COVID-19 VACCINATION CLASSIFICATION OF OPINION MINING WITH SEMANTIC KNOWLEDGE-BASED DECISION MAKING

Nikhila Polkampally[1], D. Rakesh Kumar[2],
G. Soma Sekhar[2], and Mettu Karuna Sri Reddy[1]

[1]Department of Computer Science and Engineering,
Geethanjali College of Engineering and Technology, India
[2]Department of Computer Science and Engineering, Associate Professor at
Geethanjali College of Engineering and Technology, India

ABSTRACT

The Covid-19 ontology is to classify the data using a supervised learning approach in machine learning, which has been preprocessed. Afterthe classification is done, with thehelp of opinion mining with decision-making, the classified data is stored in the database using semantic webontology using the protégé tool. The data will be retrieved through SPARQL which helps to retrieve complex queries, followed by the output based on the given query. This Covid-19 ontology helps in analyzing the risk factors and treatment plans for the respective individuals i.e., students based on their given details which include diagnosis, symptoms, and vaccination history. The information given by the students can be automatically processed and with the help of SWRL (Semantic Web Rule Language), the risk factor and treatment plans for the students are inferred from the given knowledge.

KEYWORDS

Covid-19, SPARQL, Semantic Web, SWRL, Ontology.

1. INTRODUCTION

Corona viruses square measure an enormous family of viruses. several corona viruses square measure higher famed to cause metabolism infections in humans ranging from disease to extra severe diseases like Near East metabolic process Syndrome (MERS) in 2003 related to degreed severe acute metabolic process syndrome (SARS) in 2012. COVID-19 is a disease caused by a newly identified coronavirus. Before the events in Beijing, China, in the solar calendar month of 2019, this new infection and illness were unknown. The novel virus was first known as 2019-nCov. The International Committee on Taxonomy of Viruses (ICTV) tasks experts dubbed it the SARS-Cov-2 virus because it was so similar to the one that caused the outbreak. Small droplets from the nose or lips of a person with COVID-19 coughs or exhales can spread the disease from person to person. These droplets land on the person's surroundings, and the other individual contracts COVID-19 by touching these things or surfaces, subsequently contacting their eyes, nose, or mouth.

The WHO recommends collecting expectorated sputum, endotracheal aspirate, or Bronchoalveolar lavage from the upper and lower respiratory tracts. In the research, the genetic material extracted from the saliva or mucus sample is amplified using a reverse polymerase chain reaction (RT-PCR), which involves converting an RNA molecule into a dual DNA molecule. If

there is sufficient genetic material, the search for conserved parts of the coronavirus genetic code begins. The presence of the parts denotes the presence of the illness. COVID-19 affects people differently. The majority of persons who are infected will have mild to moderate symptoms and will recover without the need for hospitalization, while others will experience severe symptoms and require hospitalization.

Ontology is a philosophical field that explores the nature and structure of reality. An ontology is used to help humans or software agents share a common understanding of information structure, reuse domain knowledge, make domain assumptions obvious, distinguish domain knowledge from operational knowledge, and assess domain knowledge. Domain ontologies, Upper ontologies, Knowledge Representation ontologies, and General ontologies are the four categories of ontologies. Here, Covid-19 general ontology is used for analyzing the risk factors and treatment plan of the student.

However, it is necessary to represent information (branch, diagnosis, vaccination history, symptoms) about each student so that analysis of students can be performed automatically using Semantic Web Ontology technology such as Resource framework description (RDF), SWRL, and SPARQL queries. This ontology is focused on Six scenarios: Diagnosis, Recommendation, Risk, Symptoms, Treatment, and Vaccination History and has shown the vaccination details of our college final year students in scenarios of Fully Vaccinated, Not Vaccinated, Started Vaccination.

The following is how the rest of the paper is organized: Section 2 examines and discusses related Covid-19 ontologies and methods to semantics. Section 3 discusses the framework of the Covid-19 Ontology. The suggested ontology is introduced in Section 4. Section 5explains the Ontology visualization tools. SWRL for the proposed treatment plan and risk factors are introduced to students in Section 6.Section 7 goes into detailed SPARQL queries for data analysis. Section 8 gives the results for the proposed Covid-19 Ontology. Finally, in Section 9, we conclude.

2. LITERATURE OVERVIEW

This project can be run on any graphical operating system like windows, ubuntu, Linux, etc. To run this project the specification is to have protégé tool 5.5 installed and it is used as the editor for giving the query.

Basic Formal Ontology is proposed by Robert *et al.* [1] It distinguishes between different sorts of ontologies and taxonomies, and shows how applied ontology, after constructing an ontology as a description of the types of entities in a certain area, draws on more traditional metaphysical principles. It explains the fundamental ideas of the Basic Formal Ontology (BFO), which is now used by more than a hundred ontology projects throughout the worldand gives examples of BFO-based domain ontologies. A standard foundation for Semantic Web technologies, Web Ontology Language (OWL), is also discussed. Throughout, it gives precise suggestions for the design and construction of domain ontologies.

Jing-Mei et al. [2] propose ontology query answering on databases. As a result, databases with OWL ontologies could solve SPARQL queries. Bi-directional methods are then examined to handle this type of customized Datalog application, re-turning accurate responses to the query using both forward and backward chaining within our logical framework.Semantic Web Performance Evaluation of Ontology Query Languages proposed by Zhijun Zhang *et al.* [3]. In this study, they will compare different ontology query languages and query systems, as well as provide user assessments. Elik Ertuğrul et al. [4] presented a pre-diagnosis-based knowledge system to anticipate Covid-19 smartphone users using the data and symptoms they have. The technology not only shares user information with specialists butalso analyzes symptoms as atool

to analyze potential Covid-19 hazards. To do that, a user must first enter details in the patient evaluation sheet, which includes aCovid-19 diagnostic test as well as other information.

Ontology of Knowledge Modeling (SARS-CoV-2) is proposed by Sherimon *et al.* [5] where the patient's data is classified as suspected, probable, and confirmed covid-19 virus and the time of execution is minimized.An Act Network COVID-19 Application Ontology is proposed by Visweswran S., (2021) *et al.* [6] where the harmonization of data elements is enabled and the data elements are critical to COVID-19 research.

Fully automatic ontology web development from RDB Scheme is proposed by Tabbasum Naz *et al.* [9] to create the ontologies easily without spending much time and overcoming the problems of existing systems that they are not fully automatic and not easily accessible. It mainly focuses on creating rules for RDB mappings and OWL Ontology development.

Sunil S Bhopal et al. [10] propose that children and young people under 25 have a reduced risk of COVID-19 death. Göğebakan K et al. [11] provide a review of recent advances in using the semantic web to prevent, diagnose, and treat diabetes mellitus. Diabetes must be prevented and detected in its early stages. To detect it in the early stages machine learning expert systems, logic, and semantic methodologies have been developed. Diabetes data can be represented as machine-processable metadata using the Semantic Web.

