



Sematic Web Mining Using Fuzzy C-means Algorithm

Wria Mohammed Salih Mohammed^{1*} and Mohamad Mehdi Saraee²

¹School of Basic Education, Computer Science, University of Sulaimani, Kurdistan, Iraq.

²Islamic Azad University, Shahreza Branch, Shahreza, Iran.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/BJMCS/2016/25471

Editor(s):

(1) Kai-Long Hsiao, Taiwan Shoufu University, Taiwan.

Reviewers:

(1) M. Bhanu Sridhar, GVP College of Engg. for Women, Vizag, India.

(2) Anonymous, Osaka Prefecture University, Japan.

(3) Alexandre Ripamonti, University of Sao Paulo, Brazil.

Complete Peer review History: <http://sciencedomain.org/review-history/14504>

Received: 4th March 2016

Accepted: 19th April 2016

Published: 7th May 2016

Original Research Article

Abstract

Sematic web mining (SWM) is the incorporation of two astonishing development research areas; semantic web and data mining. Sematic web can promote the performances and productivities of web mining. Also, data mining approaches can be applied on the semantic web data because the semantic web data is prosperous sources of knowledge to feed the data mining techniques.

In this research a SWM system is designed and implemented using fuzzy C-means (FCM) algorithm. This is performed by developing an application that is created using several techniques. The system makes up of creating the semantic web dataset, dataset query (SPARQL) and converting the semantic web dataset into traditional dataset. After that, data mining is implemented encompassing data pre-processing, fuzzy C-means algorithm and finally exploring the results.

SWM using FCM has been practiced by producing an application, which involves various techniques such as DotNetRDF, C# programming language, SPARQL for query language. The final results that obtained from the achieved system are more accurate and knowledgeable because of the combination between semantic web and Fuzzy C-means.

Keywords: Data mining; semantic web; fuzzy C-means (FCM); Sematic Web Mining (SWM).

*Corresponding author: E-mail: wria.salih@univsul.edu.iq;

1 Introduction

Semantic Web (SW) describes the meaning of web contents so that a machine can understand it. SW makes knowledge a machine-processable form. It is the continuousness of the World Wide Web, which allows users to be able use their knowledge and share the information with others via websites and applications.

Also, Data Mining (DM), which also called knowledge discovery from data (KDD), is automatically searching to find the patterns and trends that go beyond analysis. Also it can be said that DM is analyzing data from various sources and encapsulating it into potentially useful knowledge.

Applying DM on the SW data desires some processes that are called semantic web mining (SWM). The purpose of SWM is to improve the result of data mining by gaining advantages of the SW.

Clustering methods are unsupervised approaches that can be utilized to establish data into groups according to similarities among the data items. Clustering algorithms consist of hard partition and fuzzy partition; hard partition such as K-means algorithm that makes data into groups, and fuzzy clustering such as C-means that is a clustering method which allows one data item belongs to more than one group.

The main goal of this research is applying fuzzy c-means algorithm on semantic web data proficiently and productively. The proposed system is SWM using FCM clustering algorithm. Fuzzy clustering is one of the techniques that developers can obtain flexibility in DM because each data item can belong to more than one group. It also can use the fuzzy logic to improve SW and obtain more accurate results.

2 Problem Statement

The problem of this research is implementing fuzzy c-means algorithm on semantic web data because flexible clusters can obtain from fuzzy c-means algorithm. However, K-means, which is a hard clustering algorithm, is used to cluster objects into groups. In the most circumstances fuzzy c-means is more natural than hard k-means clustering algorithm. The used algorithm in this research is fuzzy C-means algorithm. A system is developed for this reason. The proposed system consists of two main parts. First part is creating the semantic web dataset, query language (SPARQL) and converting the semantic web dataset into traditional dataset. Second part, Fuzzy c-means algorithm is implemented encompassing data pre-processing, fuzzy clustering mining and finally exploring the fuzzy mining results. The proposed system for this research includes various techniques such as DotNetRDF for Semantic web part in C#, also C# programming language, and SPARQL for query language.

The purpose of using both area; semantic web and data mining together is to improve the result of data mining by gaining advantages of the SW. the reason of using Fuzzy C-means algorithm is to obtain flexible clusters because each data item can belong to more than one group. Moreover, fuzzy logic can be used to improve semantic web data and obtain more accurate results.

3 Related Work

SW is the extension of the current web and web mining is the discovering of important information from web data. [1] illustrated DM on SW using clustering metadata; they grouped the data using measuring similarities to serve as input to the hierarchical clustering algorithms. Grouping according to their similarities among the metadata is chosen and the closest data from each other is in the same group.

Moreover, [2] explain SW for web mining using roadmap. Firstly [2] mentions DM and World Wide Web. DM can be used to analyses the behavior of individual user's collection properties of documents and access the patterns of website and both classification and clustering techniques are used to analyses the documents. Furthermore, he explained how DM techniques are used to extract information also DM for and with SW,

for this reason he used classification and hierarchical clustering on SW documents. However, he does not use any specific algorithm. The paper [2] is a survey about roadmap SW joining with web mining.

In [3] SW is used with DM techniques using Two different models based on ontology mapping and ontology mapping rules. The first model that they use is on the premise of the SW data, and then based on this retrieving the SW knowledge rules. The first model is illustrated in the Fig. 1.

Cai Jiaojie et al. [3] use HowNet algorithm to establish semantic similarity of concepts.

However, second model is implementing fuzzy semantic mapping between domain ontology rules and knowledge rules. In the second model [3] uses fuzzy mapping for membership calculation, but they still use HowNet algorithm as shows in Fig. 2.

