Building an Ontology structure in deep learning model of recommender systems

Thi Phuong Trang Nguyen[1,4] , Trong Hai Duong[2] and Luu Thuy Ngan Nguyen [3]

[1]Ho Chi Minh City Open University,
[2]Institute of Science and Technology of Industry 4.0, Nguyen Tat Thanh University HCMC,
[3]University of Information Technology - Viet Nam Nationnal University HCMC,
[4]University of Information Technology - Viet Nam Nationnal University HCMC
trangntp@grad.uit.edu.vn, haiduongtrong@gmail.com,
ngannlt@uit.edu.vn

Abstract. Nowadays, most of developments in deep learning have been proved extremely beneficial for several Medication and Healthcare Tasks. Most of researches in deep learning are focus on figuring out the structured specialities under the data and creating the models to learn how to demonstrate them from raw and unfiltered data. And now, Ontology is a popular vocabulary structure to describe data meaning. In this article, we mention to the possibility of creating Ontology structure of training data in Deep learning Model. More precisely in this paper we summarize the advances in both deep learning and semantic data mining for Health And Medical Sciences Task in recently. We illustrate how learning model with ontology-based deep learning that can extract better data representations to provide more information for the learning process. Using the framework of pattern structures and Formal Concept Analysis algorithms to build a pattern concept lattice in deep learning model. Finally, we present our intentions in building a diagnostic application that uses Ontology-based deep learning.

Keywords: Formal concept analysis, Pattern structures, Ontology

1 Introduction

Today, data is growing rapidly in all areas, with data owner, individuals and businesses easily accomplishing many purposes. Most data exist in natural language. However, natural language is very ambiguous, difficult to define semantics, context of data. Because the data is collected without a specific entry, it leads to have a lot of unrelated data when they are categorized. So the most characteristic of data is not the size, speed or diversity but the value of the data. How to incorporate semantics into a smart data-processing engine to divide data into groups of relevant information is an important research objective.

This article presents the importance of building intelligent learning applications based on the ontology model of training data. Moreover, the paper presents our research

methodology for building a logical structure data in the learning process to exploit information. Analytical technologies and semantic processing technologies will convert vast amounts of data into abstract, meaningful and useful internal data for making a decision. Strict definitions of ontological linkages in ontology and natural language processing will make it easier to integrate data sources. This paper presents a method for automatically converting large amounts of data into abstract, semantic and useful form by expressing them on Ontology to support the development of intelligent data processing applications. It is also the premise for the development of Ontology-based semantic decomposition algorithms for predictive analytics, data-driven decision support.

2　　Researches related to the use of ontology in intelligent learning applications

The success of semantic Webs depends on the ontology, as well as the development of Web pages annotated by metadata that follow these ontologies. Although the benefits that ontologies provide are great, building them automatically is extremely difficult. For this reason, automated information extraction tools for building ontologies such as identity recognition systems are essential.

Ontology is a common term in semantic web models or semantic processing in data [3]. It is widely used in many fields such as: knowledge technology [14], data design and integration [5,19], information extraction, search (Yahoo and Lycos), E-commercial (Amazon and eBay), configuration (Dell and PC-order), and Smart Government Applications (DARPAs High Performance Knowledge Base). Ontology is the structure of the links between objects, attributes, events, processes in a scientific way as they are in the real world. In the information system, ontology is the representation of the fields that already exist in reality that: 1) Reflect the properties of the object as in the real world 2) can be aware of the experts in that fields 3) Formed by mechanisms that support automatic processing. Ontology facilitates the exchange of data between different sources. Recently, it has attracted considerable attention from researchers. Most of the research focuses on the language in Ontology, the inference mechanism, the Ontology editing environment, and integration [19]. In particular, the integration of ontologies is a complex task because the characteristics are naturally formed, so the language, field, and structure of different ontologies are often not the same. [8,9].

Deep learning (or Structured learning, hierarchical learning or deep machine learning) is a branch of machine learning based on a set of algorithms for finding a high-level abstraction model in data by using a the hierarchical chart which has many processing layers, including many linear and non-linear transformations. Deep Learning is part of an extended set of machine learning algorithms based on how data is represented. An image can be represented in many forms as a vector of values or in more abstract forms such as set of edges, specific image areas, etc. Some expressing ways better than the others when simplify the learning process of the computer (eg, face recognition or facial expressions [13]). One of the best benefits of deep learning is that it can learn features from original data. Consequently, there is no need for specific extraction algorithms prior to the implementation of supervised and semi-supervised data

classification steps [10]. Researches in this field are exploring better ways of expressing and creating models for learning these expressions from unclassified raw data sources. Neuron model as the way that human brain work - is simulated to describe the representation of features during deep learning [20].