A sentiment analysis-based ontology on social networking website reviews is proposed by PratikThakora *et al.* [12] this exposed ontology mainly focuses on customers' tweets with negative sentiments retrieved and analyzed using an ontology-based method. The Canada Post has identified client dissatisfaction with their delivery services. A variety of technologies are used forthe extraction of Twitter data, cleaning the extracted data, subjective analysis, ontology model construction, and sentiment analysis. The organization could utilize the findings of this study to take corrective action and develop an automatic online response to the issues.

In this ontology,the information given by the students can be automatically processed and with the help of SWRL (Semantic Web Rule Language) and SPARQL queries, the risk factor and treatment plans for the students are inferred from the given OWL knowledge.
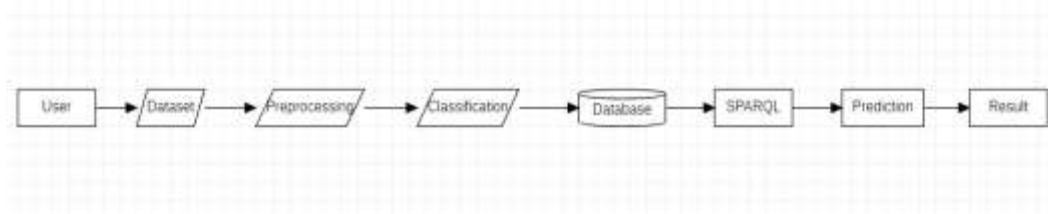
3. FRAMEWORK

Figure 1. The Framework of COVID-19 Ontology

The user collects the student's data and stores it in the dataset. The preprocessing is done after the data is collected and that preprocessed data will be classified using a supervised learning approach in machine learning. The classification is done with the help of opinion mining and with the help of decision making the classified data is stored in the database on the semantic web. The data will be retrieved through SPARQL which helps to retrieve complex queries. This Covid-19 ontology helps in analyzing the risk factors and treatment plans for the respective individuals i.e., students based on their given details which include diagnosis, symptoms, and

vaccination history. The information given by the students can be automatically processed and with the help of SWRL (Semantic Web Rule Language), the risk factor and treatment plans for the students are inferred from the given knowledge followed by the output.

4. COVID -19 ONTOLOGY

Classes, Attributes, and Relationships are the three major components of ontology [7]. Classes define broad concepts, Attributes explain the Classes in ontology (e.g., the student has a name, Id, Branch, roll number, etc.), and Relationshipslinkclasses in the same domain explicitly[4]. The developed Covid-19 ontology schema is shown in Figure 2.
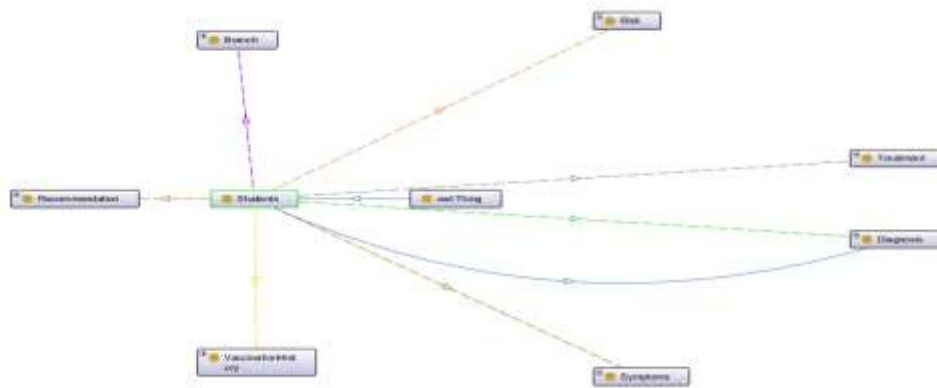
Figure 2. Schema of covid-19 ontology

4.1. Classes

Adding Classes

Step 1: Open the "Classes" tab

       Window → Tabs → Classes

Step 2: There is only one class in the empty class tree, owl: Thing, which is the superclass of everything.

Step 3: Create subclass Students. This is the subclass of owl: Thing.

Step 4: To add a subclass select owl: Thing and click on the icon shown in Figure 3 then the window will appear to enter the subclass name.
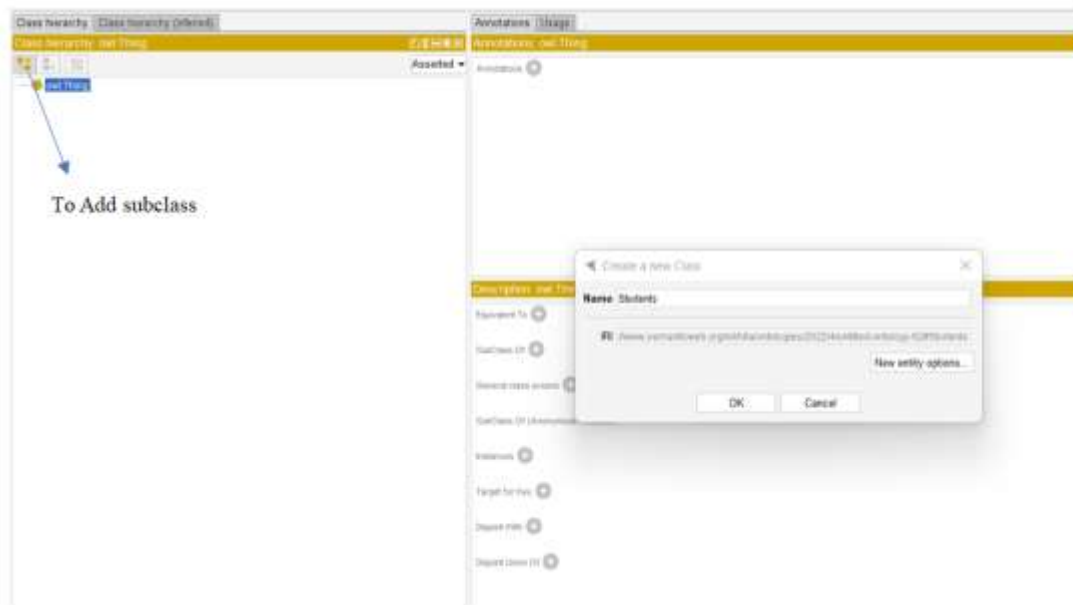
Figure 3. Adding subclass

Adding Sibling to the class and Deletinga class

To add a sibling to any class,select that class and click on the "add sibling class button" as shown in figure 4, and to delete the Class, select the class and click on the "delete selected classes button" as shown in Figure 4.

