Computer Science and Information Systems 20(3):1109–1132 https://doi.org/10.2298/CSIS220829035F

# Probabilistic Reasoning for Diagnosis Prediction of Coronavirus Disease based on Probabilistic Ontology

Messaouda Fareh, Ishak Riali, Hafsa Kherbache, and Marwa Guemmouz

LRDSI Laboratory, Faculty of sciences
University Blida1
Soumaa, B.P 270, Blida, Algeria
{farehm, ishakriali, kherbachehaf, guemmouzm}@gmail.com

**Abstract.** The novel Coronavirus has been declared a pandemic by the World Health Organization (WHO). Predicting the diagnosis of COVID-19 is essential for disease cure and control. The paper's main aim is to predict the COVID-19 diagnosis using probabilistic ontologies to address the randomness and incompleteness of knowledge. Our approach begins with constructing the entities, attributes, and relationships of COVID-19 ontology, by extracting symptoms and risk factors. The probabilistic components of COVID-19 ontology are developed by creating a Multi-Entity Bayesian Network, then determining its components, with the different nodes, as probability distribution linked to various nodes. We use probabilistic inference for predicting COVID-19 diagnosis, using the Situation-Specific Bayesian Network (SSBN). To validate the solution, an experimental study is conducted on real cases, comparing the results of existing machine learning methods, our solution presents an encouraging result and, therefore enables fast medical assistance.

**Keywords:** COVID-19, Probabilistic Ontology, Multi-Entity Bayesian Networks, Uncertainty, Reasoning

## 1. Introduction

December 31, 2019, is the day of the appearance of COVID-19 for the first time in the Wuhan region, China notified the outbreak to the World Health Organization [49]. Since 2019, the world has been affected by coronavirus-19 (COVID-19), which has caused many deaths. The world does not realize the importance of this disease and its impact on future life, until March 11, 2020, the COVID-19 epidemic is declared a pandemic by the World Health Organization (WHO) [36].

The COVID-19 virus is a highly contagious respiratory disease that has spread rapidly around the world since it was first reported in China in late December 2019. The early detection of patients at risk to develop critical illness may aid in delivering proper care and reduce mortality [46]. Current literature has already identified several risk factors.

Widely reported statistics on COVID-19 over the globe need to consider the uncertainty of the data and possible explanations for this uncertainty. The professional environment of medical practice is characterized by great complexity, many grey areas, and uncertainty as an essential component of all levels of patient care.

The term ”uncertainty” is intended to encompass a variety of aspects of imperfect knowledge, including incompleteness, vagueness, ambiguity, and others[1]. Medical infor-

---

[1] `https://www.w3.org/2005/Incubator/urw3/group/draftReport.html`

mation can be imprecise, partial, vague, or imperfect as some lab results may be missing from the feature set or due to the patient's inability to answer properly [3].

We present an example of uncertainty in medical applications cited in [41]. In a simple scenario, we could try to model the knowledge that patients showing a running nose and complaining of feeling light-headed have a common cold; we can use an axiom like *({RunNose, LightHead}, Cold)*. This rule would be correct most of the time but ignores the fact that there exist many different maladies—some of them potentially serious—that share these same symptoms and require additional interventions.

Uncertainty pervades every activity related to health care, it prompts patients to seek care and stimulates medical intervention. The inability to abolish uncertainty, furthermore, creates difficult challenges for clinicians and patients [24]. Therefore the diagnosis is particularly doubtfull, which can result in inaccurate or completely false diagnoses [23]. This is why the medical field gives great importance to the factor of uncertainty. Hence, it seems very interesting to design and implement automated healthcare systems that consider these challenges and ease the diagnosis task for the doctors [45].

Many research papers on COVID-19 used Machine Learning (ML) methods and ontologies as a knowledge base. Some contributions focus on applying ML in detecting the diagnosis of COVID-19, whereas others exploit data for predicting diagnosis.

Upon analyzing previous works, it has been observed that they use classical ontologies. ( [50], [37], [47], [25], [6], [29] and [32]), these studies cannot deal with uncertainty, and they only have precise concepts and relationships. A commonly mentioned limitation of classical ontology languages, especially within the context of knowledge representation, is their inability to model or handle uncertainty [41]. Indeed, notice that we use the axioms in an ontology as absolute information, and the consequences follow (or not) from these axioms. This leaves no space for statements which are not completely certain [41].

Despite all the advances made in the field of the semantic web, problems associated with data uncertainty and ambiguity still need to be solved in the knowledge management of a real domain. So uncertainty is inevitable when we model most application domains, like in the medical field, the symptoms are subjective and therefore imprecise and incomparable. In addition, concepts and relationships may not be described by the description logic language. One of the main flaws of classical ontology is the inability to represent and reason under uncertainty. This uncertainty can manifest during the prediction of a patient with COVID-19.

Bayesian Networks (BNs) are proposed for decision support, and they allow probabilistic reasoning. They are a probability-based inference model, increasingly used in the medical domain as a method of knowledge representation for reasoning under uncertainty for a wide range of applications, including disease diagnosis [16].

Different from other techniques, BNs have been used to interpret and explain COVID-19 data, BNs integrate multiple sources of data in a single model that provides a statistical estimates model. This last presents the uncertainty concerning mechanisms that generate the data [34].

Many papers research on COVID-19 used Bayesian networks for handling uncertainty related to COVID-19 diagnosis, we cite [18], [7], [18], [43], [34], [51], these studies propose a Bayesian network model to predict the diagnosis of COVID-19 according to different purposes.

In this paper, we have combined Bayesian networks with classical ontologies to harness the power of Bayesian networks, based on probability theory and the inference and probabilistic reasoning mechanisms proposed by BNs, on the knowledge base represented by a classical ontology, which gives us a probabilistic ontology.

Probabilistic ontologies based on Multi-Entity Bayesian Networks (MEBN) [8] allow us to describe uncertain knowledge in a reasoned and structured way and to model uncertainty using probability factors and causal links. These methods have been used in the medical field through several languages such as PR-OWL, thus giving way to the modeling of probabilistic knowledge through ontologies. These are used to describe domain knowledge, and the uncertainty associated with this knowledge in a structured and shareable way, in a format that can be read and processed by a computer.

Bayesian methods allow the system, on the one hand, to integrate expert knowledge with machine learning, to provide understandable models, and on the other hand, to provide, in a natural manner, probabilistic predictions making it possible to keep account for uncertainty when making decisions.

**Problems**: The confronted problems are as follows:

- How to predict the diagnosis of COVID-19 in an uncertain environment; and how to use Bayesian networks and ontologies to manage associated uncertainty for decision support.
- How to integrate the knowledge of an expert in a probabilistic model to achieve machine learning to use probabilistic inference for predicting diagnosis.
- How to model the uncertain knowledge of the COVID-19 diagnosis, using probabilistic ontologies, which combine the two models of Bayesian networks and classical ontologies.

**Contribution**: our main contributions in this paper are summarized as follows:

- Development of a probabilistic ontology for COVID-19, containing important concepts, such as symptoms and risk factors related to COVID-19, and probabilistic knowledge.
- Predicting COVID-19 diagnosis based on probabilistic inference, using Situation-Specific Bayesian Network.
- Collecting real anonymous data of patients for constructing dataset used for learning.
- Our proposed system's results are promising and can be exploited in real environments.

**Motivation**: Bayesian networks are a very powerful formalism; on the one hand, it deals with random uncertainty through probability theory which represents a very efficient framework for the management of degrees of influence of an attribute in another and its propagation throughout the Bayesian network and on the other hand, its ability to make inferences even with missing data, for example for a patient without doing the COVID test, and with a small number of symptoms or risk factors, we can reason about his observations to help for diagnosis. In addition to these advantages, Multi-Entity Bayesian Networks are more expressive than classical BNs because they integrate first-order logic with Bayesian probability and are more flexible in terms of inferences.

Among the limits of classical ontologies is the inability to reason under uncertainty, which leads us to build a probabilistic ontology to represent the knowledge base and the treatment of uncertainty provided by MEBN.

