IEEE International Conference on Computer Vision, 2017, 3429-3437.

2. Garcez ADA, et al. Neural-symbolic learning and reasoning: A survey and interpretation. Neuro-Symbolic Artificial Intelligence: The State of the Art, 2022, 342, 1-51.

3. Mastropietro A, et al. Multi-domain Model of Healthy Ageing: The Experience of the H2020 NESTORE Project. Italian Forum of Ambient Assisted Living, 2018, 13-21

4. Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider PF. The Description Logic Handbook: Theory, Implementation, and Applications. 2010, Cambridge University Press.

5. Batsakis S, Petrakis E, Tachmazidis I, Antoniou G. Temporal representation and reasoning in OWL 2. Semantic Web, 2017, 8(6): 981–1000.

6. Siebra C, Wac K. Engineering uncertain time for its practical integration in ontologies. Knowledge-based Systems. 2022, 251, 109152.

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

8. Lewis M, Sutton A. Understanding Exercise Behaviour: Examining the Interaction of Exercise Motivation and Personality in Predicting Exercise Frequency. J. Sport Beh. 2011, 34(1): 82-97.

9. Procko T, Elvira T, Ochoa O, Del Rio N. An Exploration of Explainable Machine Learning Using Semantic Web Technology. IEEE 16th Int. Conf. on Semantic Computing, 2022, 143-146.

10. Giunti M, Sergioli G, Vivanet G, Pinna S. Representing n-ary relations in the Semantic Web. Logic Journal of the IGPL, 2019.

11. Manea V, Hansen MS, Elbeyi SE, Wac K. Towards Personalizing Participation in Health Studies. Fourth Int. Workshop on Multimedia for Personal Health & Health Care, 2019, 32-39.

12. Manea V, Wac K. MQOL: Mobile quality of life lab: From behavior change to QOL. Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 2018, 642-647.

13. Detrano R. et al. International application of a new probability algorithm for the diagnosis of coronary artery disease. American Journal of Cardiology. 1989. 64, 304-310.

14. Figueiredo EB et al.. Semântica em prontuários eletrônicos para oncologia pediátrica: uma revisão integrativa. Journal of Health Informatics. 2023, 15(2):61-9.

15. van der Veer SN, et al. Trading off accuracy and explainability in AI decision-making: findings from 2 citizens' juries. J. American Medical Informatics Association. 2021, 28(10), 2128-2138.