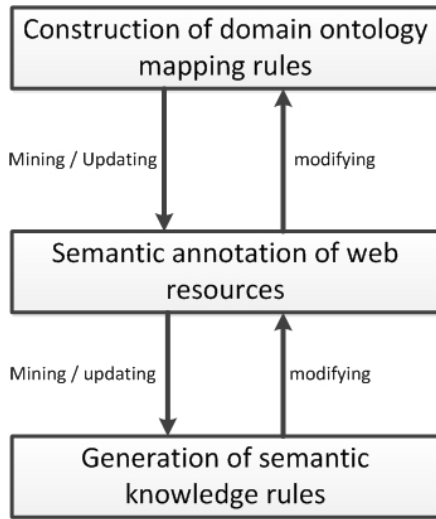


Fig. 1. Semantic web mining first module

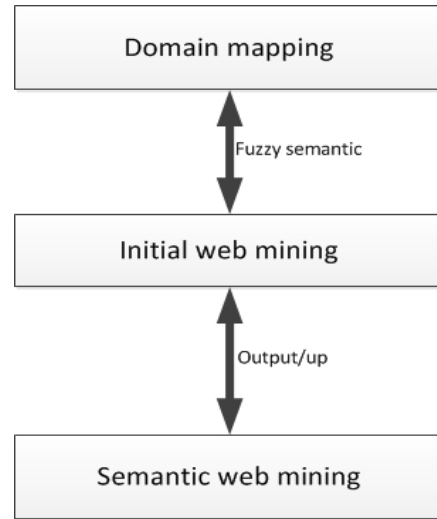


Fig. 2. Fuzzy semantic mapping (Second module)

Furthermore, [4] proposed mining complex SW data without end-user involvement. Also, they mine SW data using association rule mining. Firstly, they store data in triple format in a database, and then apply association rule mining. The algorithm that [4] proposed called SWApriori. SWApriori accepts dataset that contains triples with minimum support and minimum confidence.

The paper, which written by [5] proposed penalized posterior probability based FCM algorithm to group user pattern information to cluster the users. As a result, the similar group of the user is detected. Furthermore, fuzzy clustering depends on probability of data to be clustered. FCM the center of a cluster is the mean of the all user pattern points.

Chhaman and Usvir [6] proposed web-mining uses FCM algorithm, which is applied to the text biomedical article and the output, is an XML file that includes the created clusters. Also, [7] present web usage mining using FCM algorithm because web-log data is unsupervised data, thus FCM algorithm can support web-log data to classify the fuzzy clusters. The main aim of [7] is to discover the best solution of FCM mixing with genetic algorithm to avoid error rate of pattern.

In this research, clustering FCM algorithm on SW data is applied. Also, in SW, RDF is the basic to create SW datasets so that RDF is created on XML base. Constantly, RDFS is vocabulary for RDF; also, OWL is the extension of RDFS. Furthermore, OWL is richer and newer than RDFS.

4 Backgrounds

4.1 Semantic web

Semantic Web is a major area of interest within the field of World Wide Web. Basically SW is defined as a web of data, which linked in a way to create semantics that adhere to language, constructs and defined grammar [8]. Moreover, Berners-Lee has illustrated that SW is to make the web machine-processable. SW consists of the main blocks that are shown in Fig. 3.

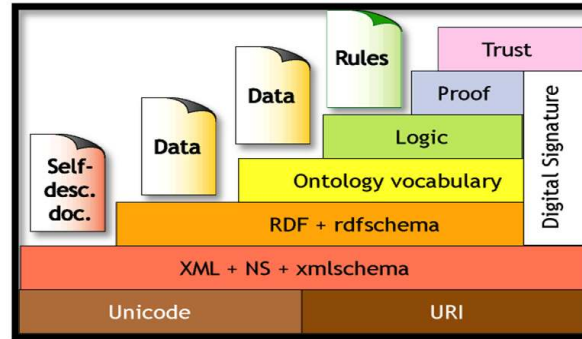


Fig. 3. Semantic web structure

XML (Extensible Mark-up Language) is used to show document and data in a text-based format that can be passed via standard Internet Protocols [9]. Additionally, it is used for transferring data from one application to another [10].

The Resource Description Framework (RDF) is a language that is produced to enrich and help the SW. RDF is similar to XML and HTML and it supports resource description and meta-data. RDF triples have three main components; subject and predicate and object [11]. Subjects and predicates are URI, but object can be URI or literal. Literals are surrounded with < > brackets [12].

As it is mentioned that RDF is based on XML and XML/RDF model is characterized as a graph of triples, the RDF syntaxes mixes with the XML tags which is called RDF/XML. The syntax of RDF/XML is more comprehensible than RDF [13].

RDFS is W3C recommendation and it is schema language for RDF [14]. RDFS provides data modeling vocabulary for RDF. Moreover, it is an extension of RDF. RDFS is a mechanism to describe related sources and relationship between sources for example we can use domain and range for properties. RDFS includes rdfs:Literal, rdfs:subClassOf, rdfs:Datatype, rdfs:range, rdfs:domain, rdfs:comment and rdfs:Resource [15].

SPARQL is a query language in SW, where user can query various RDF data sources via SPARQL [16]. If SQL is query language for relational database then SPARQL is query language for SW [17]. There are some other query languages for SW, but SPARQL is the most powerful query on SW data. Also SPARQL is W3C recommendation [18,19].

4.2 Basic of C-means algorithm

The FCM clustering algorithm is the most extensively used fuzzy clustering algorithm, Professor Jim Bezdek invented this algorithm in 1981. The FCM algorithm splits data to fuzzy groups with respect to some particular criterion. In FCM each data item belongs to two or more clusters. However, in non-fuzzy or hard

clustering, data is classified into crisp clusters and each data item has exactly one centroid. The result of FCM algorithm is a list of centroids [20].