Researchers are now focusing to building data-mining algorithms on Ontology models of such data. The article of Ontology-based Deep Learning for Human Behavior Prediction in Health Social Networks [21] introduces a deep learning model - Retristive Boltzmann Machine based on the Ontology. Observing a number of chronic diseases such as diabetes, cardiovascular disease and cancer, the author finds that one important factor that causes them is obesity. This has been a major challenge for countries with high per capita incomes, which are now a growing challenge in low- and middle-income countries. Recent researches have shown that obesity can be transmitted through social networks [6] but too few researches prove this [23]. The authors suggest using the internet and mobile devices to monitor and collect health information including distance, movement speed, running. Using these technologies, they tracked daily workout information, social interaction (texting, gaming, event participation, sports competitions, etc.) of a group of 254 people [22]. Based on that data, the authors propose a deep learning structure based on Ontology to predict the tendency of an individual to act. This model is an extension of the Retristed Boltzmann Machines model [25] by integrating Ontology into [14]. While other deep learning models such as RBMs, Convolutional Neuron Network [17], and Sum Product Networks [16] treat the disease symptoms of a patient as they were not related to each another but in fact they have strong connections in biomedical sciences. Therefore, a better model must learn these connections during the predictive model training process. To solve this, the authors suggest learning from the bottom up in the structure of Ontology. The main idea of the algorithm is that each concept will be learned from its attribute structures, related concepts and subordinate concepts because the Ontology structure clearly defines the close relationship between the attributes and concepts.

Ontology in the semantic web is a common lexical structure for describing the semantics of data. This network includes a collection of concepts in different areas, attributes and interactions that are prevalent among them. Ontology describes individuals, classes, attributes, and relationships [4]. In order to help the computer read and use content from such semantic web pages, the "An Ontology Based Deep Mining Methodology to Cluster The Content From Web Servers" is proposed to represent the idea of exploiting content based on Ontology structure. The ODMM concept - Semantic Web - is a new developing step in web technology. It makes the data easy to search, extract, display, translate. Character processing methods are proposed to solve the problem of document classification. The benefits can be easily realized by the fact that this method extracts data sets that contain synonyms, more general meaning words and homonyms. Moreover, it also allows to sort for the repeated data items or discrete data items.

The paper "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis" [27] is not only focus on solving problem of characteristic showing but also integrate information from many different sources. In experiments, the author extract characteristic of disease from two separate sources: Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET). Normally, the

other ones use separate algorithms to extract characteristic before classify them. Sometime, they combine some attribute vectors by link them together to generate a longer attribute vector. However, we propose a new method base on Deep Learning, we use "Deep Boltzmann Machine" to find the characteristic which have invisible structure in data, and offer 1 solution to synthetic them from two sources base on "Deep Boltzmann Machine" model.

There are two problems need to be focused on this paper: show characteristic and integrate data from many different sources. About characteristic showing, we use approach per small data method because it's intermediate solution between pixel approach and Region of Interest approach. This method could solve the problem with the vectors which have many attributes effectively, and we could easily recognize the change of characteristic although it's very small. About data integrating, we give it to clinical medicine view. Neuroscientists and X-ray researchers verify the brain's images by finding split fields and connect them with their vicinity , till the end is all of brain. The researchers focus on clear sample, generalize relative information from whole brain and provides clinical diagnosis. This method could extract more information and improve the accuracy of the diagnosis. With the above-mentioned researches, we realize that the using semantic technology in the development of intelligent data applications is very feasible and necessary.

Identify the value of building an Ontology model for date learning in smart learning applications. In this newspaper, we propose an automatic method to build Ontology model: OCA (Ontological Concept Analysis) for data learning in intelligent machine learning applications to solve the issues related to size and diversity of data.

3   The algorithm to build a ontology structure for data training (OCA)

Formal Concept Analysis (FCA) is a mathematical theory which is invented and developed by Rudolf Wille, used to analyze data via formal contexts and concept lattices [3]. It is widely successfully applied to lots of fields such as medicine and psychology, musicology, linguistic databases, library and information science, software re-engineering, etc… and to a variety of methods for data analysis, information retrieval, and knowledge discovery in databases. One of the significant features of FCA is the graphical visualizations of inherent data structure. From a binary table, FCA builds a concept lattice to describe all connection among objects and theirs attributes. FCA is based on the philosophical understanding of the world in terms of objects and attributes. It is assumed that a relation existed between object and attribute expresses a meaning.