Figure 4. Adding siblings and deleting the selected class

Creating Class Hierarchy

Creating class hierarchies helps to speed up the process of adding classes.
Steps to create a class hierarchy:
Step 1: Select owl:Thing.
Step 2: Click on Tools and select create class hierarchy as shown in figure 5.

Adding sibling to   For deleting
the class   selected classes

Figure 5. Creating the class hierarchy

Step 3: Now, select the root class owl: Thing -> Continue
Then the Enter hierarchy window will appear where the class ontology should be given as shown in Figure 6.
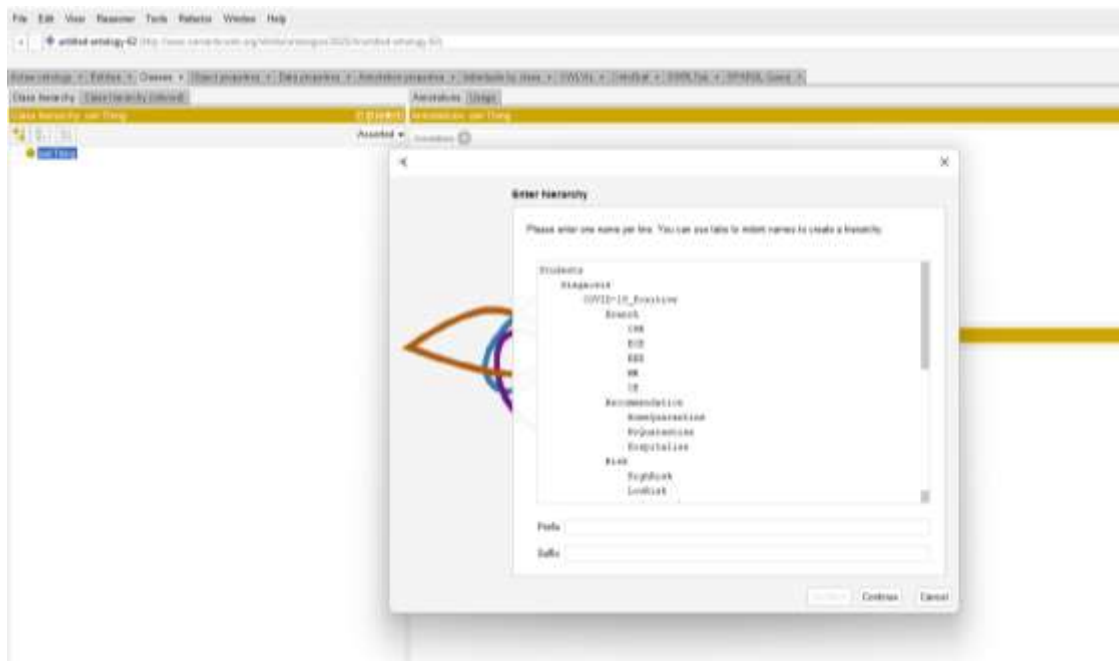
Figure 6. Enter hierarchy and click on Continue

Step 4: After entering the class hierarchy click on continue, then if we want to make the subclasses disjoint, we need to check the box.
Step 5: Click on Finish and the Class hierarchy for Covid-19 ontology is created as shown in the below figure.
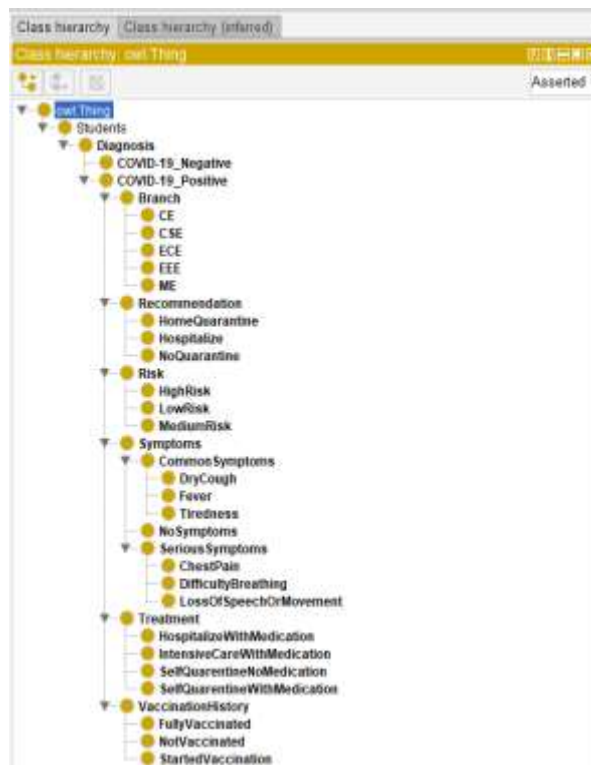
Figure 7. Covid-19 Ontology Class hierarchy.

Description of Class

One of the loaded ontologies' axioms corresponds to each line in a section[8].
The sections are:

  i. **Disjoint Union Of -**Indicates that in a DisjointUnion class axiom, this class is the main class.
 ii. **SubClassOf (Anonymous Ancestor) -** Protege analyses the specified class's predecessors and collects all of their superclasses, which are then displayed in this section.
iii. **DisjointWith -**This class is disjoint with a list of class expressions in each row. There may be two or more classes in a DisjointClasses axiom (the currently selected class is removed from the list for clarity). This section will have a yellow backdrop for inferred disconnected classes. Note that inferred disjoint classes are not displayed by default and will only be visible if the Reasoner Preferences option is activated.
 iv. **Target for Key -**Specifies a set of object and data attributes that serve as a key for instances of the current class. In OWL 2, keys are a new feature that consists of a set of properties. The particular values of 10 of these qualities put together establish distinctness for a certain individual. For example, in a constrained situation, a key consisting of the name and Roll number could be used to infer the uniqueness of individuals in the class Students.
  v. **Instances -**In a class assertion axiom, each row specifies an individual who has the current specified class as its type. This section has a yellow backdrop to show inferred instances. Inferred instances will only be presented if this option in the Reasoner Preferences is enabled.

vi.   **General Class Axioms -**Each row displays a General Class Axiom with the signature of the currently selected class (i.e., mentions the currently selected class). Inferred data is not displayed in this area.

vii.   **SubClassOf -**Each row specifies a class expression that is a subclass of the currently selected class. To put it another way, each row is a superclass of the current class. In this section, inferred superclasses are highlighted in yellow. Inferred instances will only be presented if this option in the Reasoner Preferences is enabled.

viii.   **EquivalentTo –**Each and every row has a class expression that is the same as the current class. This section has a yellow backdrop to indicate inferred equivalent classes. Inferred comparable classes will only be shown if this option in the Reasoner Preferences is enabled.

