Ontology và TOPSIS

Ngày 10 tháng 12 năm 2024

1 Module sử dụng Ontology tìm thuốc cho bệnh nhân

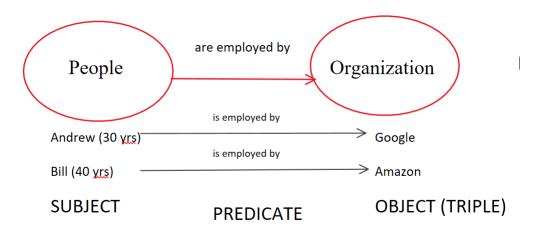
1.1 Đinh nghĩa Ontology

- Cụm từ "Ontology" là một từ ghép trong tiếng Hy Lạp được cấu thành từ hai chữ "onto" nghĩa là sự tồn tại và "logy" nghĩa là khoa học. Trong triết học, "Ontology" là ngành khoa học nghiên cứu về bản chất sự tồn tại của các thực thể và cách chúng liên quan với nhau.
- Tương tự, trong khoa học máy tính, "Ontology" là mô hình dữ liệu biểu diễn một lĩnh vực cụ thể thông qua các khái niệm, mối quan hệ và các quy tắc giữa chúng.

1.2 Cấu trúc của Ontology

- Individuals (cá thể): là thành phần cơ bản và hạt nhân nhất của ontology.
- Classes (lớp): là tập hợp của các lớp con hoặc các cá thể có đặc điểm chung.
- Properties (thuộc tính): là thuộc tính mà cá thể sở hữu.
- Relations (quan hệ): là mối quan hệ giữa các đối tượng được diễn tả thông qua các thuộc tính của chúng.

Ví dụ về cấu trúc của Ontology:



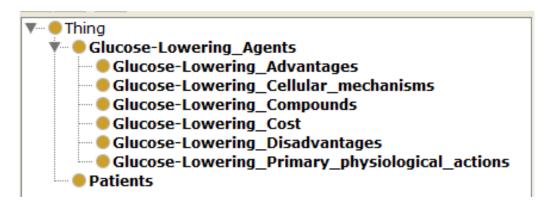
Hình 1: Cấu trúc Ontology

- 1. Classes: trong hình vẽ này ta thấy có hai class là "People" và "Organization".
- 2. Individuals: ta thấy có 4 individual được thể hiện trong hình minh họa là "Andrew", "Bill", "Google" và "Amazon".
- 3. Properties: ta dễ dàng thấy property của Andrew là 30 years old còn Bill là 40 years old.
- 4. Relation: ta thấy mối quan hệ giữa Andrew và Google, cũng như Bill và Amazon hay mọi người và tổ chức đó là "được tuyển dụng".

Dễ thấy từ hình vẽ mô tả trên, ta có thể thấy Ontology chứa dữ liệu dưới dạng bộ ba (triple) và thể hiện cấu trúc Subject-Predicate-Object rất rõ ràng.

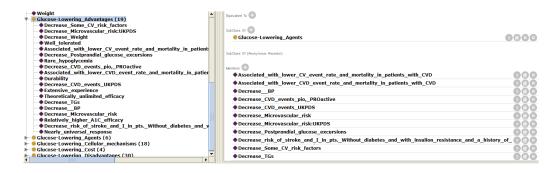
1.3 Sử dụng Ontology để tìm thuốc cho bệnh nhân

1. Sử dung phần mềm Protege để tao Ontology:

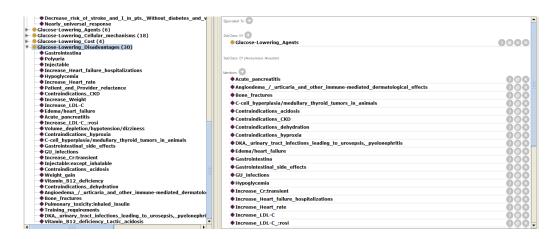


Hình 2: Class và các subclass sử dụng trong Ontology

- Glucose-Lowering Agents: class này thể hiện các tác nhân giúp giảm đường huyết qua các loại thuốc MET, SU, GLP-1,...
- Glucose-Lowering Advantages: subclass này thể hiện lợi ích của các loại thuốc.
- Glucose-Lowering Disadvanges: subclass này thể hiện han chế của các loại thuốc



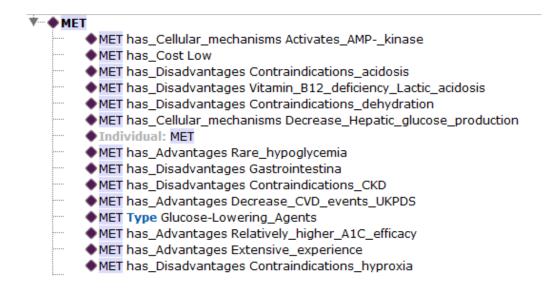
Hình 3: Subclass Glucose-LoweringAdvantages



