Decision Support System to Decide The Right Marketing Strategy Using Data Mining, Semantic Web and Ontology

Denny Satria*
Gunadarma University
Jl. Margonda Raya 100
Depok, Indonesia
dennysatria@
rocketmail.com

Jessica Permatasari†
Gunadarma University
Jl. Margonda Raya 100
Depok, Indonesia
jessicapermatasari23@
gmail.com

I Wayan S. Wicaksana‡
Gunadarma University
Jl. Margonda Raya 100
Depok, Indonesia
iwayan@staff.
gunadarma.ac.id

ABSTRACT

With the development of the internet, retail companies were began to realize that as a critical factor to succeed their business, advertisement and promotion of a product on the web influenced the level of customer purchases. And it can served as a trace of customer shopping habits, the most important thing to consider is that potential data for retail companies are from the positive and negative response of customer. Therefore, the work in this paper is about our earlier work of concepts and techniques of semantic web and ontology to retrieve the data, then the data are store in a database and process to become data mining, which are useful in decision making. By using data mining and decision support system retail companies can always communicate a better offer to their client, and also can increase the level of client satisfaction provided by companies.

Keywords

Decision Support System, Database, Data Mining, Ontology, Semantic Web

1. INTRODUCTION

Nowadays, most companies have their own website to promote their company products and services aiming at developing their business. To recognize macromarketing trends and real-time customer needs, contents in customer websites are valuable asset. It can be potential data to make a right decision for companies to decide their marketing strategy with the aim to increase sales.

The enormous amount of data from the web made companies difficult to found, accessed, and presented relevant information, this is because information content are presented primarily in natural language. The data coming from the web are not just about which products customer needs, or about the advertisement and advancement of products, but what must be considered was customer complaints against the products. The problem is computer does not have human intelligence, for example is computer can not be distinguish the customer complaints or compliments words and it can be a different concept of words in other web pages.

Different semantic patterns used in various parties may have great challenges in the comprehension of the content in the websites. Therefore this paper presents a concepts and techniques of semantic web and ontology approach for business intelligence. By adopting and integrating semantic, ontology, database system, data mining,and rule-based decision support system, the marketing and sales performance information can be captured, clustered, compared, analyzed, integrated and disseminated. This leverages organizational marketing capital to support marketing plan and shorten the promotion planning time.

The topics covered are: the semantic web including standards such as RDF(Resource Description a Framework), OWL (Ontology Web Language), XML(eXtensible Markup Language), a widespread communication medium for the Web, database system, data mining and DSS(Decision Support System).

whole of the system will help DSS to produce more helpful information for companies. So that CEO(Chief Executive Officer) can decide the right marketing strategy and implemented that with better service to customers, promos in accordance with client expectation, as well as products development based on client buying trends. All this may add to the company's client loyalty and also increase the number of new client resulting in increase revenue of the company.

2. RELATED WORKS

Now the use of semantic approach for semantic web in solving the problems to describe various information in the web is growing rapidly and to manage query and semantic reconciliation in the distributed information environment, on this problem ontology can be used as an alternative way. Ontology can identify the concept of the information and

*Student of Sarmag Program, Gunadarma University.

†Student of Sarmag Program, Gunadarma University.

‡Center of Information System Studies, Gunadarma University.

relationship. As we know that web is commonly made for human's view, not machine's. This is why machine hardly describe the information which integrated which other information in the web. Therefore, in the future we can provide better understandable web for human and machine [1].

The primary aim is web can provide the relevant data to be process in the Decision Support System to help human to make a decision. In this part, we will see several related works which explain about Semantic Web in Decision Support System.

Jethro Borsje, Leonard Levering, and Flavius Frasincar [5] described in their paper "Hermes: a Semantic Web-Based News Decision Support System", a framework that provides decision makers with the ability to extract a set of news items related to specific concepts of interest. That is because decision makers need an information system that is capable of extracting a set of relevant news items automatically. And their goal from this paper accomplished by creating a knowledge base and developing a system that classifies news with respect to the knowledge base.

"Addressing Diversity Method Semantics in Information Between Suppliers and Manufacturing on Supply Chain Management" written by Lily Wulandari and I Wayan Simri Wicaksana [8], describe about the semantic web and ontology approach to find the appropriate information between supplier and manufacturing. for finding the appropriate information the authors use semantic web and ontology for the primary vehicle to cope diversity that will be implemented to achieve interoperability. This approach will be develop as prototipe for interoperability with semantic web and ontology use OWL based on RDF/S and XML/S. The primary Tools is Protégé for the development of ontology and sources representation, also the authors use virtual server (Apache) and virtual machine (QEMU) for the simulation.