The remainder of this paper is organized as follows: Section 2 explores the background knowledge. Section 3 describes the related works, it presents a comprehensive review of related studies on COVID-19. Section 4 is devoted to the proposal of probabilistic ontology. An experiment and an evaluation were conducted to validate the proposed approach presented in Section 5. We draw conclusions and discuss future work in Section 6.

## 2. Background

### 2.1. Uncertain ontological knowledge

The Uncertainty Reasoning for the World Wide Web Incubator Group (URW3-XG) [2] of World Wide Web Consortium (W3C) [3] in his last report declares that Uncertainty is an intrinsic feature of many of the required tasks, and a full realization of the World Wide Web as a source of processable data and services demands formalisms capable of representing and reasoning under uncertainty. Although it is possible to use semantic markup languages such as OWL to represent qualitative and quantitative information about uncertainty, there is no established foundation for doing so. Therefore, each developer must come up with his/her own set of constructs for representing uncertainty. This is a major liability in an environment so dependent on interoperability among systems and applications.

The URW3-XG propose a high-level ontology to present various types of uncertainty: **ambiguity, empirical, randomness, vagueness, inconsistency** and **incompleteness.**

The new pandemic typically poses a challenge to data analytics, considering its limited information and the geographical and temporal evolution of the recent epidemic. Therefore, an accurate model for predicting the future behavior of a pandemic becomes challenging due to uncertainty [20]. In this context, we are interested in creating a probabilistic ontology to deal with this uncertainty.

### 2.2. Bayesian Networks

One of the most promising approaches to deal with uncertainty is Bayesian Networks (BN) [40]. A Bayesian Network *N* is a triplet *(V,A,P)*, where:

- *V* is a set of variables,
- *A* is a set of arcs, which together with *V* constitutes a direct acyclic graph *G=(V,A)*,
- *P* is a set of conditional probabilities of all variables given their respective parents.

The joint distribution for a BN is equal to the product of *P(node/parents(node))* for all nodes. Bayesian networks are powerful models that compactly represent the joint probability distribution defined by the set of variables under study [12].

---

[2] `https://www.w3.org/2005/Incubator/urw3/group/draftReport.html`

[3] `https://www.w3.org`

### 2.3. Multi-Entity Bayesian Networks (MEBN)

Multi-Entity Bayesian Networks [31] integrate first-order logic with Bayesian probability. MEBN logic expresses probabilistic knowledge as a collection of MEBN fragments (MFrags) organized into MEBN Theories (MTheories). Formally, an MFrag F is defined as :

$$F = (C, I, R, G, D) \text{ Where}$$

- $C$ is a finite set of values a context can take form as a value.
- $I$ is a set of input random variables
- $R$ is a finite set of resident random variables
- $G$ is a directed acyclic graph representing the dependency between input random variables and resident random variables conditional on context random variables in one-to-one correspondence.
- $D$ is a set of local conditional probability distributions where each member of $R$ has its own conditional probability distribution in set $D$.
- Sets $C$ , $I$, and $R$ are pairwise disjoint

### 2.4. Probabilistic Ontology PR-OWL

PR-OWL was developed as an extension enabling OWL ontologies to represent complex Bayesian probabilistic models. From [39], a probabilistic ontology is an explicit, formal knowledge representation that expresses knowledge about a domain of application. This includes:

- Types of entities that exist in the domain;
- Properties of those entities;
- Relationships among entities;
- Processes and events that happen with those entities;
- Statistical regularities that characterize the domain;
- Inconclusive, ambiguous, incomplete, unreliable, and dissonant knowledge related to entities of the domain; and
- Uncertainty about all the above forms of knowledge; where the term entity refers to any concept (real or fictitious, concrete or abstract) that can be described and reasoned about within the domain of application.

### 2.5. Reasoning under uncertainty

The term ”uncertainty reasoning” is meant to denote the full range of methods designed for representing and reasoning with knowledge when Boolean truth values are unknown, unknowable, or inapplicable. Commonly applied approaches to uncertainty reasoning include probability theory, fuzzy logic, subjective logic, Dempster-Shafer theory, and numerous other methodologies [13].

In the context of probabilistic reasoning using Bayesian networks, these networks are primarily used for performing probabilistic inference, which involves generating probabilistic statements regarding the variables depicted within the network.

Two types of inference can be performed on a Bayesian network: exact and approximate. Exact inference leverages the conditional independences within the network to calculate an exact posterior probability for each inference. Examples of algorithms that fall under exact inference include Bucket Elimination [4], Message Passing [40], and Junction Tree [28]. On the other hand, approximate methods such as Markov Chain Monte Carlo and variational methods are used for the second category. We refer to algorithms such as Likelihood Weighting [21], Backward Sampling [22], and Self Importance [48], which estimate probabilities by drawing from the set of possible combinations of network variables' states multiple times.

In MEBN, the process of inference is done by constructing Situation Specific Bayesian Network (SSBN) [14], which is a Bayesian network constructed by creating and combining instances of the MFrags in the MTheory. When each MFrag is instantiated, instances of its random variables are created to represent known background information, observed evidence, and queries of interest to the decision maker. The process of inference starts with a generative MTheory, adds a set of finding MFrags representing problem-specific information, and specifies the target nodes for our query. The process of MEBN inference [14] consists of:

1. Construct an SSBN, which is constructed by creating and combining instances of the MFrags in the generative MTheory.
2. When each MFrag is instantiated, instances of its random variables are created to represent known background information, observed evidence, and queries.

## 3. Related works

After studying several related works that deal with uncertainty about knowledge in the medical field, we have classified them into several categories: medical ontology-based approaches, Bayesian network-based approaches, and probabilistic ontologies-based approaches.

Various predictive models based on machine learning have been proposed in the literature to help reduce the load of COVID-19 on healthcare systems, we cite for example: [35], [52], [27], [38], [42], [1], [2], [33] and [5], these studies do not deal with uncertainty, but a few studies use a Bayesian Network model for handling uncertainty and predicting a diagnosis of COVID-19.

### 3.1. Bayesian network-based approaches

Bayesian Networks can be useful for decision-makers to make sense of complex information using a probabilistic approach.

The authors [18] present a Bayesian network that provides the basis for a practical CTA (Contact Tracing Apps) solution that does not compromise privacy. Users of the model can provide personal information about relevant risk factors, symptoms, and recent social interactions. The model then provides them feedback about the likelihood of the presence of asymptotic, mild or severe COVID-19.

In the paper [7] the authors propose a Bayesian network to predict the probability of COVID-19 infection, based on a patient's profile, the structure and prior probabilities have been amalgamated from knowledge. This paper constructs the solution of the Bayesian network created by [18] and the work by [43] for predicting COVID-19 status coupled with eventual prognoses. The network takes an input of observable symptoms and risk factors to produce a personalized probability score for disease status.

In [34], the authors use a Bayesian Network model to estimate the COVID-19 infection prevalence rate (IPR) and infection fatality rate (IFR) for different countries and regions, where relevant data are available. This combines multiple sources of data in a single model.

Terwangne et al. [51] developed a model named COVID-19 EPI-SCORE to predict the severity classification of patients hospitalized with COVID-19. The purpose of the study is to assess the COVID-19 severity classification. In this approach, Bayesian network analysis was used to build the model for predicting the accuracy of severity classification.

### 3.2. Ontologies-based approaches

This section provides a review of recent articles that have discussed the significant contributions using ontologies-based approaches.

In [37], the author proposed an approach adopted in their study that employs the use of an improved Case-Based Reasoning (CBR) model for reasoning tasks in the classification of suspected cases of COVID-19. Knowledge representation in the proposed framework was achieved using an ontology-based knowledge formalization technique.

In [47] the authors developed an ontology representing major novel coronavirus (SARS-CoV-2) entities, Ontology has a strong scope on chemical entities suited for drug repurposing. COVID-19 Ontology is used to attain the semantic interoperability and mapping between various entities and relationships in the COVID Knowledge SuperGraph.

The paper [6] discusses the advantages of using ontologies for describing and modeling psychological research questions. The authors use and apply CCOnto as a theoretical and formal description system to categorize psychological factors that influence student behavior during the COVID-19 situation.