In fact, the concepts could be found in documents at different levels depending on the transparency of the type of document being considered, for example, some documents contain clear concepts in the form of definitions such as "a tiger is a mammal" or "mammals such as tigers, lions or elephants". And we must re-search the model to find the classification or relationship of concepts in the documents. And the problem is how we could find them. There is 1 solution is we should find how the concepts from the document are analyzed and used more than find their clear definitions. And FCA could help to solve it.

However, FCA has not cared about the relationship between objects. In 1 ontology, the relationship between objects is mandatory. So we propose upgrade FCA to new level: "Ontological Concept Analysis" (OCA) to build an Ontology structure use for Deep Learning application.

3.1 FCA (Formal Concepts Analysis)

Definition FCA. FCA starts with a formal context and builds a set of formal concepts orrganized within a concept lattice. A formal context $C = (G, M, I)$ is a representation of correlation between set data object G, set data attribute M and the incidence relation of the context. And then formal concept will be clarified after understanding the derivative operator.

In Table 1, a cross table for a formal context is shown. The rows are the set of objects G, the columnts are the set of attibutes M, and every cross represents an element of the relation I. The triplet (lion, 4legs, Relation I) forms a formal context.

Table 1. Example about a formal context

	Small	Big	2legs	**4legs**	Feath-ers	Hair	Fly	Hunt	Run	Swim
Dove	x		x		x		x			
Hen	x		x		x					
Duck	x		x		x		x			x
Goose	x		x		x		x			x
Lion		x		**x**		x		x	x	

Suppose that the correlation of object $g \in G$ to the attribute $m \in M$ is written as gIm or $(g, m) \in I$ and read that object g has the attribute m. For a set $A \subset G$, A' is defined as the set of attributes common to the objects in A, written $\{m \in M \mid gIm$ for all $g \in A\}$. Similarly, for an attribute set B we define B' as the set of objects which have all attributes in B, written $\{g \in G \mid gIm$ for all $m \in B\}$. As a result, a formal concept of the context (G, M, I) is a pair (A, B) with $A \subseteq G, B \subseteq M, A' = B$ and $B' = A$. In a formal concept (A, B): A is the extent – set of objects - and B is the intent – set of attributes. Ţ(G, M, I) denoting the set of all concepts of context (G, M, I) is considered as the concept lattice of whole data. [4]. An example of formal concept for the context in table1

Pick a set of objects: A ={duck}
Derive attribute: A' = {small, 2legs, feathers, fly, swim}
Derive objects: (A')' = {small, 2legs, feathers, fly, swim}' = {duck, goose}
Formal concept: (A'',A')=({duck, goose},{small, 2legs, feathers, fly, swim}).
Figure 1 shows a lattice for the context in Table 1.

3.2 Pattern concept structure

Pattern structures are an extension of FCA for mining complex data [1]. Pattern structure is defined by a set of objects, a set of descriptions associated with the set of objects,

and a similarity operation on descriptions, matching a pair of descriptions to their common part.
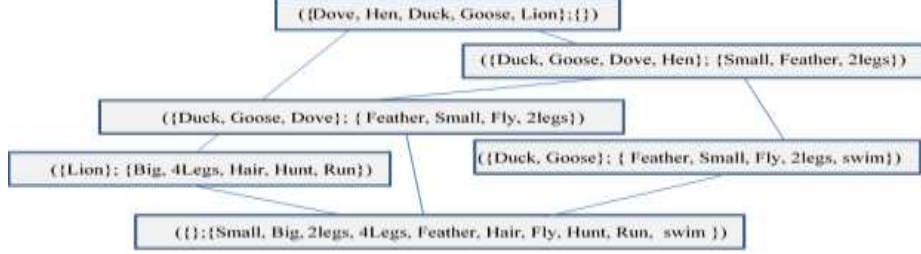

Fig. 1. An exaple a lattice for the context in Table 1

We want to make a basic ontology structure but FCA only make a concept lattice structure, so we combine pattern structure in [1] to FCA for build ontology structure.