This method is based on the following equation:

$$J_m = \sum_{i=0}^n \sum_{k=0}^n u_{ij}^m \|x_i - c_j\|^2$$

When $1 \leq m < \infty$

Where u_{ij}^m is the value of membership of x_i , in c_j ($c =$ cluster). $\|x_i - c_j\|^2$ is measuring distance between x_i and c_j , it means it is the distances between center and each point. In the above equation u_{ij} and c_j usually update until the final result is realized, u_{ij} and c_j are keeping updated according to the following functions:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}} \quad c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}$$

K is the number of steps. The iteration is stopped if $|u_{ij}^{(k+1)} - u_{ij}^{(k)}| < \partial$ where ∂ a termination creation between 0 and 1, this procedure is updated J_m [21]. The algorithm of fuzzy C-means, as shows in Fig. 4, consists of the following steps:

- 1- Initialize u_{ij}^0
- 2- Calculate the centroids vectors (c_j):

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}$$

- 3- Update $U(k)$ and $U(k+1)$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

- 4- If $|u_{ij}^{(k+1)} - u_{ij}^{(k)}| < \partial$ then stop and return results. Otherwise return to step 2 [22].

In fuzzy C-means clustering algorithm, the data is in the form of matrix. Furthermore, all values in Matrix are between 0 and 1, and every point is compared to centroids.

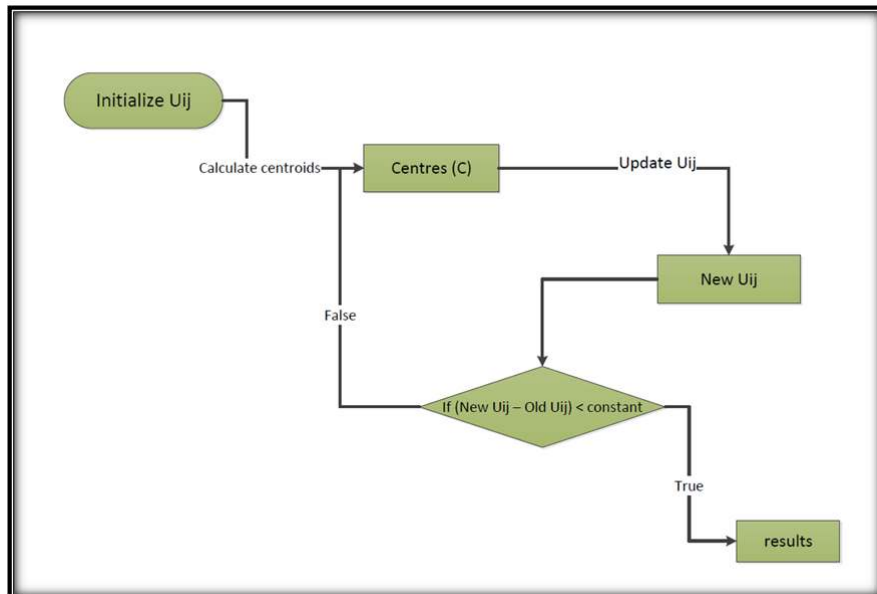


Fig. 4. C-means algorithm

The following example shows the comparison between K-means and C-means:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

K-means algorithm

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \\ 0.6 & 0.4 \\ 0.7 & 0.3 \\ 0.8 & 0.2 \end{bmatrix}$$

C-means algorithm

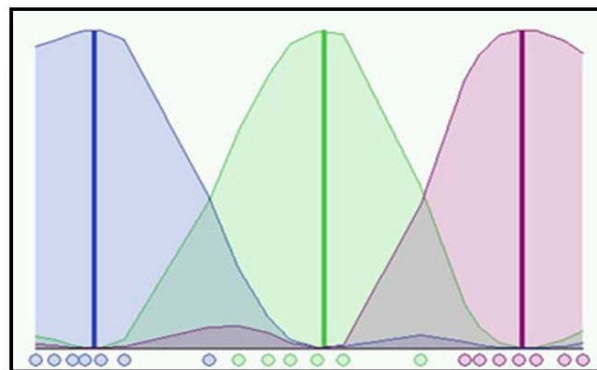


Fig. 5. Fuzzy C-means (3 clusters)

Fig. 5 highlights that the data consists of three clusters, it is in the 37th steps of running algorithm, and $\partial = 0.01$, where ∂ is a termination criterion between 0 and 1 [23,21].

4.3 Using data mining techniques on semantic web data

The SWM is a combination of data mining and SW; this includes semantic knowledge extraction and semantic mining [24,25]. SW provides meaning for web data and creating heterogeneous data structure. Furthermore, web mining is extracting important and useful knowledge and interesting from heterogeneous data and web data [26]. The purpose of SWM is to improve the result of web mining by obtaining advantages of the SW. Moreover, making use of web mining for accumulation of the SW by retrieving useful patterns, structures, similar meanings and semantic relations from existing web resources [27,28]. Both SW and DM are complementing each other and generating more chances for researchers in the world. As we know most of the data on the web is unstructured data which only human can understand it. However, by using SW it can be seen that data can be processed by machine more efficiently. Hence, the combination between them makes the process easier for web miner and web semantic developer to develop the mining semantic web data [27]. SWM, which shows in Fig. 6, is used in different fields such as biomedical, education and clinical scenarios [26,29].