In [29] the proposed solution is to design a COVID-19 ontology model, as well as an alert system combining vital sign parameters and symptom parameters, to identify suspected early cases of chronic obstructive pulmonary disease patients with COVID-19.

Furthermore, [25] presented the Coronavirus Infectious Disease Ontology (CIDO) which is a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis within the field of informatics.

In [32] the authors proposed a project nominated “Ontological and bio-informatics analysis of anti-coronavirus drugs and their implication for three drug re-purposing against COVID-19”. The authors fixed that their ontology-based bio-informatics strategy lead to drug prediction for COVID-19.

### 3.3. Probabilistic Ontology-based approaches

Various approaches have been made to represent uncertainty in ontology; one of them is the probabilistic ontology based on the Bayesian network.

One of the reasons that make the research in ontology languages focusing on deterministic approaches has limited expressiveness of knowledge uncertainty [44]. There is current research based on extending OWL and combined with Bayesian networks, so, it can represent uncertain knowledge, and especially, probabilistic information [17].

Until today, there is no standard language for representing probabilistic ontologies, the W3C consortium recommends using BNs for semantic web uncertainty processing[4]. There are some contributions that have been proposed in the literature to deal with uncertain knowledge in ontologies, we cite some examples:

- BayesOWL [15]: the authors propose a framework to model uncertainty in semantic web ontologies based on Bayesian networks, it is a probabilistic generalization of OWLlanguage. It provides a set of rules for the direct translation of OWL ontology into a Bayesian network, it also provides a method for incorporating probability constraints while constructing a Bayesian network.
- OntoBayes: Yang and Calmet [53] present an integration of the web ontology language OWL with Bayesian networks, called OntoBayes. This model makes use of probability and dependency-annotated OWL to represent uncertain information in BN structures
- In [19], the author proposed an ontology-based approach for constructing Bayesian networks, the proposed approach supports the construction of the structure of Bayesian networks and the conditional probabilistic tables of nodes of BN. This method enables the modification of Bayesian networks based on existing ontologies.
- In [26] a new methodology has been presented for using causal knowledge to extend and improve a standard hierarchical medical ontology. The structure of the variables and the symptoms of patients is obtained based on the medical dictionary of the terminology of regulatory activities (Medical Dictionary for Regulatory Activities Terminology.

Among the most important contribution of probabilistic ontologies language is a PR-OWL. PR-OWL is an upper ontology written in the Web Ontology Language (OWL) that provides constructs for representing probabilistic ontologies based on Multi-Entity Bayesian Networks [31]. Improvements in OWL compatibility in the second release of PR-OWL enable ontology designers to express uncertainty associated with an existing OWL ontology [10].

PR-OWL 2 has also been adopted by other domains such as maritime [30] and fraud detection in Brazil [9].

### 3.4. Analysis

- The medical field, in general, is full of uncertain knowledge, especially the COVID-19 diagnosis, despite the success of ontologies, classical ontologies have been widely used to model and represent data and reasoning with the knowledge of COVID-19, like the approaches [6], [25] and [32], however, classical ontologies do not provide adequate support to deal with uncertain knowledge. This is because classical ontologies are based on Boolean logic which does not allow representing uncertain data. So, After studying existing works, we find that works using classical ontologies ( [50], [37], [47], [25], [6], can't deal with uncertainty and they only have precise concepts and relationships.

---

[4] `https://www.w3.org/2005/Incubator/urw3/group/draftReport.html`

- In the work [29], the reasoning process used by the adaptation COVID-19 is a rule-based technique. For example, a buccal temperature greater than 38.5 ◦C and the patient having a headache is a rule for detection of COVID-19, the problem is that this rule is deterministic, i.e. all patients with observed temperature and headache values, must have a positive diagnosis of COVID-19, which is incorrect. The limit of this reasoning process is not in all of the suspected cases with the observation cited above, we can deduce the diagnosis. This is caused by the uncertainty of medical knowledge related to COVID-19 diagnosis, so, we need a formalism for representing uncertainty like the probabilistic theory using the stochastic rule, for example, if a patient has a 38.5 ◦C and he has a headache, we can detect COVID-19 with the probability of 0.8.
- Several techniques have been proposed to confront the problem of reasoning under uncertainty using Bayesian Networks [18], [7], [18], [43], [34], [51]. BNs are a well-established technique for dealing with uncertainty, they exploit probabilistic reasoning to provide information about causal relationships for a set of variables modeling a given domain. They are credible for decision support.
- PR-OWL is the most expressive language for representing uncertain knowledge, and probabilistic ontology compared to other existing approaches. The meta-model of PR-OWL is based on multi-entity Bayesian networks (MEBNs) that combine the expressivity of first-order logic with Bayesian probability theory, for this, we opt for PR-OWL for modeling COVID-19 probabilistic ontology.
- From our study on ontologies dealing with medical diagnosis, it was found that there are no probabilistic ontologies proposed for medical diagnosis, and also there is no probabilistic ontology for COVID-19 diagnosis.

## 4. Proposed modeling

In this paper, we have constructed a probabilistic ontology, which combines the notion of classical ontology for representing the knowledge base, and Bayesian Multi-Entity Networks for treatment and reasoning under uncertainty. Probabilistic ontologies based on Bayesian Networks make it possible to describe uncertain knowledge in a reasoned and structured way and to model uncertainty by probability distributions and causal links.

In fact, a classic ontology can't express uncertain knowledge because it is based on deterministic logic. Moreover, a probabilistic ontology is equipped with a mechanism of probabilistic reasoning based on the Multi-Entity Bayesian Network (MEBN) inference engine.

A MEBN has several advantages over standard Bayesian networks:

1. MEBN offers a very high level of expressiveness based on first-order logic to better represent the real world and perceived reality.
2. MEBN represents a simple formalism for quick and effective modeling especially when it comes to a problem that contains a lot of repetitive knowledge structures.
3. MEBN offers highly flexible inference mechanisms based on the generation of Situation Specific Bayesian Networks (SSBNs)
4. Modularity, MFrags can be easily added or removed from the modeled system without any loss in the structural coherence of the network.

We first discuss the various forms of uncertainty addressed in this paper before introducing our ontology.

### 4.1. The uncertainty types

The uncertainty types tackled in this paper are incompleteness and randomness, defined in the high-level ontology proposed by the Uncertainty Reasoning for the World Wide Web Incubator Group of the World Wide Web Consortium. They are defined as follows:

1. **Randomness** - the award is an instance of a class for which a statistical law governs whether instances are satisfied.
   We give an example of this uncertainty type in relation to our case study: For the diagnosis of COVID-19, it is difficult to predict it because there isn't an exact law that determines in a certain way the diagnosis from a set of symptoms, so we can't say for example, if the patient has a cough, fever, and headache, then, sure and certain that he is affected with the diagnosis, but there is uncertainty in the influence of these symptoms on the diagnosis, which induces us to use probability theory to model this randomness.
2. **Incompleteness** - information about the world is incomplete, it can take the form of missing information [23]. So, in our case study, with only a subset of symptoms and factors, the system can reason on the diagnosis, using probability distributions.

In the rest of this section, we propose a probabilistic ontology for COVID-19 diagnosis. Figure 1 shows our proposed process for predicting COVID-19 diagnosis, this process starts with constructing the probabilistic ontology, which is divided into three steps, constructing entities, rules, and the probabilistic components, then the reasoning of the diagnosis using the proposed ontology, based on probabilistic inference offered by Multi-Entity Bayesian Network. The details of these steps are presented in the following sections.