Follow [1] A Pattern structure Þ is a triple (G, (D,⋒),ζ), where G, D are the set of objects and the set of descriptions, and ζ: G → D maps an object to a description, (D, ⋒) is a complete meet-semilattice on D.

Following The Galois, relating sets of objects and descriptions in a Pattern structure (G, (D,⊓),ζ) is:

$$A' := \prod_{o \in A} \zeta(o) \qquad for\ A \subseteq G$$

$$d' := \{o \in G \mid d \subsetneq \zeta(o) \quad for\ d \in D$$

With the meaning: Giving a subset of objects A ($\subseteq G$), A' returns the description that it is common for all objects in A. Giving a description d, d' is the set of all objects which have all description in D follow operation $\subsetneq$: $x \subsetneq y \Leftrightarrow x \cap y = x$.

A pattern concept of Þ(G, (D,⋒),ζ), is a pair (A, d) where A ⊆ G and d ∈ D, A' = d and d' = A, A is called the concept pattern extent and d is called the pattern intent.

3.3 Ontological Concept Analysis

We combine FCA and Pattern structure for build ontology structure. We need to define identities for formal concepts, so we will add semantic descriptors to the creation of concept lattices.

In this case, a OCA model (A pattern concept) is a pair (G,D) where G is the set of object, include two part, the first part is object (O) and identification (I), so G= O+I, and the second part is the descriptions set (D).

Build concept lattices from the formal concept. This a very naive and intuitive algorithm to build formal concept from an object or attribute until the whole set is scanned. The aggregation of these two approaches generate all possible formal concepts from binary table.

From Object:

```
For each object A in G
  Compute derivative of set A to get the attribute set B
  = A'
  Compute derivative one more time of A' to obtain all
  objects having those attributes C = (A')'
  So The formal concept is the pair (C, B) or (A'', A')
  If it has not t∈ I in A', the formal concept name is FCᵢ
  Else the formal concept name is FAₜ
End
```

From attribute:

```
For each object B in D
  Compute derivative of set B to get the object set A =
  B'.
  Compute derivative one more time of B' to obtain all
  attributes containing in those objects C = (B')'
  So The formal concept is the pair (A, C) or (B'', B')
  If it has not t∈ I in B', the formal concept name is FCᵢ
  Else the formal concept name is FAₜ
End
```

The upper concept is called superconcept and the lower one is subconcept. Superconcept has more extent and less intent than subconcept. The top node of the lattice corresponds to the concept of everything and the bottom corresponds to the concept of nothing.

Table 3 shows a set of formal concepts for the context in Table 2, for simplicity some intents are not given.

Table 2. A formal context in OCA model

	m_1	m_2	m_3	m_4
$o_1 ⋊ i_1$	x			x
$o_2 ⋊ i_2$		x		
$o_3 ⋊ i_3$	x		x	
$o_4 ⋊ i_4$	x		x	x
$o_5 ⋊ i_5$	x	x	x	

Table 3. A set generate formal concepts in Table 2 following that Intersection method

Generate Formal concepts	
Fi_5	$\{(o_5 ⋊ i_5),(m_1,m_2,m_3)\}$
$Fi_3i_4i_5$	$\{(o_3 ⋊ i_3,o_4 ⋊ i_4,o_5 ⋊ i_5),(m_1,m_3)\}$
Fi_1i_4	$\{(o_1 ⋊ i_1,o_4 ⋊ i_4),(m_3,m_4)\}$
Fi_4	$\{(o_4 ⋊ i_4),(m_1,m_3,m_4)\}$
$Fi_1i_3i_4 i_5$	$\{(o_1 ⋊ i_1,o_3 ⋊ i_3,o_4 ⋊ i_4,o_5 ⋊ i_5),(m_1)\}$
Fi_2i_5	$\{(o_2 ⋊ i_2,o_5 ⋊ i_5),(m_2)\}$

These two approaches are inefficient since there are too many concepts are generated multiple times. So we must delete some duplicate concepts. Following it in Intersecion method below:

Intersection Method

```
For each attribute m∈M, compute the attribute extent {m}'
```

```
For any two sets in this list, compute their intersec-
tion.
  If it is not yet contained in the list, add it;
  otherwise reject that intersection. The two sets also
contain the case of one set and itself. Each two sets
generate a set of intersection objects
Repeat until no new extents are generated.
If G is not yet contained in the list, add it. G is the
top node or the concept of everything
For every extent intersection set A in the list, compute
the corresponding intent A'.
```

Table 4 shows a set generate formal concepts in Table 2 following that Intersection method