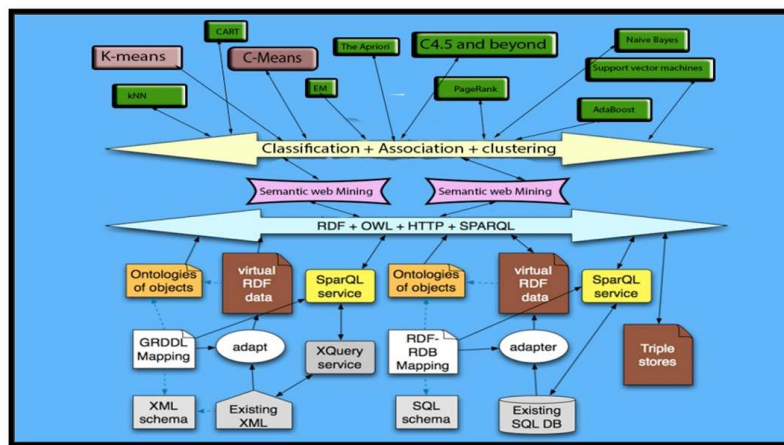


Fig. 6. Semantic web mining

Mining SW data requires collecting various resources from multiple websites in the form of RDF, OWL, RSS or FOAF [30]. After that, it needs cleaning by using different techniques for example removing log entries that are not necessary for mining such as scripts and graphics [31]. Next, knowledge is extracted and analyzed from the SW data.

5. Methodology

Mining semantic web using fuzzy C-means algorithm and DotNetRDF. DotNetRDF is used with writing programming language codes. How the proposed approach works, can be summarized as follows:

SW dataset can be created or bring it from external sources, following by checking the dataset for validity. As a result, the valid dataset and triples are obtained.

Writing programming language with DotNetRDF for SW data to extract important information from the data. In this paper, C# programming code in Visual studio and DotNetRDF.

Finally, the table of data, which contains URIs and literals, and the final results, is ready to mined using fuzzy c-means. Applying fuzzy c-means produces the centroids, fuzzy clusters.

This system is classified into two main parts, the first part is Semantic web and second part is Data Mining.

5.1 Semantic web part

First of all, we need a dataset to implement. The dataset can be created using DotnetRDF tools in Visual studio. Although, dataset can be imported as either external dataset from a machine or an external online link. The dataset needs to be checked for validity because it is possible that the external SW data is invalid dataset. Then, the valid RDF, RDFS or Owl is obtained from the SW dataset, as well as, triples (subjects, predicates, objects) as shows in Fig. 7. Next step is converting the triples to traditional dataset.

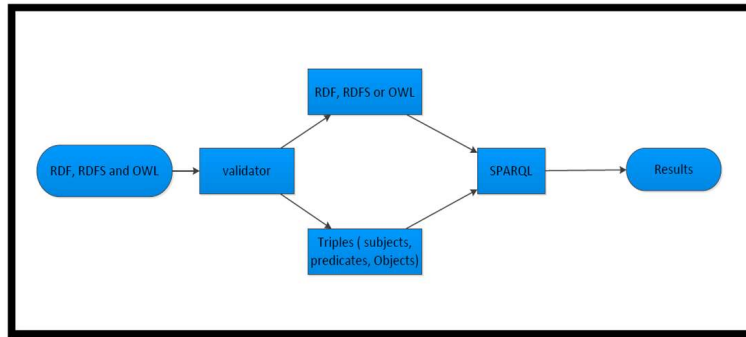


Fig. 7. Validate and SPARQL of SW datasets

5.2 Converting semantic web dataset to traditional dataset

To convert SW data to traditional dataset, there are two methods to undertake that, in this project both of them are utilized. The first method is using SPARQL query to convert SW dataset to traditional dataset and the second method is using programming language code. SPARQL derives information from SW dataset. Firstly, the semantic web data are required to be analyzed. The important information can be retrieved from SW dataset. Second method is using programming language to retrieve important information from RDF, RDFS or OWL. The problem with both methods is that for each dataset, particular query or particular programming code is required to write. Also, DotNetRDF is required if programming language is used to retrieve important information from the SW dataset.

5.3 Data mining part

The second step is mining of the web semantic dataset using fuzzy C-Means algorithm that is one of the most widely used and well-known algorithm. Data mining part is implemented according to Fig. 8.

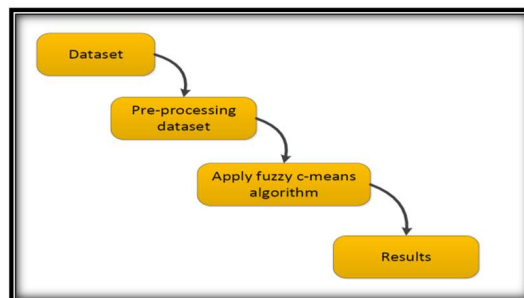


Fig. 8. Data mining processes

Dataset: the traditional dataset, which is produced from the semantic web data, is dataset for the system. Also, we can add dataset from other sources.

Preprocessing dataset: is cleaning dataset from the noisy data, and arranging dataset.

Apply fuzzy c-means algorithm: is applying the proposed algorithm on the dataset. And finally we have the results.

The reason to apply FCM algorithm in this paper is that it allows one data item belongs to more than one clusters. But any data item in the hard cluster algorithms belongs to one cluster only. The FCM algorithm works by conveying membership to each data point equivalent to each cluster center on the base of distance between the cluster and the center. Noticeably, the summation of membership of any data item is equal to one.