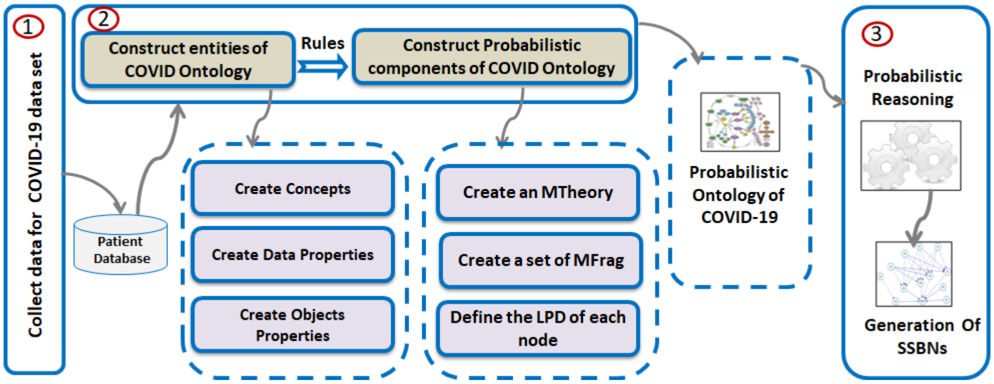

**Fig. 1.** Proposed modeling for aided diagnosis of COVID-19 based probabilistic ontology

The primary concern with the research on COVID-19 is the barrier prompted by the lack of adequate COVID-19 clinical data [11]. So, the first step to developing a probabilistic ontology is to collect anonymous data to use in the learning step.

The construction process of probabilistic ontology is described in the following major steps:

### 4.2. Collecting data for constructing COVID-19 dataset

It is worth emphasizing the significant problems facing any researcher with objectives like ours. There is a severe lack of transparency and important datasets are inaccessible. Diagnosing the COVID-19 pandemic is a multidisciplinary effort that requires significant data.

Data and their analytics are crucial components of such decision-making activities, one of the most difficult aspects is collected of accurate and detailed clinical data, even if these data have an imperfection.

We collected anonymous clinical data, and we analyzed it with expert domain, to facilitate better clinical decisions and treatment. Our dataset is made up of 300 real cases from patients in a hospital. This dataset contains anonymized medical data from patients that were tested for COVID-19. The important features of the data consist of symptoms and clinical signs of patients, risk factors, a set of radiological test results, and blood tests. The lack of such a dataset prompted us to build it ourselves.

The dataset includes patients that have been diagnosed positively with COVID-19 and those that have a negative diagnosis. Therefore both positive and negative diagnostic cases are present in the dataset. This latter is divided into two parts, one part is used for learning Bayesian Network parameters, which represents 80% of the dataset. The the remaining part is used for model testing.

Following discussions with the domain expert, we have identified the key **clinical symptoms** that are important for detecting COVID-19. These symptoms include Fever, Asthenia, Shortness of breath, Diarrhea, Dry cough, and Headache.

- **Biological symptoms** are:
  - Blood urea,
  - Creatinine,
  - Blood Sugar,
  - D_Dimers: is a fibrin degradation product, a small protein fragment present in the blood after a blood clot is degraded by fibrinolysis.
  - WBC: The normal level of lymphocytes on the complete blood count
  - HGB: Blood Hemoglobin (HB) is the amount of hemoglobin in 100 ml of blood
  - SpO2: Blood saturation, it estimates a patient's condition.
  - IGM/IGG: Are immunoglobulins produced by the immune system to provide protection against SARS-CoV-2. Anti-SARS-CoV-2 IgM and IgG can therefore be detected in samples from affected patients.
  - CRP: The dosage of protein C.
  - I.N.R/TP: I.N.R is an indicator of blood coagulation. TP is the prothrombin level is a biological test that evaluates the effectiveness of blood clotting in the body.
- **Radiological symptom** is a TDM, is an imaging test that scans an area of the body, such as the lungs and takes cross-sectional images of the area using a beam.
- **Risk factors** are Renal failure, Cardiovascular history, Diabetes, High blood pressure, Morbid obesity, and Chronic lung disease.

## 4.3. Constructing the probabilistic ontology of COVID-19

In this section, we present the development of the COVID-19 probabilistic ontology. While there is robust literature on ontology engineering and knowledge engineering for Bayesian Networks, the literature contains little guidance on how to model a probabilistic ontology. To fill the gap, Carvalho [10] proposed the Uncertainty modeling Process for Semantic Technologies, which describes the main tasks involved in creating probabilistic ontologies.

The First step in this model is to define the **entities, attributes, and relationships** by looking at the set of goals defined. After establishing the entities, their attributes, and relationships, we can proceed to specify the **rules** governing our Probabilistic Ontology (PO). Subsequently, we can create the **probabilistic components** by defining the groups and their elements. These components will aid in the implementation of the PO. In this step we define the Multi-Entity Bayesian Network, this network is composed of an MTheory and several MFrags, each MFrag contains the nodes and the corresponding probability distributions.

**1. Entities and properties of COVID Ontology:** After the data collection process, and after several discussions with the domain expert, we determined the domain entities with their properties, to model the entities of COVID-19 ontology, figure 2 shows the entities and attributes of COVID Ontology, it contains five classes.

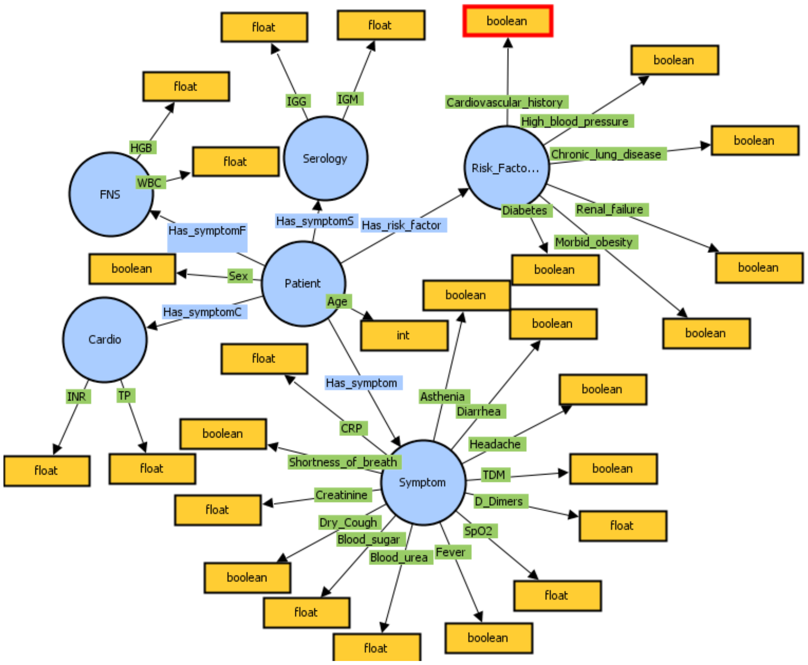

**Fig. 2.** Entities, attributes and relationships of COVID Ontology

The entities are presented in figure 2, which contains the following five classes.

- **Patient class** with the properties age and sex.
- **Symptom class**: with the properties Fever, Asthenia, Diarrhea, Blood urea, Headache, Creatinine, Blood Sugar, TDM, Shortness of breath, D_Dimers, SpO2, Dry cough, and CRP.
- **Risk factors class**: with the properties: Renal failure, Cardiovascular history, diabetes, High blood pressure, Morbid obesity, and Chronic lung disease.
- **Cardio class**: with the properties TP and INR.
- **Serology class**: with the properties IGG, and IGM.
- **FNS class**: with the properties HGB and WBC.

**2. Rules:** Generally, probabilistic rules are initially described using qualitative probability statements. The implementation of a probabilistic ontology requires specifying numerical probabilities and stochastic rules. Probability values can be obtained from domain experts or calculated from observation.

To determine stochastic rules, we propose to create a classical Bayesian Network structure, where all variables and dependencies must be linked, according to the causal relationship between the nodes, to extract the probability values. Then, we applied Expectation Maximisation (EM) for automatic learning parameters.

The stochastic rules are defined according to causal links between the nodes of the Bayesian network, taking into consideration the values of the probability distribution.

We present an example of stochastic rule between two nodes: TP which can take two values low and normal and Cardiovascular_history, which can take two values true and false, on a patient P.

**Declaration of a rule on TP node:**

```
if any P have (Cardiovascular_History=false ) [
   low = 0.05,
   normal = 0.95
  ]else [
   if any P have (Cardiovascular_History=true ) [
   low = 0.82,
   normal = 0.18
   ]else [
    low =0.5,
    normal = 0.5
 ]
```

**3. Probabilistic components of COVID Ontology:** The probabilistic COVID ontology aims to predict whether a patient is touched by COVID-19 or not, based on a set of evidence which represents symptoms and risk factors. Its construction pass through two steps.