Hình 4: Subclass Glucose-LoweringDisadvanges

Property name	Property type	Domain	Range		
has_Advantages	Object	Glucose-Lowering_Agents	Glucose-Lowering_Advantages		
has_Cellular_mechanisms	Object	Glucose-Lowering_Agents	Glucose-Lowering_Cellular mechanisms		
has_Compounds	Object	Glucose-Lowering_Agents	Glucose-Lowering_Compounds		
has_Cost	Object	Glucose-Lowering_Agents	Glucose-Lowering_Cost		
has_Disadvantages	Object	Glucose-Lowering_Agents	Glucose-Lowering_Disadvantages		
has_Primary_physiological_actions	Object	Glucose-Lowering_Agents	Glucose-Lowering_Primary physiological_actions		
has_History_of_Diseases	Object	Patients	Glucose-Lowering_Disadvantages		
has_Adverse Drug_Reactions	Object	Patients	Glucose-Lowering_Agents		
Not_recommended	Object	Patients	Glucose-Lowering_Agents		
ID_No	Data	Patients	xsd: string		

Hình 5: Các relation trong Ontology



Hình 6: Thông tin về thuốc MET

- 2. Sử dụng SPARQL Query để truy vấn data trong Ontology:
 - 1. Cấu trúc cơ bản của SPARQL:
 - PREFIX: định nghĩa các tiền tố URI
 - SELECT: xác định biến truy vấn
 - WHERE: định nghĩa điều kiện lọc dữ liệu

Ví dụ ta cần lấy thông tin về những cá thể thuộc class Person và tên của họ từ một dataset:

```
PREFIX ex: <a href="http://example.org/">http://example.org/</a>
SELECT ?person ?name
WHERE {
    ?person a ex:Person. \Tìm tất cả các cá thể kiểu "person"
    ?person ex:hasName ?name. \Tìm tên của từng cá thể qua thuộc tính hasName
}
```

2. Giải thích:

- (a) PREFIX: Quy ước đường dẫn rút gọn tới địa chỉ triển khai Ontology là ex.
- (b) SELECT: Xác định kết quả trả về của truy vấn cho hai biến ?person và ?name.
- (c) WHERE: Xác định cách các biến và mối quan hệ giữa chúng được lấy ra từ data.

Hệ thống hỗ trợ ra quyết định cho bệnh nhân tiểu đường sẽ dựa vào tiền sử của bệnh nhân, truy vấn xem tiền sử này liên quan đến tác dụng phụ của các loại thuốc nào để loại bỏ chúng ra khỏi danh sách khuyến nghị cho bệnh nhân.

2 Module xếp hạng thuốc tiểu đường sử dụng TOPSIS

Sau khi đã tìm được thuốc phù hợp, việc tiếp theo ta cần làm là LOẠI BỔ các nhóm thuốc không phù hợp và XÉP HẠNG nhóm thuốc còn lại để tìm ra lựa chọn tối ưu nhất cho bệnh nhân. Chúng ta sẽ phân tích ví dụ sau: Giả sử có một bệnh nhân mẫn cảm với GLP-1 và có tiền sử bệnh increasing LDL-C, Edema và một bảng thống kê rủi ro thuốc kèm chi phí.

Properties	Antidiabetic medications							
	MET	GLP-1	SGLT2	DPP-4	TZD	SU	Insulin	
Нуро	3	3	3	3	3	7	7	
Weight	1	1	1	3	5	7	7	
Renal/GU	7	7	5	3	3	7	7	
GI Sx	5	5	3	3	3	3	3	
CHF	3	3	3	3	5	3	3	
CVD	1	3	3	3	3	5	3	
Bone	3	3	3	3	5	3	3	
Cost	1	3	3	3	1	1	3	

MET: metformin (Biguanides); SU: sulfonylureas; Hypo: hypoglycemia; GU: genitourinary; GI Sx: glycemic index symptom; CHF: congestive heart failure; CVD: cardiovascular diseases.

Hình 7: Rủi ro thuốc và chi phí

Sau khi sử dụng Ontology, ta đã LOẠI BỔ được 3 loại thuốc không phù hợp vơi bệnh nhân là GLP-1, TDZ và SGLT2, ta sẽ sắp xếp thứ tự ưu tiên của 4 loại thuốc sau đây:

Antidiabetic medications	Properties							
	Hypo	Weight	Renal/GU	GI Sx	CHF	CVD	Bone	Cost
MET	3	1	7	5	3	1	3	1
DPP-4	3	3	3	3	3	3	3	3
SU	7	7	7	3	3	5	3	1
Insulin	7	7	7	3	3	3	3	3

Hình 8: Danh sách chưa được sắp xếp

Bước 1: Xây dựng ma trận quyết định

Hình 9: Ma trận quyết định

Bước 2: Chuẩn hóa ma trận quyết định

$$R = \begin{bmatrix} R_{ij} \end{bmatrix}_{m \times n} = \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix},$$

$$r_{ij} = a_{ij} / (\sqrt{\sum_{k=1}^{m} a_{kj}^2}), i = 1, 2, \dots, m, j = 1, 2, \dots, n.$$