6 Results

6.1 Creating semantic web data

One of the uses of DotNetRDF in Visual Studio 2010 is creating SW datasets. Creating the SW dataset requires programming codes. SW dataset is RDF, RDFS or OWL, and then triple is produced from the SW dataset in order to support users to understand SW dataset easily. Triples are consisted of three main components; subjects, predicates, and objects which are shown in Fig. 9.

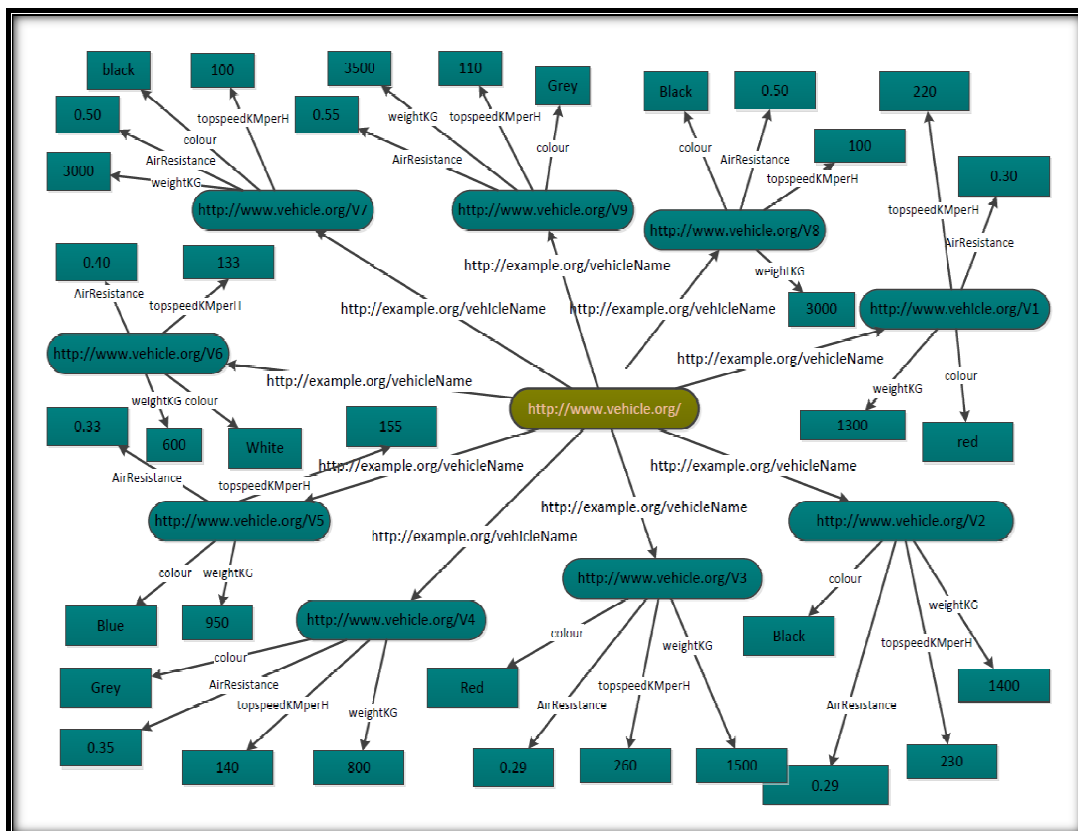


Fig. 9. The graph of vehicle RDF

6.2 Validating semantic web dataset

RDF can be created using DotNetRDF or it can bring from online link. As well as the semantic web data can be copied and past into the application. Then the semantic web data can be checked for validity before converting into triples and traditional dataset (Fig. 10).