In the **first step**, we design the different components of probabilistic COVID ontology PR-OWL2, which is based on a Multi-Entity Bayesian Network.

The MEBN language represents knowledge as a collection of MEBN Fragments (MFrags), which are organized into MEBN Theories (MTheories). An MFrag represents a repeatable pattern of knowledge that may apply to multiple domain entities. An MFrag consists of

Random Variables (RVs), a fragment graph whose nodes represent RVs, and Local Probability Distributions (LPDs) of some of these RVs. The repeated structure is represented by allowing RVs to have arguments that can be filled in with domain entities. These arguments are called ordinary variables to distinguish them from random variables. MFrags may have three kinds of nodes.

For constructing the MEBN, we define all components of MTheory and the set of MFrags of this MTheory. Creating the MFrags is done by defining the nodes, resident nodes, context nodes, and input nodes. Figure 3 shows the MFrags of COVID Ontology.

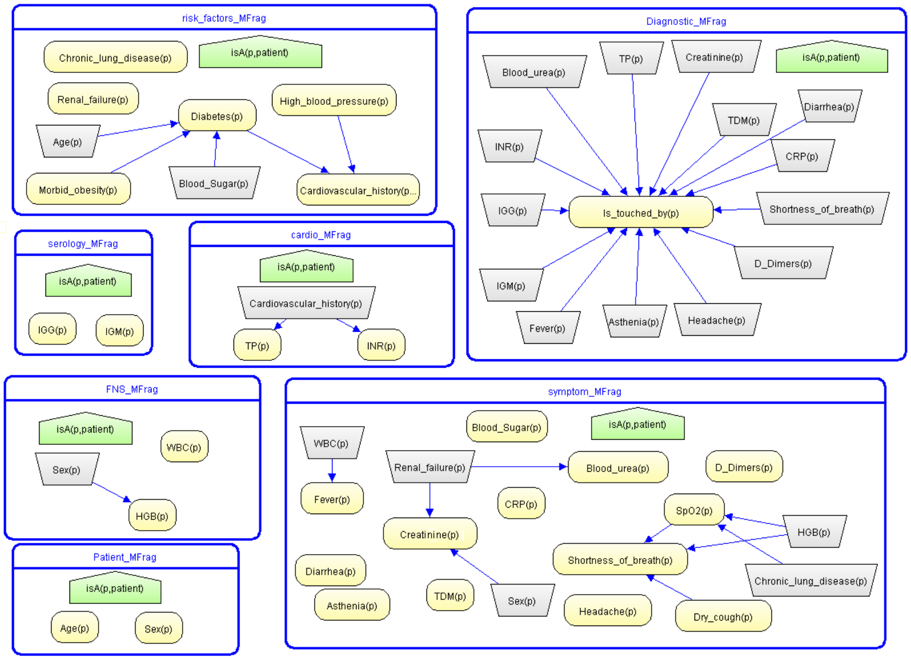

**Fig. 3.** MFrags of COVID Ontology

COVID probabilistic ontology presented in figure 3 contains 7 MFrags defined as follows:

- **Patient MFrag** contains the different attributes of the patient: age and sex, which are defined as resident nodes.
- **Risk-factors MFrag** contains the different risk factors for COVID-19 disease, it is defined by the resident nodes Chronic lung disease, Renal failure, High blood pressure and Morbid obesity which influence Diabetes node, this node is influenced also by input nodes Blood sugar and Age.
- **Serology MFrag** contains the different attributes of Serology: IGG and IGM, which are defined as resident nodes.
- **Cardio MFrag** contains the different attributes of Cardio: TP and INR, which are defined as resident nodes, and influenced by input node Cardiovascular history.

- **The FNS MFrag** contains the different attributes of FNS: WBC and HGB, which are defined as resident nodes. HGB is influenced by input node Sex.
- **Symptom MFrag** contains the different symptoms of Diarrhea, Asthenia, TDM, Blood sugar, CRP, Headache, and Shortness of breath which are influenced par Dy cough and SpO2. This later is influenced by input nodes HGB and Chronic lung disease. In this MFrag we define the Creatinine node influenced by input node renal failure and sex.
- **Diagnosis MFrag** with is represented by the input nodes IGG, IGM, INR, Blood urea, TP, Creatinine, TDM, Diarrhea, CRP, Shortness of breath, D_Dimers, Headache, Asthenia and Fever.

In the **second step**, we define the LPD of each node based on statistical study applied to collected data of COVID-19 anonymous patients. So, to complete the probabilistic ontology we add the stochastic rules between the random variables which be used in the reasoning process, but before this task, we compute the conditional probabilities between the variables, so that we can define a pseudo (called LPD) code for each variable.

All the components of the Multi-Entity Bayesian Network presented by the different MFrags with its nodes showing in Figure 3 are represented in the PR-OWL probabilistic ontology, we present a fragment that represents the PR-OWL code of the Asthenia node using the PR-OWL language.

```
<NamedIndividual
rdf:about=”file:/unbbayes−4.22.18/COVID.owl#Asthenia_1”>
        <rdf:type   rdf:resource=”&pr−owl2;OrdinaryVariableArgument”/>
            <pr−owl2:hasArgumentNumber  rdf:datatype=”&xsd;integer”>
            1
            </pr−owl2:hasArgumentNumber>
        <pr−owl2:isArgumentOf
         rdf:resource= ”file:/unbbayes−4.22.18/COVID.owl  #MEXPRESSION_Asthenia”/>
        <pr−owl2:typeOfArgument
         rdf:resource= ”file:/unbbayes−4.22.18/COVID.owl#symptom_MFrag.p”/>
</NamedIndividual>

<!−− file:/unbbayes−4.22.18/COVID.owl#Asthenia_Table —>
<NamedIndividual  rdf:about=
         ”file:/unbbayes−4.22.18/COVID.owl#Asthenia_Table”>
        <rdf:type  rdf:resource=”&pr−owl2;DeclarativeDistribution”/>
        <pr−owl2:hasDeclaration  rdf:datatype=”&xsd;string”>
         [  false=0.89,
            true=0.11
         ]
         </pr−owl2:hasDeclaration>
</NamedIndividual>
```

## 4.4. Probabilistic inference using a probabilistic ontology of COVID-19

When a query is submitted, we use the proposed reasoning in MEBN, consisting construct a BN to answer the query, this process is called SSBN construction. Reasoning in MEBN consists of the generation of a Situation-Specific Bayesian Network (SSBN), a minimal Bayesian network sufficient to solve a set of target nodes for which it is necessary to calculate the probability.

The reasoning process for predicting COVID-19 diagnosis is realized by generating an SSBN for each patient, the set of nodes in SSBN represents a set of evidence of a patient, which represents the different values of symptoms and risk factors.

To construct the SSBN for each query of a patient, the process of reasoning is an iterative bottom-up process on the MFrags having a prior knowledge base of prior probability distribution across the resident random variables. This algorithm involves d-separation and inferring an intermediate Bayesian Network obtained from the set of Resident random variables for every iteration until the iterations are terminated.

The output of this process is an SSBN, which is a minimal Bayesian Network sufficient to obtain the posterior distribution for a set of instances of target random variables, given a set of occurrences of random variables. A standard Bayesian network inference algorithm is applied to SSBN.

The algorithm used for Bayesian inference is the junction tree algorithm; it propagates evidence through the whole structure of BN. Junction tree inference is used to perform efficient probabilistic inference in Bayesian networks. It involves transforming the network into a junction tree, where each node represents a cluster of related variables and joint probability distributions over those variables. The tree is constructed as a set of maximal cliques, and messages containing information about the joint probability distributions are passed between nodes to perform inference.

Finally, the answer to the query is obtained by inspecting the posterior probabilities of the target nodes.

In order to illustrate the mechanism of the inference in our system, let's suppose that we have a patient named "p2", with a set of evidences.