Table 4. A formal concepts for the context in table 2 in OCA model

Formal concepts From object		Formal concepts From attribute	
$o_1 \rtimes i_1$	$\{(o_1 \rtimes i_1, o_4 \rtimes i_4),(m_1,m_4)\}$	m_1	$\{(o_1 \rtimes i_1, o_3 \rtimes i_3, o_4 \rtimes i_4, o_5 \rtimes i_5),(m_1)\}$
$o_2 \rtimes i_2$	$\{(o_2 \rtimes i_2, o_5 \rtimes i_5),(m_2)\}$	m_2	$\{(o_2 \rtimes i_2, o_5 \rtimes i_5),(m_2)\}$
$o_3 \rtimes i_3$	$\{(o_3 \rtimes i_3, o_4 \rtimes i_4, o_5 \rtimes i_5),(m_1,m_3)\}$	m_3	$\{(o_3 \rtimes i_3, o_4 \rtimes i_4, o_5 \rtimes i_5),(m_1,m_3)\}$
$o_4 \rtimes i_4$	$\{(o_4 \rtimes i_4),(m_1,m_3,m_4)\}$	m_4	$\{(o_1 \rtimes i_1, o_4 \rtimes i_4),(m_3,m_4)\}$
$o_5 \rtimes i_5$	$\{(o_5 \rtimes i_5),(m_1,m_2,m_3)\}$		

Draw the concept lattice. For Draw the concept, We follow step of drawing Hasse diagram. Below is the pseudo code of this procedure

```
Step 1: Draw a small circle for the extent G at the top -
the concept of everything. G is the node containing all
objects in data table.
Step 2: From G, draw to its sub-extent which has most el-
ements. Repeat this step to the least element extent in
the FC set.
  Sub-step2: Every attribute is written slightly above
  the circle of its attribute extent and every object is
  written slightly below the circle that is exactly below
  the circles that are labeled with the attributes of the
  object.
Step 3: If there is a path from bottom node to G for an
extent, remove the upper extent.
Step 4: If there is a path from bottom attribute to G for
an intent, remove the lower intent.
```

```
Step 5: If the circles has no name, set its name by its
top attribute.
```

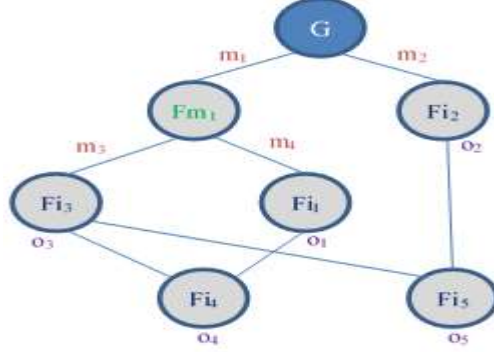The final concept latice in Table 2 shows in Figure 2

Fig. 2. The concept latic for the context in table 2

In the produced concept lattice, every extent has one/multi path(s) from its position to G containing attributes as described in table.

Concept scaling. In reality, obtained data does not only contain one value in each attribute as Yes/No but also a range of continuous values such as mark of a student [0-10], series of color [grey, black, white, blue…] … Thus, we need to scale these attribute values into binary forms which is called scaling transformation and was developed by Ganter and Wille [12]. Scaling transformation is a process of converting data tables with general attributes, e.g. nominal, ordinal, etc., to data tables with yes/no attributes. For instance, attribute mark with range from [0-10] could be replaced by four yes/no attributes: mark-weak, mark-average, mark-good, mark-excellent corresponding to mark intervals [0- 4], [5- 6], [7-8], [9-10]. The derived context is obtained by a replacement. Each many valued attribute is substituted by the scale attribute in the pre-defined range, then the new attribute is marked X at the conresponding data row. If the attribute in a many-valued context has two negated values as TRUE/FALSE, MALE/FEMALE, we do not need to either name or define a range for new attributes. (see more in table 5)

Table 5. Example changed many valued context to one valued context

	Mark
Student 01	3
Student 02	5
Student 03	8

	Mark-Weak	**Mark-Average**	**Mark-Good**	**Mark-Ex-cellent**
Student 01	x			
Student 02		x		
Student 03			x	