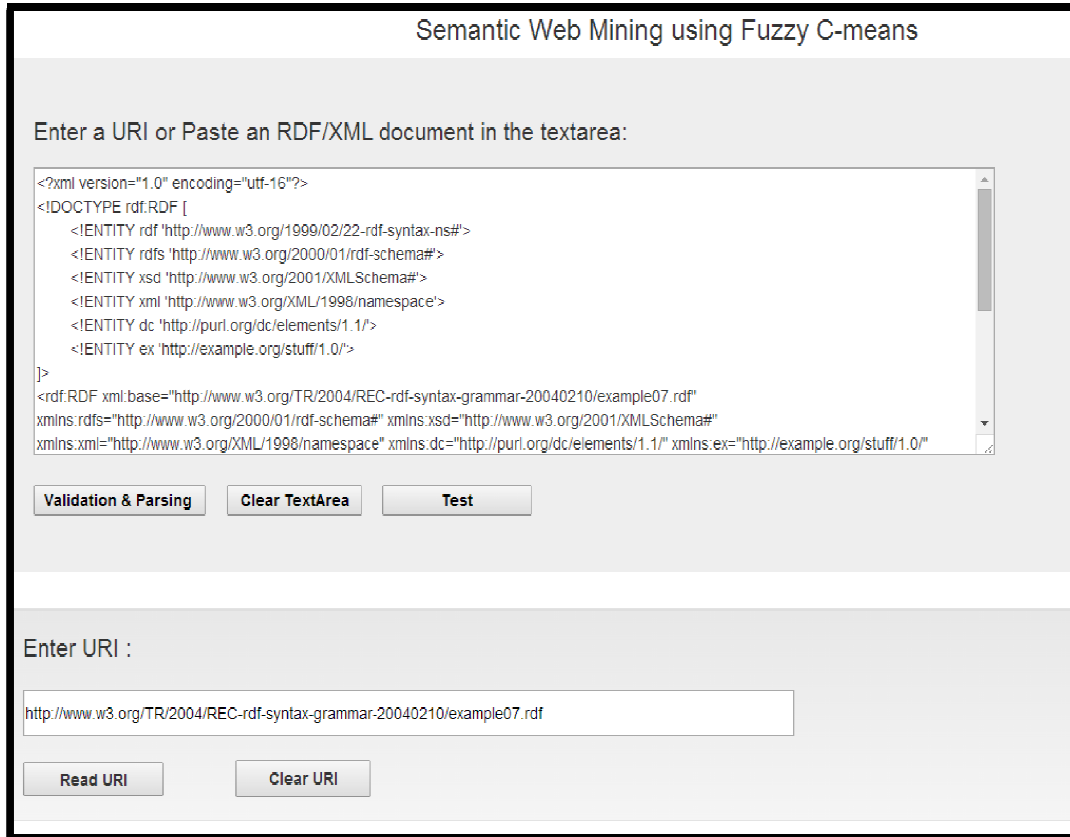


Fig. 10. Bring OWL from URI

6.3 Triples

Triples consist of subjects, predicates and objects. The three main components of the Triple can be separated as explained in the Fig. 11. In the above application when the user clicks on (validation & parsing) button, it will show the triples and valid RDF (Fig. 11).

6.4 Results of the data mining part

In DM part, fuzzy C-means algorithm is applied. Firstly, we need to upload the excel file which is ready to be mined, then, the excel file shows in the system as explained in Fig. 12.

After uploading the excel file, all the column's names are appeared in the drop down list as shows in figure 13. Thus, we need to specify one of the columns that we want to mine according to it.

The original RDF/XML document :

```
<?xml version="1.0" encoding="utf-16"?>
<!DOCTYPE rdf:RDF [
  <ENTITY rdf http://www.w3.org/1999/02/22-rdf-syntax-ns#>
  <ENTITY rdfs http://www.w3.org/2000/01/rdf-schema#>
  <ENTITY xsd http://www.w3.org/2001/XMLSchema#>
  <ENTITY xml http://www.w3.org/XML/1998/namespace>
  <ENTITY dc http://purl.org/dc/elements/1.1/>
  <ENTITY ex http://example.org/stuff/1.0/>
]>
<rdf:RDF xml:base="http://www.w3.org/TR/2004/RFC-rdf-syntax-grammar-20040210/example07.rdf" xmlns:rdf="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:xml="http://www.w3.org/XML/1998/namespace" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:ex="http://example.org/stuff/1.0/" xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
    <ox:editor rdf:nodeID="autos1" />
    <dc:title>RDF/XML Syntax Specification (Revised)</dc:title>
  </rdf:Description>
  <rdf:Description rdf:nodeID="autos1">
    <ex:fullName>Dave Beckett</ex:fullName>
    <ex:homePage rdf:resource="http://purl.org/net/dajobe/" />
  </rdf:Description>
</rdf:RDF>
```

Subject	Predicate	Object
http://www.w3.org/TR/rdf-syntax-grammar	http://example.org/stuff/1.0/editor	_:autos1
http://www.w3.org/TR/rdf-syntax-grammar	http://purl.org/dc/elements/1.1/title	RDF/XML Syntax Specification (Revised)
_:autos1	http://example.org/stuff/1.0/fullName	Dave Beckett
_:autos1	http://example.org/stuff/1.0/homePage	http://purl.org/net/dajobe/

Fig. 11. Validation RDF and creating Triples from URI source

Import traditional dataset or any excel data :

Name	AirResistance	topspeedKMperH	colour	weightKG
http://www.vehicle.org/V10.3	220		red	1300
http://www.vehicle.org/V20.32	230		black	1400
http://www.vehicle.org/V30.29	260		red	1500
http://www.vehicle.org/V40.35	140		gray	800
http://www.vehicle.org/V50.33	155		blue	950
http://www.vehicle.org/V60.4	130		white	600
http://www.vehicle.org/V70.5	100		black	3000
http://www.vehicle.org/V80.6	105		red	2500
http://www.vehicle.org/V90.55	110		gray	3500

No file chosen

Fig. 12. Showing the traditional dataset after uploaded excel file

Select Target :

- Name
- AirResistance
- topspeedKMperH
- colour
- weightKG

Fig. 13. Select the target column

We can see the results as it is shown in Fig. 14, if we choose topSpeedKMperH column.

Uij1	Uij2
0.84241681117096	0.15758318882904
0.980298885627424	0.0197011143725763
0.570289473262615	0.429710526737385
0.651005526187301	0.348994473812699
0.528922662678302	0.471077337321698
0.390529512130494	0.609470487869506
0.2953517188387	0.7046482811613
0.14086773614993	0.859132263850069
0.0494870828146414	0.950512917185359

Fig. 14. The Matrix of Uij

It can be noticed, in Fig. 14, that each value can be divided into two numbers and the total of these two numbers needs to equal one, for example (0.9, 0.1), (0.3, 0.7) and (0.4, 0.6). To make the values between 0 and 1 all the values need to be divided by the largest number in the selected dataset. The result of Uij1 minus one is equal to Uij2.

As a result, we have matrix of Uij and keep updating Uij until we have the final matrix, figure 14 is the final matrix for the above DM example. The summation of every both values in the same level is equal to one, For example

$$\begin{aligned}
 0.84241681117096 + 0.15758318882904 &= 1 \\
 0.980298885627424 + 0.0197011143725763 &= 1 \\
 0.570289473262615 + 0.429710526737385 &= 1
 \end{aligned}$$

These values, which can range between zero to one, are the membership values. For example, the membership of 0.0197011143725763 is very low. But the membership of 0.980298885627424 is very high, even it is close to crisp membership. Every values in the dataset has two values, both values are the memberships of different clusters.