The SSBN generated for the query IsTouchedBy (p2) is shown in figure 4. For "p2" with a set of its evidences, we can see in the SSBN that there is a chance of $56,25\%$ that p2 is not suffering from COVID-19.

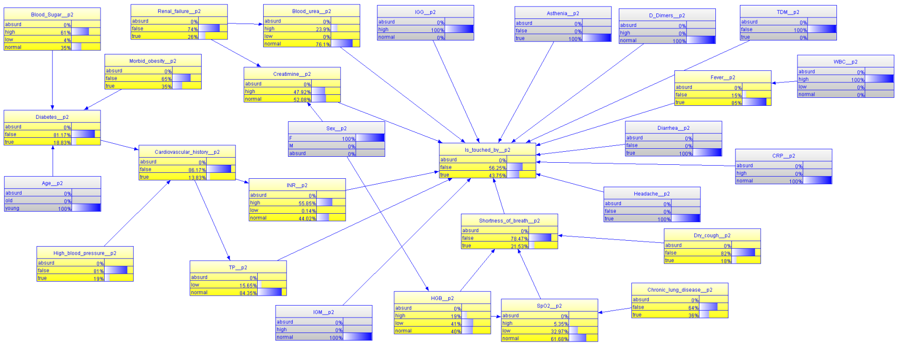

**Fig. 4.** The generated SSBN from evidences of p2

## 5. Experimentation

We have constructed our COVID-19 probabilistic ontology, to apply probabilistic inference for the diagnosis task, it is represented by PR-OWL language using the tool

UnbBayes[1]. UnBBayes is a probabilistic network framework written in Java, it has a graphical interface.

In this section, we present the performance of our proposed modeling, tested on 20% of the dataset dedicated to model testing. We have used metrics are accuracy, precision, recall, and F1-score.

- **Accuracy** is the probability that the test will provide correct results, that is, be positive in sick patients and negative in healthy patients, it is the probability of the true positives and true negatives.
- **Precision** is the rate of the positive predictions that are actually positive.
- **Recall** it is the rate of true positives and indicates the classifier's ability to detect people with COVID-19 correctly.
- **F1-score** can be used as a single measure of performance of the test for the positive class. The F1-score is the harmonic mean of precision and recall.

This metric mainly is very applicable to cases of diagnosis in medicine since it allows for increased confidence and acceptability of results, an illustration of these parameters on our probabilistic ontology inference can be found in Table 1.

**Table 1.** Evaluation measures of our COVID-19 diagnosis model using ML metrics

| **Measures** | **Values** |
|---|---|
| Precision | 0.95 |
| Recall | 1 |
| F1-score | 0.97 |
| Accuracy | 0.97 |

The diagnosis accuracy metric was used to evaluate the ability of a diagnostic COVID to correctly identify by the probabilistic inference. In the medical world, the use of sensitivity, specificity, Positive predictive value (PPV), and Negative Predictive Value (NPV) metrics are more frequent, than other metrics.

- **Specificity** metric is the capacity of classifying healthy patients as negatives. It is the rate of true negatives.
- **Sensitivity** metric can be calculated in the same way as recall. it measures the ability to give a positive result when a hypothesis is verified, it refers to the ability to detect a maximum number of patients.
- **Positive Predictive Value (PPV)** It is the ratio of patients truly diagnosed as positive to all those who had positive test results (including healthy subjects who were incorrectly diagnosed as patients).
- **Negative Predictive Value (NPV)** It is the ratio of subjects truly diagnosed as negative to all those who had negative test results.

This metric mainly is very applicable to cases of diagnosis in medicine since it allows for increased confidence and acceptability of results, an illustration of these parameters on our probabilistic ontology inference can be found in Table 2.

---

[1] `https://sourceforge.net/projects/unbbayes`

**Table 2.** Evaluation measures of our COVID-19 diagnosis model using Diagnostic test evaluation

| Measures | Values |
|---|---|
| Specificity | 94.12% |
| Sensitivity | 100% |
| Positive Predictive Value | 95.92% |
| Negative Predictive Value | 100.00% |

**Discussion**: According to the presented results in Table 1 and Table 2:

- We note a high-specificity value in our system, which means it will correctly rule out almost everyone who doesn't have the disease and won't generate many false-positive results. The specificity value with 94% will correctly return a negative result for 94% of people who don't have the disease but will return a positive result (a false-positive for 6% of the people who don't have the disease and should have tested negative.
- Sensitivity measures how often a test correctly generates a positive result for people who have the condition that's being tested for (also known as the "true positive" rate). The Sensitivity value in our system is 100%, so, it will flag everyone who has the disease and not generate false-negative results. Therefore, it will correctly return a positive result for 100% of people who have the disease, but it won't return a negative result for the people who have the disease (the false Negative value is 0%).
- The PPV is the probability that the disease is present when the test is positive, in our system, we have a 95.92% value of PPV, which is a high value that indicates that a positive test result is likely correct, in 5% of healthy subjects who were incorrectly diagnosed as patients.
- NPV is the probability that the disease is not present when the test is negative, in our system we have a 100% value of NPV, so all subjects who were diagnosed as healthy are healthy subjects, and there are no patients who were incorrectly diagnosed as healthy.
- We can conclude that using the reasoning-based probabilistic ontology increases the prediction quality in terms of precision, recall, accuracy and F1-Score and handle the uncertainty related to the diagnosis of COVID-19. So, our system gives very good performance values, applied to the dataset instances used for the test.

A comparative analysis of the performance of our proposed approach was carried out with the machine learning methods: Logistic regression (LG), Support Vector Classifier(SVC), Decision tree (DT), Random Forest (RF), Gaussian Naive Bayes (GNB), and our model of Probabilistic Ontology (PO). The results of the performed evaluation are summarized in Table 3.

From Table 3, we can note that the experimental results for various machine learning models reveal that they provide less values of evaluation parameters regarding the probabilistic approach.

The best value of precision (1), is presented using SVC and FR methods, but they haven't the best value of recall which is 0.75, which resulted in a 0.83 of F1-Score, however, GNB and DT presented the best value of recall (1), but, it presented respectively 0.5

**Table 3.** Evaluation results using machine learning methods

| Measures | LR | SVC | DT | RF | GNB | **PO** |
|---|---|---|---|---|---|---|
| Precision | 1 | 1 | 0.75 | 1 | 0.5 | **0.95** |
| Recall | 0.75 | 0.75 | 1 | 0.75 | 1 | **1** |
| F1-score | 0.83 | 0.83 | 0.83 | 0.83 | 0.66 | **0.97** |
| Accuracy | 0.87 | 0.87 | 0.75 | 0.87 | 0.5 | **0.97** |

and 0.75 for precision value, which decreases its F1-Score measure to 0.5 for GNB and 0.75 for DT. LR presented intermediary values of precision and recall.

The proposed probabilistic model has exhibited promising predictive ability. It can be seen from Table 3 that the prediction for the probabilistic inference have the highest precision, recall, F1-Score, and accuracy and are, respectively, about **0.95, 1, 0.97**, and **0.97**, which signifies that the prediction basing on probabilistic inference is near to the prediction of the expert domain.

Histogram in Figure 5 outline the evaluation results using the standard evaluation measures defined in Table 3.

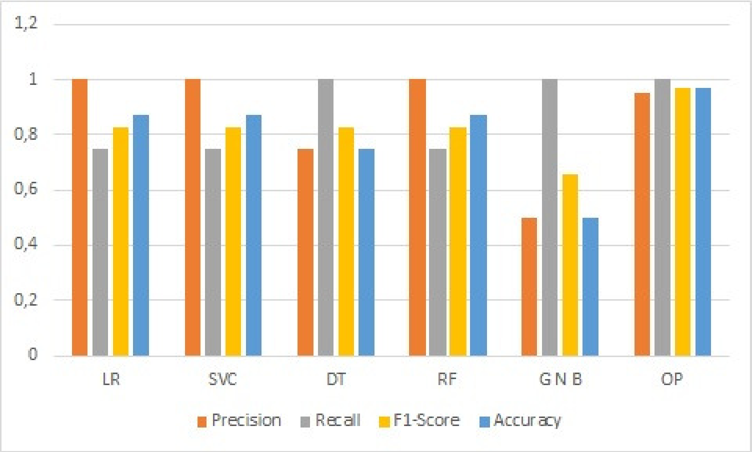

**Fig. 5.** Histogram of evaluation results

**Discussion**: The validity of a diagnostic test is evaluated in terms of its ability to detect subjects with COVID-19 as well as its capacity to exclude subjects without COVID-19. We can see from Figure 5 that:

- **Recall**: Our probabilistic ontology presents the best value of recall score (1), it is among the best learning machine methods which have the same value of recall, the two methods are: GNB and DT.
- **Precision**: Regarding precision value, OP achieved a very good value (0.95). It presents the second-best value after the methods of RH, SVC, and RD.