The process of building the ontology structure (OCA) is presented in Figure 3.
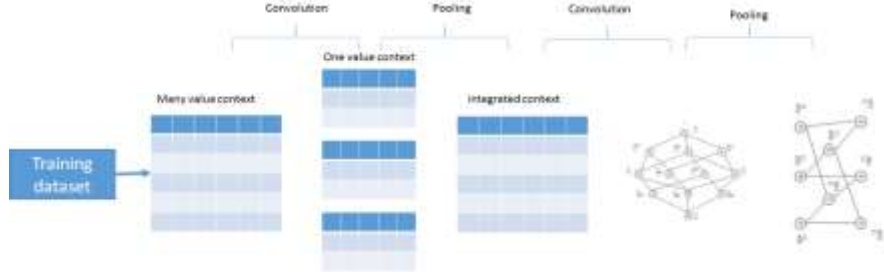

Fig. 3. The process of building the ontology structure (OCA)

4   Experiments

This article presents the diagnostic system for patients with acute cerebral infarction at the People's Hospital 115, Ho Chi Minh City, Vietnam.

System data includes: medical records of 100 patients older than 60 years of age with cerebral infarction. Data is collected from July 2017 to December 2018. Cerebral infarction patients are divided into 9 groups acccorrding to the 10th International Classification of Diseases (ICD10).

Our experiments method is as follows: We divide the data of 100 patients into two parts, the first part (training) includes medical records of 90 patients, part 2 (predicted) data of 10 patients. Based on the training data, we used the prediction methods on 10 patients in two tasks. Specifically, We use symptom list to diagnose the prognosis of disease incidence (List of diseases likely to occur with risk rates recommended by the doctor). To accomplish the above process, we create a deep Ontology-based learning structure of disease data and associated pathology records to predict an individual's action tendency. We reuse the ORBM model from the article [18]. This model is the result of the extension of the Retristed Boltzmann Machines model by integrating Ontology into the data training process. In the experimental scope of the article, we present only the results of automatic creating ontology model process of disease data for the above application and achived results from two predictions. We did not describe training process settings in detail, in this article, for training process; we are re-using the method from the article [18]. We only changed the ontology model (SMASH Ontology) used by the article, instead of the fixed model, we will generate the ontology model during the training process of the data.

In this case, a OCA is a set of tuples with three elements: i) patient ID (i.g P001), ii) The disease ID (i.g Cerebral infarction), iii) Set of disease symptoms (i.g large atherosclerosis).
An example of a inputted data is given below:

$$\left\langle \begin{array}{c} [(P001); \ (\text{``Nhồi máu não khác''})] \ ; \ [(\text{"Có tiền sử rung nhĩ"}); \\ (\text{"TOAST}: xơ \ vữa \ mạch \ máu \ lớn\text{"})] \end{array} \right\rangle$$

This input represents a patient data with three part: patient P001 suffering from cerebral infarction due to large Atherosclerosis has the following symptoms: medical history Atrial fibrillation, the classification of TOAST disease is large atherosclerosis. In order to standardize the data, we modeled the three components according to the following symbols: Patients are coded according to patient ID, the name of the disease are coded acccorrding to the 10th International Classification of Diseases (ICD10) (see more in table 6), and the symptom are encoded follow the table 7. So the example is transformed like this: $\langle[(P001);\ (I63.8)]\ ;[m7;m10.1]\rangle$

Table 6. The disease set (name and code)

Disease (Vietnamese name)	ICD10 code
Nhồi máu não (Cerebral infarction)	I63
Nhồi máu não do huyết khối động mạch vành	I63.0
Nhồi máu não do tắc mạch động mạch vành	I63.1
Nhồi máu não do tắc nghẽn không đều hoặc hẹp động mạch thận	I63.2
Nhồi máu não do huyết khối động mạch não	I63.3
Nhồi máu não do tắc mạch động mạch não	I63.4
Nhồi máu não gây ra bởi tắc nghẽn hoặc co thắt không xác định của động mạch não	I63.5
Nhồi máu não do huyết khối mạch máu não, không gây ngộ độc	I63.6
Nhồi máu não khác	I63.8
Nhồi máu não, không xác định	I63.9

Some variables (attributes), we initially collected from the training data are described in Table 7, which consists of 14 variables.