The dataset is divided to two clusters so that there are two centroids. In Fig. 15, the list of the centroids has been clearly seen; the loops of both centroids are continuing until the most proper centroids are found then the loop will stop. The final values for centroids are (201.0178, 124.654560). Also, the clustering will be grouped according to final centers.

centre_one	centre_two
220	300.539740610524
225.752160902114	299.454164920371
231.332865152214	264.741265137359
215.335934777181	218.933138933605
209.082981301145	193.30435995826
204.853891549325	167.884106539349
201.741923701711	144.173328296816
201.093159191198	130.785126847429
201.017830662926	124.654560935817

Fig. 15. The centers

Furthermore, we can see the chart of matrix U_{ij} and clusters as shows in the Figs. 16 and 17 respectively.

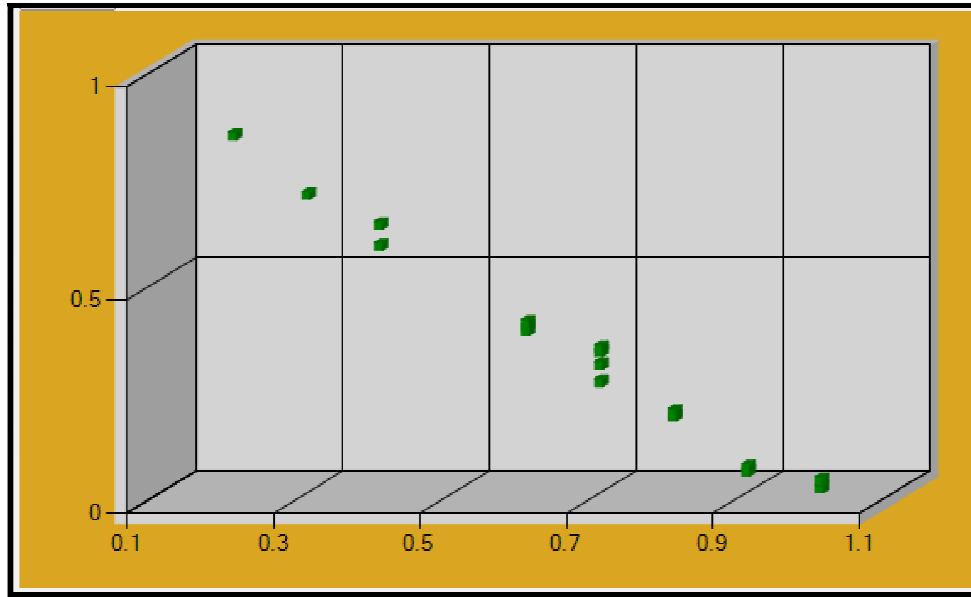


Fig. 16. Matrix U_{ij} chart

As we noticed that every points has value between 0 and 1 and X-axis is the complementation of Y-axis for example (0.98, 0.02) is the top point and so on.

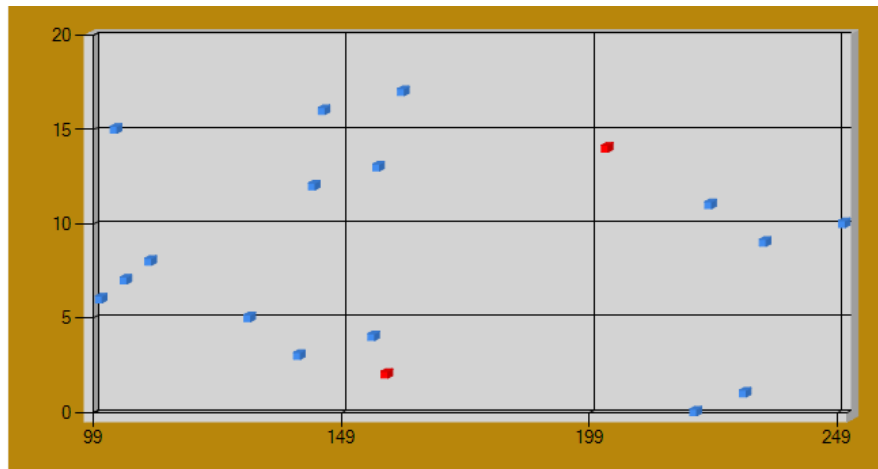


Fig. 17. The clusters chart

In Fig. 17, the dataset is classified into two groups, also the red points are centroids for each clusters. Fuzzy C-means is cleared in Fig. 17 because any data item can belong to more than one centroids.

7 Conclusion and Future Works

The future of the World Wide Web is going in the direction of SW, which is a rich source of data. In order to mine SW data, DM techniques can be used to analyse SW data from a variety of online sources to retrieve important information. Clustering techniques in DM can be used to cluster a huge amount of SW data into groups. This research explains the importance of applying FCM algorithm rather than hard (CRISP) K-means because in FCM data items belong to more than one cluster.

The proposed solution needed diverse steps and achievements. First step in this proposed system was the selection of SW dataset or the creation from different sources because I chose various datasets for different reasons such as online sources, datasets from imported external data or creating datasets in this system.

Second step was validation of semantic web datasets because I understand that every SW dataset is not valid. After checking the validity, the triples are produced.

Thirdly, the transformation of SW dataset to traditional dataset using C# language is required writing a set of codes to transfer SW dataset to traditional dataset. In contrast, SPARQL can be written to transfer SW dataset to traditional dataset.

Fifthly, the reason for applying the specific DM algorithm was to make the system more advanced.

Finally, the results of DM consist of the following three results; matrix U_{ij} , clusters and centroids of the clusters. The results include graphs of the results such as the scatter chart for matrix U_{ij} and the scatter chart for clusters and centers.