- **F1-Score**: For F1-Score, which combines precision and recall, the proposed modeling presents a high value (0.97), greater than the learning machine methods.
- **Accuracy**: For accuracy value, Our modeling based on probabilistic ontology presents the best value of accuracy (0.97) compared with machine learning methods.

The conducted evaluation of our modeling demonstrates that the probabilistic model based on Multi-Entity Bayesian Networks using ontology structure has improved the performance of the probabilistic ontology reasoning task, applied to COVID-19 diagnosis, so, probabilistic knowledge helps in improving tasks like disease prediction and risk propagation.

The current paper demonstrates the great potential of the probabilistic model as Multi-Entity Bayesian Network in tackling the uncertainty in knowledge for diagnosing COVID-19 by facilitating complex decision-making and fact interrogation.

## 6. Conclusion

In the medical field, doctors are regularly called upon to make several decisions; some of these decisions are taken easily when the diagnosis is easy, the treatment chosen is effective, and the risks are zero, while in other cases, such as COVID-19, the right decision to make is not obvious, because it is a new disease, it appeared recently in December 2019. The information necessary for its diagnosis is filled with uncertainty.

We aim to handle the uncertainty associated with the diagnosis of the new COVID-19 pandemic. Unlike the commonly known influenza, it came with limited information and very high uncertainty. Therefore, knowledge regarding the new epidemic needs to be treated due to the absence of a prior case similar to the recent pandemic. In this work, we are interested in tackling two types of uncertainty associated with the COVID-19 process: randomness and incompleteness.

Many studies applied ontology to represent knowledge, Ontologies are used for structuring and sharing knowledge because the common good practice of ontology engineering hinges on reusing and integrating existing ontologies. With the emergence of the immense quantity of data from various contexts, the need for shareable integration of domain knowledge increases. Ontologies are the key element for interoperability. However, classical ontologies do not have built-in mechanisms for representing or inferring with uncertain knowledge.

Network is a well-established technique for handling uncertainty within the artificial intelligence (AI) community. They exploit Bayes's probabilistic reasoning to provide insights into the causal relationships between the contributors and outcomes of an event.

The objective of this proposal is to predict the diagnosis of COVID-19 using the uncertainty management techniques in the field of AI, particularly the theory of probability, and the associated techniques, such as Bayesian networks and probabilistic ontologies.

The aim was to develop a probabilistic ontology for the aided diagnosis of COVID-19 infection in an individual, to take the necessary measures before reaching danger.

This model combines the advantages of classical ontologies for representing the knowledge base, the Multi-Entity Bayesian network for the uncertainty treatment associated with the prediction process, and the provided probabilistic inference mechanism.

We tested our modeling system on the collected dataset cases, and it gave good results in terms of precision and recall. The advantage of a Bayesian network is that it can predict the diagnosis even in the case of missing values, without requiring some biological tests. With only a few symptom values, the system can respond to the given observations.

To treat the various symptoms and risk factors associated with COVID-19 variants, we envisage following the development of these variants, and testing our method on big data.

## References

1. Al-Qaness, M.A., Ewees, A.A., Fan, H., Abd El Aziz, M.: Optimization method for forecasting confirmed cases of covid-19 in china. Journal of Clinical Medicine 9(3), 674 (2020)
2. Albahri, A.S., Hamid, R.A., Alwan, J.K., Al-Qays, Z., Zaidan, A., Zaidan, B., Albahri, A., AlAmoodi, A.H., Khlaf, J.M., Almahdi, E., et al.: Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (covid-19): a systematic review. Journal of medical systems 44, 1–11 (2020)
3. Ansari, A.Q., Biswas, R., Aggarwal, S.: Proposal for applicability of neutrosophic set theory in medical ai. International Journal of Computer Applications 27(5), 5–11 (2011)
4. Becker, A., Naim, P.: Les réseaux bayésiens: modèles graphiques de connaissance. Eyrolles (1999)
5. Bhuvana, J., Mirnalinee, T., Bharathi, B., Sneha, I.: Efficient generative transfer learning framework for the detection of covid-19. Computer Science and Information Systems (00), 33–33 (2022)
6. Bolock, A.E., Abdennadher, S., Herbert, C.: An ontology-based framework for psychological monitoring in education during the covid-19 pandemic. Frontiers in Psychology 12, 673586 (2021)
7. Butcher, R., Fenton, N.: Extending the range of symptoms in a bayesian network for the predictive diagnosis of covid-19. medRxiv pp. 2020–10 (2020)
8. Carvalho, R.N., Laskey, K.B., Costa, P.C.: Pr-owl–a language for defining probabilistic ontologies. International Journal of Approximate Reasoning 91, 56–79 (2017)
9. Carvalho, R.N., Matsumoto, S., Laskey, K.B., da Costa, P.C.G., Ladeira, M., Santos, L.L.: Probabilistic ontology and knowledge fusion for procurement fraud detection in brazil. In: URSW (LNCS Vol.). pp. 19–40. Springer (2013)
10. Carvalho, R.N.: Probabilistic ontology: representation and modeling methodology. George Mason University (2011)
11. Chiroma, H., Ezugwu, A.E., Jauro, F., Al-Garadi, M.A., Abdullahi, I.N., Shuib, L.: Early survey with bibliometric analysis on machine learning approaches in controlling covid-19 outbreaks. PeerJ Computer Science 6, e313 (2020)
12. Conrady, S., Jouffe, L.: Bayesian networks and BayesiaLab: a practical introduction for researchers, vol. 9. Bayesia USA Franklin (2015)
13. Costa, P.C., Laskey, K.B., Blasch, E., Jousselme, A.L.: Towards unbiased evaluation of uncertainty reasoning: The urref ontology. In: 2012 15th International Conference on Information Fusion. pp. 2301–2308. IEEE (2012)
14. Da Costa, P.C.G., Laskey, K.B., Laskey, K.J.: Pr-owl: A bayesian ontology language for the semantic web. In: Uncertainty reasoning for the semantic web I: ISWC international workshops, URSW 2005-2007, revised selected and invited papers. pp. 88–107. Springer (2008)
15. Ding, Z., Peng, Y.: A probabilistic extension to ontology language owl. In: 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the. pp. 10–pp. IEEE (2004)
16. Eom, J.H., Kim, S.C., Zhang, B.T.: Aptacdss-e: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. Expert Systems with Applications 34(4), 2465–2479 (2008)