Description of the class "Students"
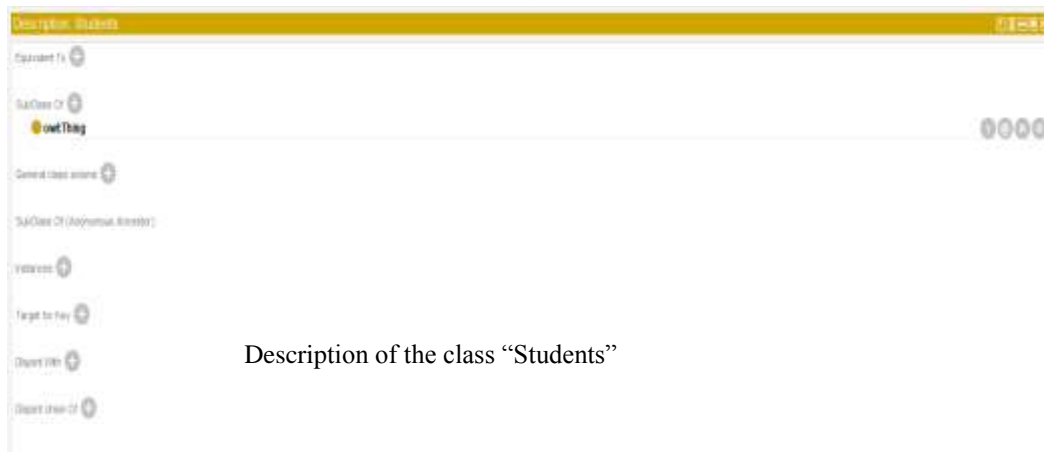
Figure 8. Description of the class

4.2. Owl Properties

OWL Properties [4] represent the relationships that exist between two objects.

There are two main characteristics:

a. Object properties: connect objects
b. datatype properties: attach an object to an rdf: literal or an XML Schema datatype.

Annotation property: Annotation information can be added to classes, individuals, and properties using this OWL attribute.

Object Properties (Domain and Range):

i.   Make sure the "Object Properties" tab is open.
ii.   Window → Tabs → Object Properties.

Data Properties:

i.   Ensure that the "Data Properties" tab is open.
ii.   Window → Tabs → Data Properties.

4.2.1. Object Properties

In OWL all properties are a sub-property of topObjectProperty.
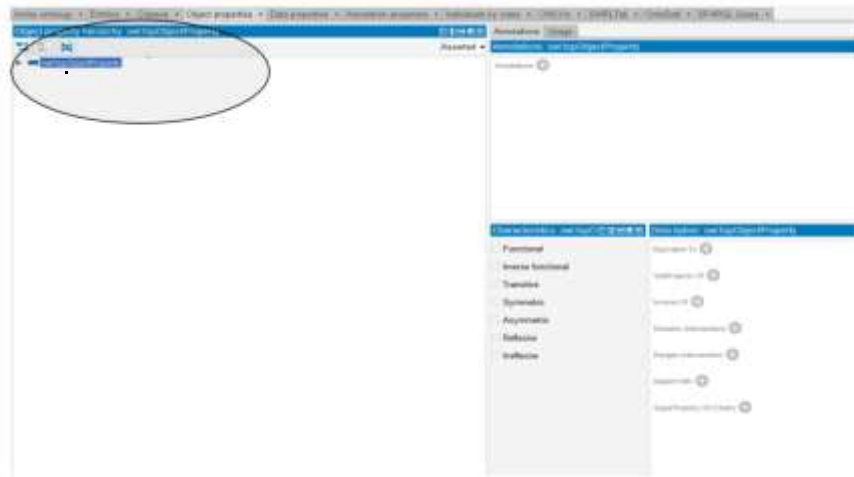
Figure 9. Object Properties super class owl:topObjectProperty

Adding sub-properties

Select the "add sub property button" and name the property has_risk_of.
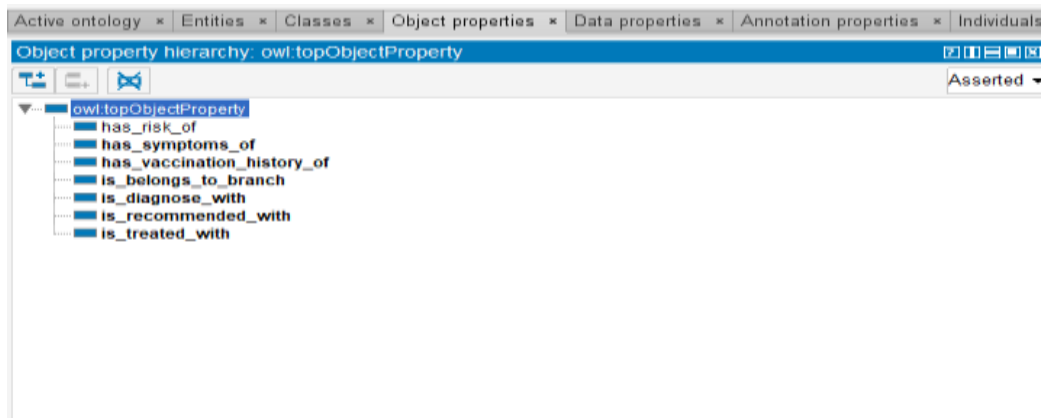Similarly, add all object properties as shown in figure 10.

Figure 10. Object Property Hierarchy

Adding Domains and ranges

Now, for each and all object properties add domains and ranges as shown in the figure below.

Figure 11. Domains and ranges for the object property has_risk_of.