Table 7. The symptom set

Attributes name	Code	Attributes name	Code
Nhóm tuổi[1]* (*Age*)	m1	Điểm NIHSS* (*NIHSS index*)	m8
Giới tính (*Gender*)	m2	XHN có triệu chứng (*XHN has symptoms*)	m9
Tiền sử hút thuốc lá (*History of smoking*)	m3	NMN theo TOAST* (*NMN by TOAST*)	m10
Tiền sử tăng huyết áp (*History of hypertension*)	m4	HATT lúc điều trị* (*HATT at treatment*)	m11
Tiền sử đái tháo đường (*History of diabetes mellitus*)	m5	HATTr lúc điều trị* (*HATTr at treatment*)	m12
Tiền sử rối loạn lipid máu (*History of dyslipidemia*)	m6	Glucose máu* (*HATTr at treatment*)	m13
Tiền sử rung nhĩ (*History of atrial fibrillation*)	m7	Liều thuốc TSH* (*Dose of TSH*)	m14

---

* multi-valued attributes

Table 8 describes a part of the monotonous relationship table extracted according to the attributes listed in Table 7. The algorithm then generates a conceptual network model for the training process.

In this section, we describe the diagnostic application for patients undergoing treatment at the 115 People's Hospital, Vietnam. Based on the medical records of the group of patients being treated at the hospital, we built the application to support the disease prediction. From the training data, we have created an ontology structure on cerebral infarction. We compared the times and the accuracy rate of building ontology disease structure by OCA with CoO and DO methods in [15,28]. Figure 4 shows that the concepts for the ontology of Cerebral infarction were generated achieving 80% accuracy over the results obtained by the two methods (the remaining 20% belonged to the concepts that were not set name), but with the time, our modeling time is three times faster than the two methods. With the unnamed concepts, we will modify them in the time when the deep learning model active.

Table 8. A formal context of training data

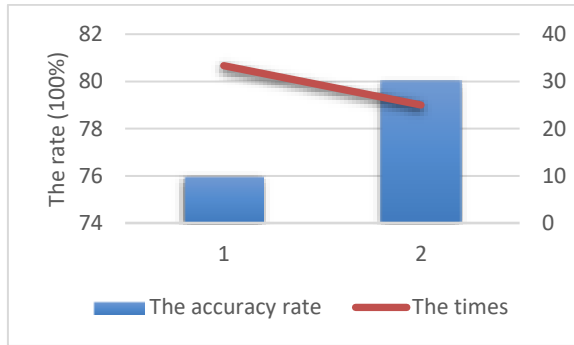| 0	I	m1.1	m1.2	m1.3	m2	m3	m4	m5	m6	m7	m8.1	m8.2	m9	m10.1	m10.2	m10.3	m10.4	m11.1	m11.2	m12.1	m12.2	m13.1	m13.2	m14.1	m14.2
P001	I63.0		x						x		x		x	x				x		x		x			x
P002	I63.3	x			x		x		x			x			x			x	x	x		x			x
P003	I63.1			x		x	x		x		x		x			x		x		x		x		x	
P004	I63.4	x			x		x		x			x			x			x		x		x		x	
P005	I63.3	x				x	x		x		x	x						x		x		x			x

Fig. 4. Comparison of the times and the accuracy rate of our ontology structure building with CoO and DO

5 Conclusion

In this paper, we have presented a novel approach for analyzing data for make ontology structure that using it in deeplearning model within the framefork of Fomal Concept Analysis and pattern structures.

With common machine learning applications, we need to develop data-based algorithms that thoroughly exploit the information of the data. By creating a hierarchical model of data structures, we can take advantage of the semantic relationship of data to extract more information.

In the ontology structure created from our method, the generated concept names are simply named, although there are still some unname concepts. But with the deep learning model, it still work the best. In its process, some unname concepts can be named based on input data relationships.With the Ontology structure is generated by our method, the Deep Learning model will work better, because it could extract more semantic information easily. In the next paper, we will show the result when integrating the above Ontology model in data learning in Deep Learning model.