17. Fareh, M.: Modeling incomplete knowledge of semantic web using bayesian networks. Applied Artificial Intelligence 33(11), 1022–1034 (2019)
18. Fenton, N.E., McLachlan, S., Lucas, P., Dube, K., Hitman, G.A., Osman, M., Kyrimi, E., Neil, M.: A privacy-preserving bayesian network model for personalised covid19 risk assessment and contact tracing. MedRxiv pp. 2020–07 (2020)
19. Fenz, S.: An ontology-based approach for constructing bayesian networks. Data & Knowledge Engineering 73, 73–88 (2012)
20. Fong, S.J., Li, G., Dey, N., Crespo, R.G., Herrera-Viedma, E.: Composite monte carlo decision making under high uncertainty of novel coronavirus epidemic using hybridized deep learning and fuzzy rule induction. Applied soft computing 93, 106282 (2020)
21. Fung, R., Chang, K.: Weighting and integrating evidence for stochastic simulation in bayesian networks, uncertainty in artificial intelligence 5 (1989)
22. Fung, R., Del Favero, B.: Backward simulation in bayesian networks. In: Uncertainty Proceedings 1994, pp. 227–234. Elsevier (1994)
23. Hamel, O., Fareh, M.: Missing types prediction in linked data using deep neural network with attention mechanism: Case study on dbpedia and uniprot datasets. In: Information Technology for Management: Approaches to Improving Business and Society: AIST 2022 Track and 17th Conference, ISM 2022, Held as Part of FedCSIS 2022, Sofia, Bulgaria, September 4–7, 2022, Extended and Revised Selected Papers. pp. 212–231. Springer (2023)
24. Han, P.K., Klein, W.M., Arora, N.K.: Varieties of uncertainty in health care: a conceptual taxonomy. Medical Decision Making 31(6), 828–838 (2011)
25. He, Y., Yu, H., Ong, E., Wang, Y., Liu, Y., Huffman, A., Huang, H.h., Beverley, J., Hur, J., Yang, X., et al.: Cido, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. Scientific data 7(1), 181 (2020)
26. Hu, H., Kerschberg, L.: Evolving medical ontologies based on causal inference. In: 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp. 954–957. IEEE (2018)
27. Jangam, E., Barreto, A.A.D., Annavarapu, C.S.R.: Automatic detection of covid-19 from chest ct scan and chest x-rays images using deep learning, transfer learning and stacking. Applied Intelligence pp. 1–17 (2022)
28. Jensen, F.: Bayesian updating in recursive graphical models by local comutations. Computational Statistics and Data Analysis 4, 269–282 (1990)
29. Kouamé, K.M., Mcheick, H.: An ontological approach for early detection of suspected covid-19 among copd patients. Applied System Innovation 4(1), 21 (2021)
30. Laskey, K.B., Haberlin, R., Carvalho, R.N., da Costa, P.C.G.: Pr-owl 2 case study: A maritime domain probabilistic ontology. In: STIDS. pp. 76–83 (2011)
31. Laskey, K.B., Costa, P.C., Janssen, T.: Probabilistic ontologies for knowledge fusion. In: 2008 11th international conference on information fusion. pp. 1–8. IEEE (2008)
32. Liu, Y., Chan, W., Wang, Z., Hur, J., Xie, J., Yu, H., He, Y.: Ontological and bioinformatic analysis of anti-coronavirus drugs and their implication for drug repurposing against covid-19 (2020)
33. Mondal, M.R.H., Bharati, S., Podder, P., Podder, P.: Data analytics for novel coronavirus disease. informatics in medicine unlocked 20, 100374 (2020)
34. Neil, M., Fenton, N., Osman, M., McLachlan, S.: Bayesian network analysis of covid-19 data reveals higher infection prevalence rates and lower fatality rates than widely reported. Journal of Risk Research 23(7-8), 866–879 (2020)
35. Nour, M., Cömert, Z., Polat, K.: A novel medical diagnosis model for covid-19 infection detection based on deep features and bayesian optimization. Applied Soft Computing 97, 106580 (2020)
36. Organization, W.H., et al.: Coronavirus disease 2019 (covid-19) situation. Geneva: Report-121 (2020)

37. Oyelade, O.N., Ezugwu, A.E.: A case-based reasoning framework for early detection and diagnosis of novel coronavirus. Informatics in Medicine Unlocked 20, 100395 (2020)
38. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers in biology and medicine 121, 103792 (2020)
39. Paulo Cesar, G.: Da Costa, Bayesian semantics for the semantic web. Ph.D. thesis, Ph. D. Thesis, Fairfax, VA, USA (2005)
40. Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan kaufmann (1988)
41. Peñaloza, R.: Introduction to probabilistic ontologies. Reasoning Web. Declarative Artificial Intelligence: 16th International Summer School 2020, Oslo, Norway, June 24–26, 2020, Tutorial Lectures 16 pp. 1–35 (2020)
42. Pirouz, B., Shaffiee Haghshenas, S., Shaffiee Haghshenas, S., Piro, P.: Investigating a serious challenge in the sustainable development process: analysis of confirmed cases of covid-19 (new type of coronavirus) through a binary classification using artificial intelligence and regression analysis. Sustainability 12(6), 2427 (2020)
43. Prodhan, G., Fenton, N.: Extending the range of covid-19 risk factors in a bayesian network model for personalised risk assessment. medRxiv pp. 2020–10 (2020)
44. Riali, I., Fareh, M., Bouarfa, H.: Fuzzy probabilistic ontology approach: A hybrid model for handling uncertain knowledge in ontologies. International Journal on Semantic Web and Information Systems (IJSWIS) 15(4), 1–20 (2019)
45. Riali, I., Fareh, M., Ibnaissa, M.C., Bellil, M.: A semantic-based approach for hepatitis c virus prediction and diagnosis using a fuzzy ontology and a fuzzy bayesian network. Journal of Intelligent & Fuzzy Systems 44(2), 2381–2395 (2023)
46. Rodriguez-Morales, A.J., Cardona-Ospina, J.A., Gutiérrez-Ocampo, E., Villamizar-Peña, R., Holguin-Rivera, Y., Escalera-Antezana, J.P., Alvarado-Arnez, L.E., Bonilla-Aldana, D.K., Franco-Paredes, C., Henao-Martinez, A.F., et al.: Clinical, laboratory and imaging features of covid-19: A systematic review and meta-analysis. Travel medicine and infectious disease 34, 101623 (2020)
47. Sargsyan, A., Kodamullil, A.T., Baksi, S., Darms, J., Madan, S., Gebel, S., Keminer, O., Jose, G.M., Balabin, H., DeLong, L.N., et al.: The covid-19 ontology. Bioinformatics 36(24), 5703–5705 (2020)
48. Shachter, R.D., Peot, M.A.: Simulation approaches to general probabilistic inference on belief networks. In: Machine intelligence and pattern recognition, vol. 10, pp. 221–231. Elsevier (1990)
49. Singhal, T.: A review of coronavirus disease-2019 (covid-19). The indian journal of pediatrics 87(4), 281–286 (2020)
50. Stancin, K., Poscic, P., Jaksic, D.: Ontologies in education–state of the art. Education and Information Technologies 25(6), 5301–5320 (2020)
51. de Terwangne, C., Laouni, J., Jouffe, L., Lechien, J.R., Bouillon, V., Place, S., Capulzini, L., Machayekhi, S., Ceccarelli, A., Saussez, S., et al.: Predictive accuracy of covid-19 world health organization (who) severity classification and comparison with a bayesian-method-based severity score (epi-score). Pathogens 9(11), 880 (2020)
52. Vrbačič, G., Pečnik, Š., Podgorelec, V.: Hyper-parameter optimization of convolutional neural networks for classifying covid-19 x-ray images. Computer Science and Information Systems 19(1), 327–352 (2022)
53. Yang, Y., Calmet, J.: Ontobayes: An ontology-driven uncertainty model. In: International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06). vol. 1, pp. 457–463. IEEE (2005)

**Messaouda Fareh** is a doctor in computer sciences, and a lecturer at the computer sciences Department, University of Blida 1, Blida, Algeria. She is a member of the Laboratory LRDSI. Her research interests include ontology engineering and knowledge, information heterogeneity, uncertain knowledge of semantic and linked data, and data mining.

**Ishak Riali** is a Ph.D. in data science, in the computer science department at University of Blida1, Blida, Algeria. His primary research interests are artificial Intelligence, ontology engineering, uncertain Knowledge, Knowledge-Based Systems, semantic interoperability, and machine learning for Linked Data.

**Hafsa Kherbache** received her BSc (2019) and MSc (2021) degrees in Software Engineering from the University of Blida1. Her interests include knowledge and ontological engineering, semantic web, and probabilistic knowledge.

**Marwa Guemmouz** obtained her BSc degree in 2019 and her MSc in Software Engineering in 2021 from the University of Blida1. Her interests include knowledge and ontological engineering, semantic web, and probabilistic knowledge.

*Received: August 29, 2022; Accepted: April 04, 2023.*