Future research should therefore concentrate on the investigation of detecting semantic web dataset automatically and convert it to traditional dataset. This will help user to avoid writing SPARQL or programming language for each semantic web dataset manually.

Another possible area of future research would be to evaluate SWM using FCM system utilizing extra-complicated semantic web data from various sources. This will examine the effectiveness and accurateness of SWM using FCM system.

Finally, implementing other types of the fuzzy algorithms on semantic web data. Only FCM clustering is applied. Other types of fuzzy algorithms data mining can be applied as well.

Competing Interests

Authors have declared that no competing interests exist.

References

- [1] Maedche Alexander, Zacharias Valentin. Clustering ontology-based metadata in the semantic web. s.l.: Springer Berlin Heidelberg; 2002.
- [2] Bettina Berendt et al. A roadmap for web mining: From web to semantic Web. s.l.: Springer Berlin Heidelberg; 2004.
- [3] Cai Jiaojie et al. Domain ontology mapping based semantic web mining models research. Atlantis Press. 2010;374-381.

- [4] Saraee Mohamad, Ramezani Reza, Nematbakhsh Mohammad Ali. SWApriori: A new approach to Mining Association rules from semantic web data; 2012.
Available:www.academia.edu
(p. under publication)
- [5] Chitraa V, Antony Selvadoss, Thanamani. Web log data analysis by enhanced fuzzy C means clustering. India: International Journal on Computational Sciences & Applications. International Journal on Computational Sciences & Applications (IJCSA). 2014;4(2):81-95.
- [6] Lakheyam Chhaman, Kaur Usvir. A survey on web usage mining with fuzzyc-means clustering algorithm. International Journal of Computer Science and Mobile Computing. 2013;2:160-163.
- [7] Niware Deepak Kumar, Chaturvedi Setu Kumar. Web usage mining through efficient genetic fuzzy C-Means. IJCSNS International Journal of Computer Science and Network Security. 2014;6(14):113-117.
- [8] Hebler John, Matthew Fisher, Blace Ryan, Perez-Lopez Andrew. Semantic web programming. s.l.: John Wiley & Sons; 2011.
- [9] Benz Brian, Durant John. XML programming bible. 1st. New York: Wiley Publishing, Inc; 2003.
- [10] Rod Stephens, Brian Hochgurtel. Visual basic.NET and XML: Harness the power of XML in VB.NET applications. New York: John Wiley & Sons, Inc; 2002.
- [11] Powers Shelley. Practical RDF. 1st. United State: O'Reilly; 2003.
- [12] Yu Liyang. A developer's guide to the semantic web. London New York: Springer Heidelberg Dordrecht; 2011.
- [13] Jain Kamna. D-RDF: Dynamic resource description framework. s.l.: Iowa State University; 2007.
- [14] Allemang Dean, Hendler James. Semantic web for the working ontologist. 2nd. Waltham, USA: Morgan Kaufmann; 2011.
- [15] McBride Brian, Brickley Dan, Guha RV. W3C; 2004.
Available:<http://www.w3.org/TR/rdf-schema/>
(Cited: 13 June 2014)
- [16] Paige Richard F. Theory and practice of model transformations. 2nd. Zurich, Switzerland: Proceedings of 2nd International Conference, ICMT; 2009.
- [17] Groppe Sven. Data management and query processing in semantic web databases. 1st. London, New York: Springer; 2011.
- [18] Prud'hommeaux Eric. W3C; 2013.
Available:<http://www.w3.org/TR/rdf-sparql-query/#basicpatterns>
(Cited: 17 06 2014)
- [19] DuCharme, Bob. Learning SPARQL. Simon St. Laurent. s.l.: O'Reilly; 2011.
- [20] Wolfram Stephen. Wolfram Mathematica; 2014.
Available:<http://reference.wolfram.com/applications/fuzzylogic/Manual/12.html>
(Cited: 20 June 2014)

- [21] Polimi. Politecnico Di Milano; 2014.
Available:http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html
(Cited: 21 June 2014)
- [22] Bora Dibya Jyoti, Gupta Anil Kumar. A comparative study between fuzzy clustering algorithm and hard clustering algorithm. *International Journal of Computer Trends and Technology (IJCTT)*. 2014; 10:108-113.
- [23] Singha Tanmoy. Clustering with fitness based adaptive differential clustering with fitness. Jadavpur, India: Jadavpur University, Kolkata-700032; 2012.
- [24] Saitta Lorenza. Information Society Technologies (IST). European Knowledge Discovery Network of Excellence. 2003;1-68.
- [25] David L. Olson, Dursun Delen. *Advanced data mining techniques*. 1st. s.l.: Springer; 2008.
- [26] Han Jiawei, Kamber Micheline. *Data mining: Concepts and techniques*. 3rd; 2011.
- [27] Tut Zen, Zen Tut; 2014.
Available:<http://www.zentut.com/data-mining/what-is-data-mining/>
(Cited: 4 July 2014)
- [28] Hotho Andreas, Sure York, Getoor Lise. Mining for and from the semantic web. 10th International ACM SIGKDD Conference on Knowledge Discovery and Data Mining KDD; 2004.
- [29] Quboa Qudamah K, Saraee Mohamad. A state-of-the-art survey on semantic web mining. *Intelligent Information Management*. 2013;5:10-17.
- [30] Zorrilla ME, et al. Web usage mining project for improving web-based learning sites. Madrid: s.n.; 2013.
- [31] Turban Efraim, et al. *Business intelligence a managerial approach*. 2nd. New Jersey: Pearson; 2011.

© 2016 Mohammed and Saraee; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:

The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)

<http://sciencedomain.org/review-history/14504>