Hình 10: Chuẩn hóa ma trận quyết định

$$R_{4\times8} = \left[\begin{array}{cccccccccc} 0.279 & 0.096 & 0.560 & 0.693 & 0.500 & 0.151 & 0.500 & 0.224 \\ 0.279 & 0.289 & 0.240 & 0.416 & 0.500 & 0.452 & 0.500 & 0.671 \\ 0.650 & 0.674 & 0.560 & 0.416 & 0.500 & 0.754 & 0.500 & 0.224 \\ 0.650 & 0.674 & 0.560 & 0.416 & 0.500 & 0.452 & 0.500 & 0.671 \end{array} \right].$$

Hình 11: Kết quả

Bước 3: Xác định trong số của các thuộc tính và chi phí (Hypo, Weight, Cost...):

$$W = [w_1 \ w_2 \ w_3 \ w_4 \ w_5 \ w_6 \ w_7 \ w_8] = [1\ 1\ 1\ 1\ 2\ 2\ 1\ 1]$$

Hình 12: Ma trận trọng số

Bước 4: Xây dựng ma trận chuẩn hóa trọng số

$$V = \begin{bmatrix} V_{ij} \end{bmatrix}_{m \times n} = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{m1} & \cdots & v_{mn} \end{bmatrix},$$

$$v_{ij} = w_j r_{ij}$$
, $i = 1, 2, ..., m$, and $j = 1, 2, ..., n$

Hình 13: Ma trân chuẩn hóa trong số

$$V_{4\times8} = \left[\begin{array}{ccccccccc} 0.279 & 0.096 & 0.560 & 0.693 & 1.000 & 0.302 & 0.500 & 0.224 \\ 0.279 & 0.289 & 0.240 & 0.416 & 1.000 & 0.905 & 0.500 & 0.671 \\ 0.650 & 0.674 & 0.560 & 0.416 & 1.000 & 1.508 & 0.500 & 0.224 \\ 0.650 & 0.674 & 0.560 & 0.416 & 1.000 & 0.905 & 0.500 & 0.671 \\ \end{array} \right]$$

Hình 14: Kết quả

Bước 5: Xác định các giải pháp lý tưởng và tiêu cực:

$$\begin{split} A^* &= \{V_1^*, V_2^*, \dots, V_n^*\} \quad \textit{v\'oi} \quad V_j^* = \min_i V_{ij} \text{ , } i = 1, 2, \dots, m \\ A^- &= \{V_1^-, V_2^-, \dots, V_n^-\} \quad \textit{v\'oi} \quad V_j^- = \max_i V_{ij} \text{ , } i = 1, 2, \dots, m \end{split}$$

Hình 15: Phương pháp

$$\begin{split} A^* &= \left(\min_i v_{i1}, \min_i v_{i2}, \min_i v_{i3}, \min_i v_{i4}, \min_i v_{i5}, \min_i v_{i6}, \min_i v_{i7}, \min_i v_{i8} \right) \\ &= \left(0.279, 0.096, 0.240, 0.416, 1.000, 0.302, 0.500, 0.224 \right), \\ A^- &= \left(\max_i v_{i1}, \max_i v_{i2}, \max_i v_{i3}, \max_i v_{i4}, \max_i v_{i5}, \max_i v_{i6}, \max_i v_{i7}, \max_i v_{i8} \right) \\ &= \left(0.650, 0.674, 0.560, 0.693, 1.000, 1.508, 0.500, 0.671 \right). \end{split}$$

Hình 16: Kết quả

Bước 6: Tính toán khoảng cách cho từng giải pháp:

$$S_{i}^{*} = \sqrt{\sum_{j=1}^{n} \left(V_{ij} - V_{j}^{*}\right)^{2}}, \quad i = 1, 2, ..., m,$$

$$S_{i}^{-} = \sqrt{\sum_{j=1}^{n} \left(V_{ij} - V_{j}^{-}\right)^{2}}, \quad i = 1, 2, ..., m.$$

Hình 17: Phương pháp

$$S_1^* = 0.424,$$
 $S_1^- = 1.458,$ $S_2^* = 0.775,$ $S_2^- = 0.911,$ $S_3^* = 1.424,$ $S_3^- = 0.526,$ $S_4^* = 1.067,$ $S_4^- = 0.664.$

Hình 18: Kết quả

Bước 7: Tính độ gần tương đối với mỗi phương pháp

$$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}, \quad i = 1, 2, ..., m.$$

Hình 19: Phương pháp

$$C_1^* = 0.775,$$

 $C_2^* = 0.540,$
 $C_3^* = 0.270,$
 $C_4^* = 0.384.$

Hình 20: Kết quả

Vậy ta kết luận phương án thuốc được sử dụng ưu tiên lần lượt là $\mathrm{MET} > \mathrm{DPP-4} > \mathrm{SU} > \mathrm{Insulin}$