Adding Sibling property and deleting property

To add a sibling property, select the "sibling property button" and to delete the property select the "delete selected properties button" as shown in figure 12.
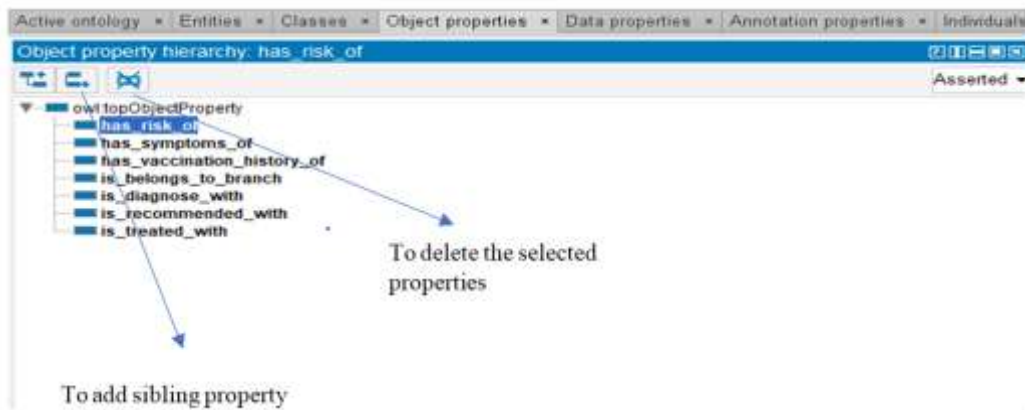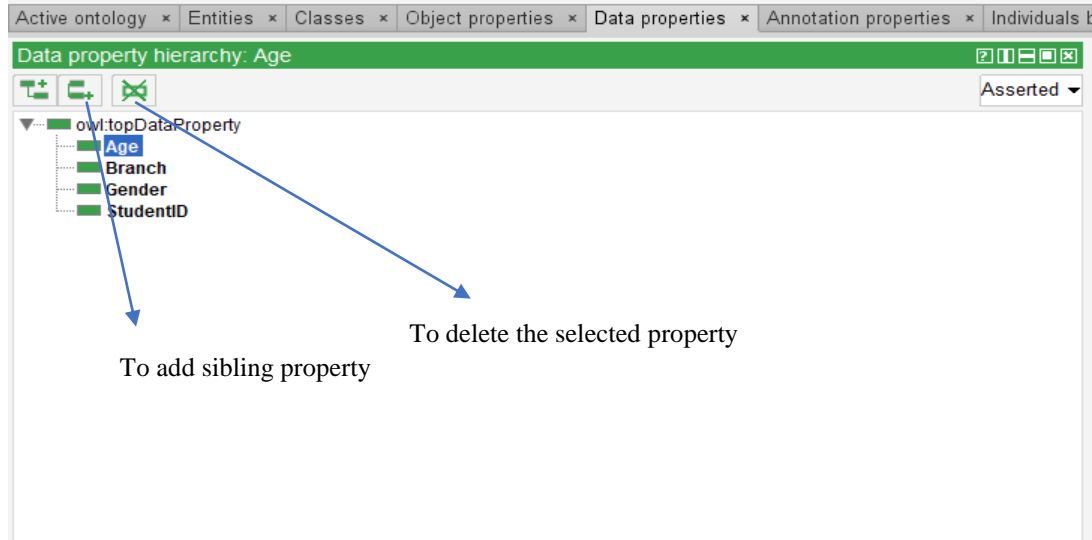

Figure 12. Adding sibling property and deleting selected property.

4.2.2. Data Properties

In OWL all properties are a sub-property of topDataProperty as shown in the figure below.
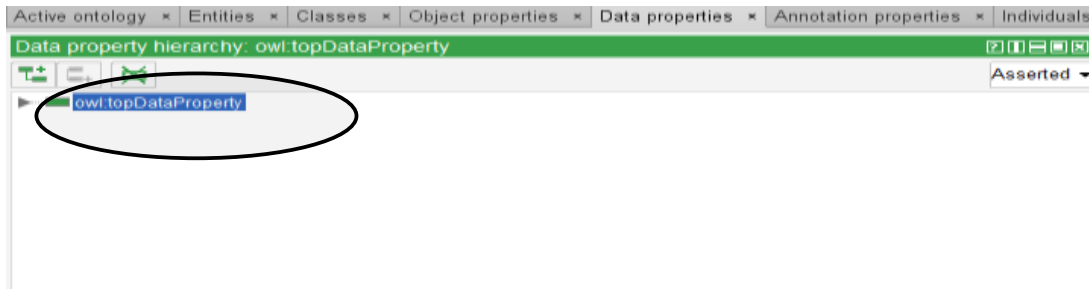
Figure 13. Data Properties superclass owl:topDataProperty

Adding Sub Data Property

Select the "add sub property button" as shown in Figure 14 and name the property StudentId.

Add Sub property
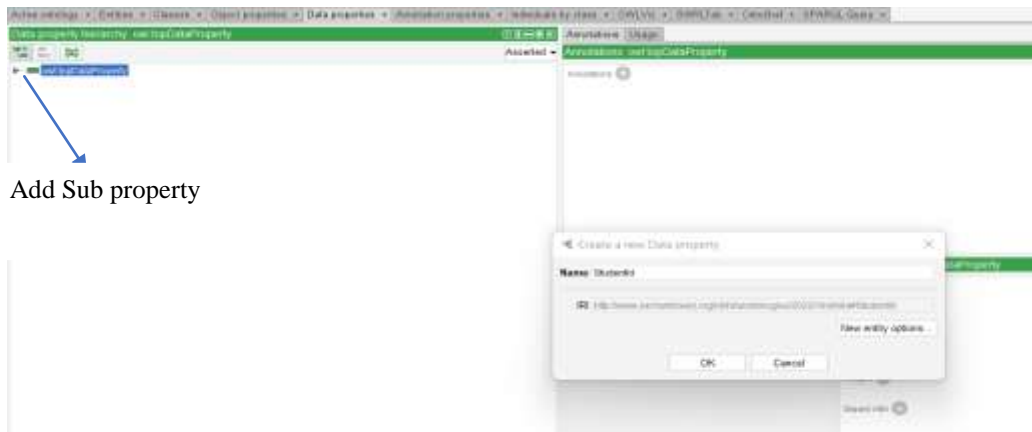
Figure 14. Adding sub-data property.

The Data property hierarchy for covid-19 ontology is shown in the figure below.
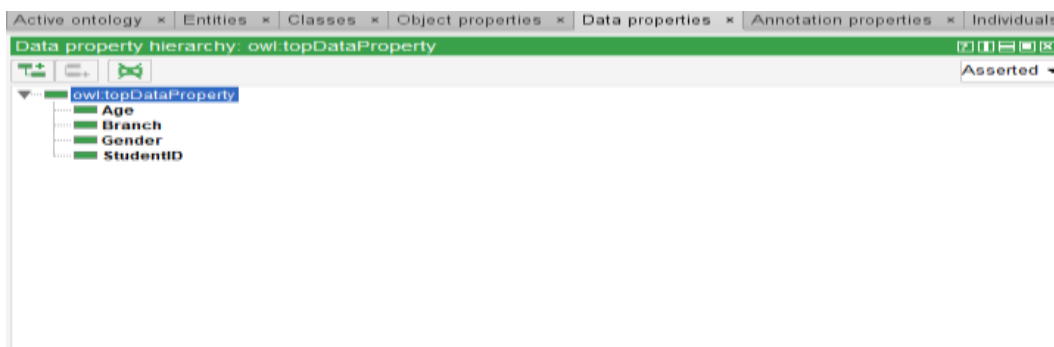
Figure 15. Data Property Hierarchy.

Adding Sibling property and deleting property

To add a sibling property, select the "sibling property button" and to delete the property select the "delete selected properties button" as shown in figure 16.

Figure 16. Adding sibling property and deleting selected property.