References

1. Aleksey Buzmakov, Elias Egho, Nicolas Jay, Sergei O. Kuznetsov, Amedeo Napoli & Chedy Raïssi, "On mining complex sequential data by means of FCA and pattern structures" International Journal of General Systems Vol. 45 , Iss. 2,2016, p. 135-159
2. Baker, K.S. and C.L. Chandler, "Enabling long-term oceanographic research: Changing data practices, informationmanagement strategies and informatics", Deep Sea Research Part II: Topical Studies in Oceanography, 2008. 55(18): p. 2132-2142.
3. Berners-Lee, T., J. Hendler, and O. Lassila, "The semantic web", Scientific american, 2001. 284(5): p. 28-37.
4. Cai, M., W. Zhang, and K. Zhang, ManuHub: a semantic web system for ontology-based service management in distributed manufacturing environments. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2011. 41(3): p. 574-582.
5. Castillo, J.A.R., et al. "Information extraction and integration from heterogeneous, distributed, autonomous information sources-a federated ontology-driven query-centric approach", in Information Reuse and Integration, 2003. IRI 2003. IEEE International Conference on. 2003. IEEE.
6. Christakis, N.A. and J.H. Fowler, "The spread of obesity in a large social network over 32 years", New England journal of medicine, 2007. 357(4): p. 370-379.
7. Dmitry I. Ignatov, "Introduction to Formal Concept Analysis and Its Applications in Information Retrieval and Related Fields", In RuSSIR 2014, Nizhniy Novgorod, Russia, CCIS vol. 505, Springer 42-141
8. Duong, T.H., N.T. Nguyen, and G.S. Jo, "A hybrid method for integrating multiple ontologies", Cybernetics and Systems: An International Journal, 2009. 40(2): p. 123-145.
9. Duong, T.H., et al., "Complexity Analysis of Ontology Integration Methodologies: A Comparative Study", J. UCS, 2009. 15(4): p. 877-897.
10. Fiorini, R.A. "A Cybernetics Update for Competitive Deep Learning System", in 2nd International Electronic Conference on Entropy and Its Applications. 2015. Multidisciplinary Digital Publishing Institute.
11. Ganter, B., G. Stumme, and R. Wille, "Formal concept analysis: Methods and applications in computer science", TU Dresden, http://www. aifb. uni-karlsruhe. de/WBS/gst/FBA03. shtml, 2002
12. Ganter, B. and R. Wille, "Conceptual scaling, in Applications of combinatorics and graph theory to the biological and social sciences", 1989, Springer. p. 139-167

13. Glauner, P.O., "Deep Convolutional Neural Networks for Smile Recognition", arXiv preprint arXiv:1508.06535, 2015.

14. Gruber, T.R., "A translation approach to portable ontology specifications", Knowledge acquisition, 1993. 5(2): p. 199-220.

15. Hong Son Nguyen, Minh Hieu Le, Chan Quan Loi Lam, Trong Hai Duong: "Smart interactive search for Vietnamese disease by using data mining-based ontology". J. Information Telecommunication 1(2): 176-191 (2017)

16. Huynh, A.L., H.S. Nguyen, and T.H. Duong, "Triple Extraction Using Lexical Pattern-based Syntax Model", in Advanced Computational Methods for Knowledge Engineering. 2016, Springer. p. 265-279.

17. LeCun, Y., et al., "Gradient-based learning applied to document recognition", Proceedings of the IEEE, 1998. 86(11): p. 2278-2324.

18. NhatHai Phan, Dejing Dou, Hao Wang, David Kil, Brigitte Piniewski: "Ontology-based deep learning for human behavior prediction with explanations in health social networks". Information Sciences. 384: 298-313 (2017)

19. Noy, N.F. and D.L. McGuinness, "Ontology development 101: A guide to creating your first ontology", 2001, Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880, Stanford, CA.

20. Olshausen, B.A., "Emergence of simple-cell receptive field properties by learning a sparse code for natural images". Nature, 1996. 381(6583): p. 607-609.

21. Phan, N., et al., "Ontology-based deep learning for human behavior prediction in health social networks". 2015: p. 433-442.

22. Phan, N., et al. "Analysis of physical activity propagation in a health social network", in Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. 2014. ACM.

23. Pate, R.R., et al., "Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine", Jama, 1995. 273(5): p. 402-407.

24. Poon, H. and P. Domingos. "Sum-product networks: A new deep architecture. in Computer Vision Workshops" (ICCV Workshops), 2011 IEEE International Conference on. 2011. IEEE.

25. Smolensky, P., "Information processing in dynamical systems: Foundations of harmony theory", 1986, DTIC Document.

26. Skiena. S, "Hasse Diagrams", In Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica. Reading, MA: Addison-Wesley, p. 163, 169-170, and 206-208, 1990.

27. Suk, H.I., et al., "Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis". Neuroimage, 2014. 101: p. 569-82.

28. Trong Hai Duong, Ngoc Thanh Nguyen, Cuong Duc Nguyen, Thi Phuong Trang Nguyen, Ali Selamat: "Trust-Based Consensus for Collaborative Ontology Building". Cybernetics and Systems 45(2): 146-164 (2014)