Adding Domains and Ranges

To add a domain, select the property and click on the domain(intersection)+ button and select the domain, then click on "OK" as shown in the figure below.
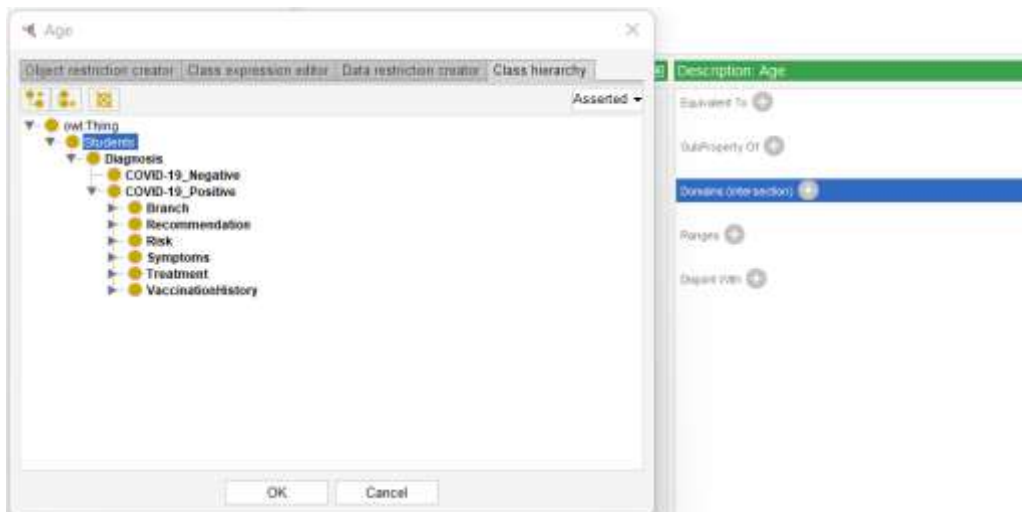
Figure 17. Adding Domain to the selected property.

To add ranges, select the property and click on the ranges + button, and in the "Built-In Datatypes Tab" select the range, then click on "OK" as shown in the figure below.
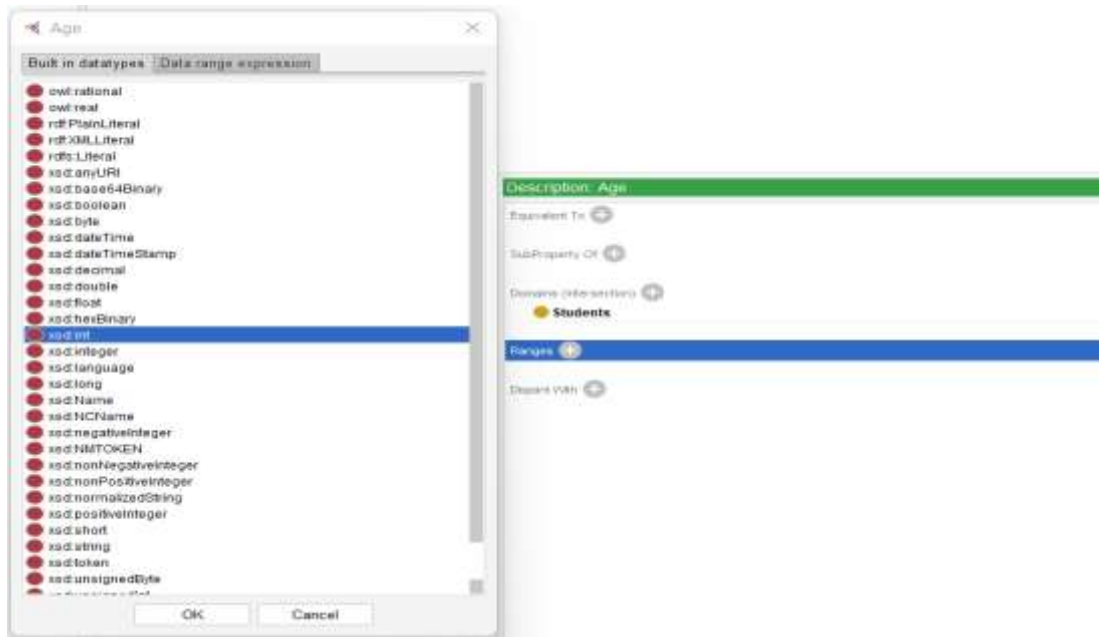
Figure 18. Adding ranges.

4.2.3. Annotation Property

Individuals, class names, property names, and ontology names can all be annotated using annotation properties. Ontology properties connect ontologies to other ontologies and are particularly useful for importing data from other ontologies. Using annotations, the OWL language allows you to comment on any subject that has a URI. Classes, properties, and ontologies can all be annotated. In addition to allowing, you to comment on a subject, the OWL language allows you to define annotation properties in an open-ended manner.
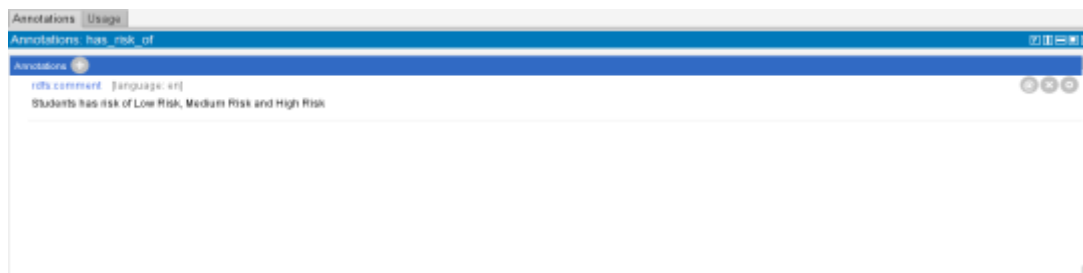
Figure 19. Adding Annotation property to the object property.

4.3. Individuals

Individuals (instances) are an ontology's fundamental, "bottom-up" components. Certain data properties are intentionally omitted for the sake of brevity.

Adding Individuals

The final step in developing the ontology is to include individuals. To do so, go to the "individuals" tab and select the "Add individual" icon.
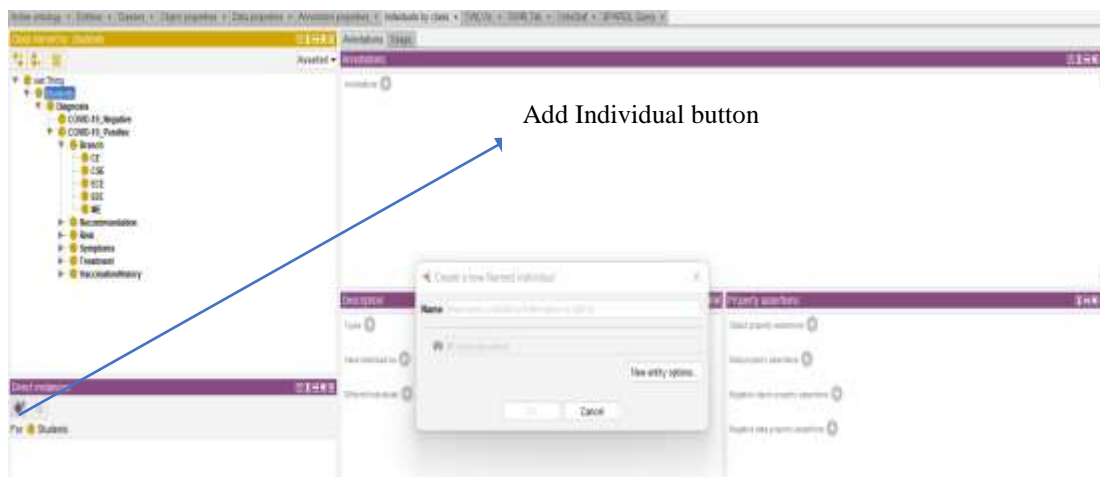
Figure 20. Adding Individuals.

5. ONTOLOGY VISUALIZING TOOLS

There are 3 tools to visualize an ontology in protégé. They are:

   i. VOWL
   ii. OWLViz
   iii. OntoGraf

5.1. VOWL (Visualization of Ontology Web Language)

To visualize the ontology graphically, we must first install the ProtégéVOWL plugin. The jar file can be downloaded from http://vowl.visualdataweb.org/protegevowl.html. After downloading the VOWL jar, place it in Protégé's plugins folder, restart Protégé, and activate the plugin via Window → Tabs → VOWL.

That's all. Simply click the "VOWL" tab in Protégé to see the ontology created in graphical form, as shown below.
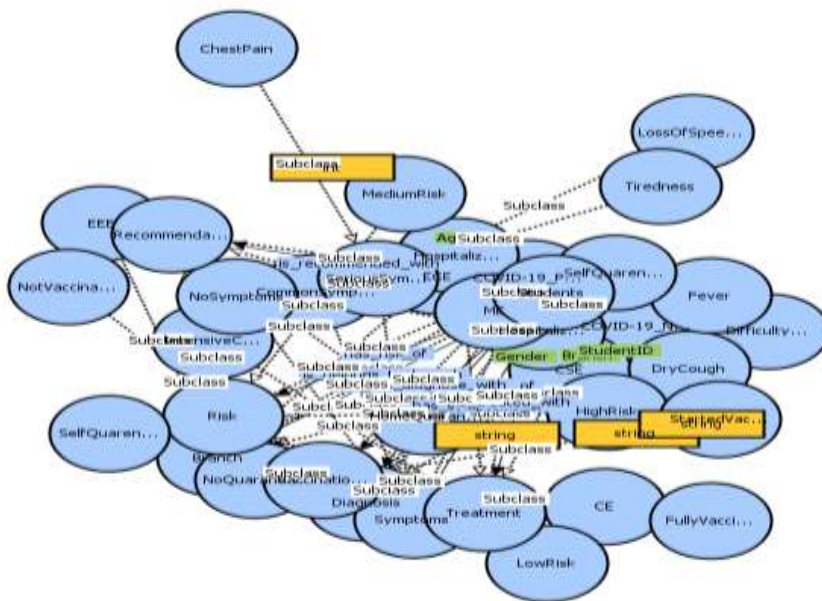
Figure 21. VOWL Representation of Covid-19 Ontology.

5.2. OWLViz

To use OWLViz, you'll need the Protege-OWL editor. It allows you to compare the claimed and inferred class hierarchies in an OWL Ontology by observing and incrementally browsing them. OWLViz operates in tandem with the Protege-OWL editor, using the same color scheme to differentiate between primitive and specified classes, computed class hierarchy alterations, and red inconsistencies. OWLViz can save both the asserted and inferred representations of the class hierarchy to a variety of graphical forms.

The figure shown below is the Inferred hierarchy.

Figure 22. Inferred OWLViz hierarchy.

The figure shown below is the Asserted hierarchy.
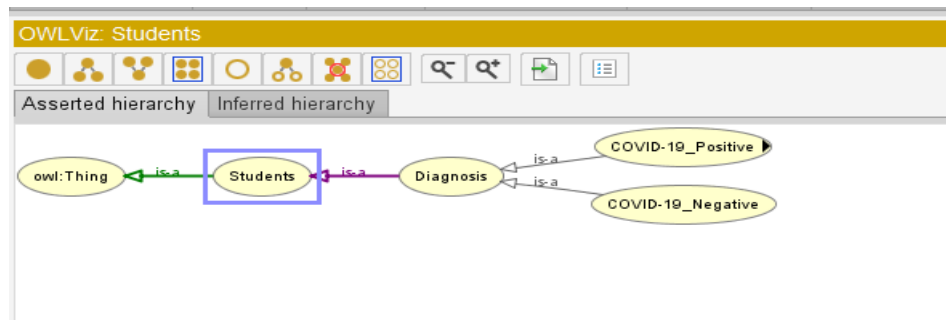
Figure 23. Asserted OWLViz hierarchy.

5.3. OntoGraf

OntoGraf allows you to interactively navigate the relationships in your OWL ontologies. Various layouts are offered for automatically structuring the structure of your ontology. Equivalence, subclasses, individuals, domain/range object characteristics, and subclasses are all supported. To build the desired display, relationships and node types can be filtered.
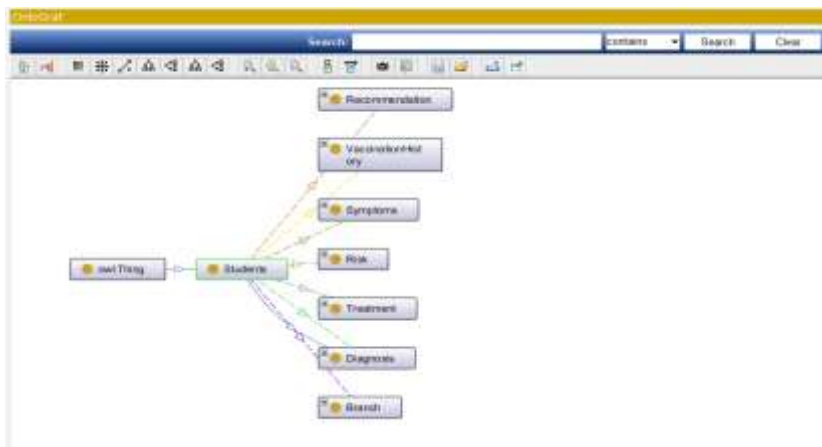

Figure 24. OntoGraf Representation

6. SWRL (SEMANTIC WEB RULE LANGUAGE)

The SWRLTab is a Protégé plugin that allows working with SWRL rules and SQWRL queries in a development environment.

Figure 25. Rules for analyzing risk factors and treatment plan of Covid-19 students.

Based on the above rules the risk factor and the treatment plan for the students is analyzed.

The Drools rule engine is being used.

i.   Select 'OWL+SWRL-> Drools' to send SWRL rules and relevant OWL knowledge to the rule engine.

ii.   To start the rule engine, press the 'Run Drools' button.

iii.   Click the 'Drools->OWL' button to convert the knowledge inferred from the rule engine to OWL knowledge.

iv.   The SWRLAPI supports an OWL profile called OWL 2 RL and performs reasoning using an RL-based reasoner for OWL 2.

See the 'OWL 2 RL' sub-tab for more information on this reasoner.

Figure 26. Drools Rule Engine

7. QUERYING COVID-19 ONTOLOGY USING SPARQL QUERIES

SPARQL is a semantic query language for databases that can obtain and alter data stored in the Resource Description Framework (RDF) format [6].

SPARQL query to know the count of diagnosis is shown in the figure below.

Figure 27. The total count of the student's diagnosis is retrieved.

SPARQL query for analyzing risk factors and treatment plans for the covid-19 student.

```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX eo: <http://www.semanticweb.org/nikhila/ontologies/2022/1/Students#>
SELECT DISTINCT ?Name ?StudentID ?Treatment ?Risk
        WHERE { ?Name eo:is_diagnosed_with eo:COVID-19_Positive.
                ?Name eo:StudentID ?StudentID.
                ?Name eo:is_treated_with ?Treatment.
                ?Name eo:has_risk_of ?Risk.}
```

Figure 28. Analyzing risk factors and treatment plans.

8. OUTPUTS

The result obtained using Covid-19 ontology is shown in the figure below. The ontologyis queried with the help of SPARQL(with the input of diagnosis of students and details of covid-19 diagnosis), student details are extracted and their treatment plan and risk factors were given based on student's vaccination history, symptoms, and diagnosis.

The total number of students count based on diagnosis(COVID-19_Positive and COVID-19_Negative) is retrieved through SPARQL in the figure below.

Figure 29. The result of the total students count based on the diagnosis.

The analysis is done for the students whose diagnosis is COVID-19_Positive and the data is retrieved with the help of SPARQL and SWRL in the figure below.

Figure 30. Result of successful analysis

The total number of students count has a risk of MediumRisk, LowRisk and HighRisk is retrieved with the help of Semantic web rule language and SPARQL based on the symptoms and vaccination history data given by the students in the figure below.

Risk	Total
MediumRisk	"51"^^<http://www.w3.org/2001/XMLSchema#integer>
HighRisk	"11"^^<http://www.w3.org/2001/XMLSchema#integer>
LowRisk	"52"^^<http://www.w3.org/2001/XMLSchema#integer>

Figure 31. The result of the total students count based on Risk factor.

Unlike, many other complex results can be retrieved using SPARQL AND SWRL.

9. CONCLUSION AND FUTURE WORKS

To summarise, we presented the COVID-19 ontology for data collection and analysis, as well as the ontology's classes, objects, and data properties, as well as sample SPARQL queries. Furthermore, we show how to use SWRL rules for an automated treatment plan and risk assessment recommendations for factors. The goal of our research is to gather student data to assist medical practitioners and other researchers in carrying out analyses, forecasting, as well as classifications. In the future, we will intend to create a semantic annotation interface to that student data from Covid-19 can be automatically annotated using the COVID-19 ontology.

REFERENCES

[1] Robert Arp, Barry Smith, and Andrew D. Spear (2015). Building Ontologies with Basic Formal Ontology. London: The MIT Press. p1-245.

[2] Jing Mei, Li Ma, and Yue Pan. (2012). Ontology Query Answering on Databases, p1-14.

[3] Zhijun Zhang and John.A. Miller. (2010). Ontology Query Languages for the Semantic Web a Performance Evaluation, p1- 34.

[4] Çelik Ertuğrul, D. and Çelik Ulusoy, D. (2021). A knowledge-based self – pre-diagnosis system to predict Covid-19 in smartphone users using personal data and observed symptoms. Expert Syst. 2, 1–27. https://doi.org/10.1111/exsy.12716.

[5] Sherimon, V., Sherimon, P.C., Mathew, R., Kumar, S.M., Nair, R. V., Shaikh, K., Al Ghafri, H.K. and Al Shuaily, H.S. (2020). Covid-19 Ontology Engineering-Knowledge Modeling of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Int. J. Adv. Comput. Sci. Appl. 11, 117–123. https://doi.org/10.14569/IJACSA.2020.0111 115.

[6] Visweswaran, S., Ph, D., Samayamuthu, M.J., Morris, M., Weber, G.M., Ph, D., Macfadden, D., Trevvett, P., Klann, J.G., Ph, D., Gainer, V., Murphy, S.N. and Ph, D. (2021). Development of a COVID-19 Application Ontology for the ACT Network.

[7] UganYasavur, Reza Amini, Christine Lisetti and NapthaliRishe. (2013). Ontology-Based Named Entity Recognizer for Behavioral Health. International Florida Artificial Intelligence Research Society Conference, p1- 6.

[8] Akerkar, R. (2009). Foundations of the Semantic Web: XML, RDF & Ontology.

[9] Tabbasum Naz, Maham Shuja, Syed Khuram Shahzad, Muhammad Atif. (2018). Fully automatic OWL generator from RDB schema. 22 (1), 8.

[10] Sunil S Bhopal, S.S., Bagaria, J., Olabi, B., and Bhopal, R. (2021). Children and young people remain at low risk of COVID-19 mortality. Lancet Child Adolesc. Heal. 5, e12–e13. https://doi.org/10.1016/S2352- 4642(21)00066-3.

[11] Göğebakan, K. and Sah, M. (2021). A Review of Recent Advances for Preventing, Diagnosis, and Treatment of Diabetes Mellitus using Semantic Web. International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-6, https://doi.org/10.1109/HORA52670.2021.9461282.

[12] PratikThakora, Dr. SreelaSasib. (2015). Ontology-based Sentiment Analysis Process for Social Media Content. Ontology-based Sentiment Analysis Process for Social Media Content. 53 (1), 7.

AUTHORS

Nikhila Polkampally presently pursuing my Btech in the CSE Department at Geethanjali College of Engineering and Technology.

Rakesh Kumar Donthi presently working as an Associate Professor in the CSE Department for Geethanjali College of Engineering and Technology. I received Ph.D. in CSE from NIT Patna. I did my Btech from Astra, a project in IIIT Hyderabad. My research fields are semantic web, Machine learning, Software Engineering, and Compiler Design.

Dr. G. Soma Sekhar presently working as an Associate Professor in the CSE Department for Geethanjali College of Engineering and Technology. I received Ph.D. in CSE from Acharya Nagarjuna University. My research field is Computer Networks.

Mettu Karuna Sri Reddy presently pursuing my Btech in the CSE Department at Geethanjali College of Engineering and Technology.