"The evolution of Intelligent Computer Software and The semantic Web" written by Jens Pohl [4] describe about the vision of a Semantic Web environment in which ontology-based Web services with intelligent capabilities are able to discover each other and individually or in self-configured groups perform useful tasks, is not only feasible but imminently realizable. A distinction between human intelligences and component capabilities within a more general definition of intelligence, which may be embedded in computer software. The primary vehicle in the quest for intelligent software has been the gradual recognition of the central role played by data and information, rather than the logic and functionality of the application. The three milestones in this evolution have been:

- The separation of data management from the internal domain of the application
- The development of standard data exchange protocols such as XML that allow machine interpretable structure and meaning to be added to data exchange packages
- And, the ability to build information models that are rich in relationships and are thereby capable of supporting

Fabio Forno, Laura Farinetti, and Sean Mehan [2] in their paper "Can Data Mining Techniques Ease The Semantic Tagging Burden?", describe about the important to increase web accessible and investigate how an important and powerful data mining technique, Latent Semantic Indexing (LSI), might help in the design and implementation of tools that guide users in semantic tagging tasks. semantically tagged resources is to make available tools which allow users to explore and understand relevant ontologies and to present relevant categories with which to tag new data. So they applied LSI to a large portion of the Open Directory Project (ODP) catalogue and computed statistical information concerning category relationships in the ODP data set, also they incorporated structural information by modifying the construction process of the LSI space. based on that, they do comparative experiment where a machine generated classification of new documents was evaluated against a classification created by a group of human users.

For that various related work, we can compare one to each other, but one of the important concepts is Semantic Web and ontology,although that several of papers above implemented that in different method, but the key is the concepts of semantic approach.

3. SOLUTION APPROACH

This section will discuss the concept and technique of semantic web, ontology, data mining, and DSS, and also will be explained about:

- How to get data from the web?
- How the information obtain is store in the database ?
- How to measure positive and negative response from the customer?

This study was limited to the concept and techniques of how retail companies can build a system that can get valuable data from customer shopping habits in the web, which are useful in DSS to improve marketing strategies. There is figures too to describe that, in more conceptual.

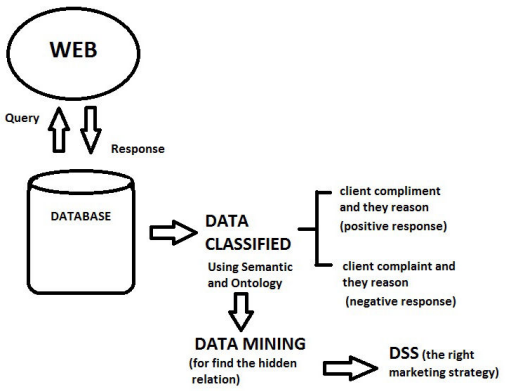

Figure 1: Database Collection and Data Mining Model using Semantic and Ontology

In the next part, will explain in detail about the performance of the system based on the image above and also about semantic web and ontology web languange.

3.1 Semantic Web

Semantic Web is a group of methods and technologies to allow machines to understand the meaning - or "semantics" - of information on the World Wide Web [6]. The original vision is the availability of machine-readable metadata would enable automated agents and other software to access the Web more intelligently. The agents would be able to perform tasks automatically and locate related information on behalf of the user.

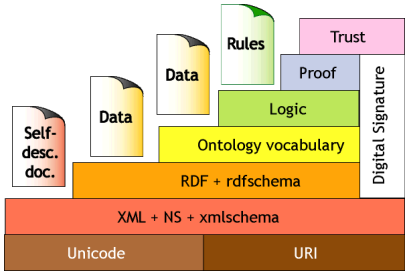

Figure 2: Semantic Web Architecture, courtesy of Tim Berners-Lee

From that picture, in a layered model, an upper layer takes as input the output of a lower layer. Hence, we should look to the lower layers in order to identify the root causes. Careful investigation of the layers in the Semantic Web Architecture demonstrates the importance of the bottom four layers because in these layers the entities (examples taken from real life) get described, modelled and implemented in various IT systems. It will be these entities which, when processed through the upper layers and enriched with for example metadata and logic, will serve for business integration or decision support system.

the term of "Semantic Web" is not formally defined it is mainly used to describe the model and technologies proposed by the W3C(World Wide Web Consortium). From picture 2 these technologies include the RDF, a variety of data interchange formats (e.g. RDF/XML, N3, Turtle, N-Triples), and notations such as RDF Schema (RDFS) and OWL, all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain. Many of the technologies proposed by the W3C already exist and are used in various projects. The Semantic Web as a global vision, however, has remained largely unrealized and its critics have questioned the feasibility of the approach.

3.2 OWL

Ontology is use for structuring information so that linkage information can be understood by machines. Ontology facilitates the uniformity and sharing of domain knowledge by four basic modeling primitives: concept, relation, instance and axiom [Gruber 1993; Neches et al. 1991].
OWL uses the following basic semantic elements to define ontology[3]:

- Class describes concept by a set of individuals (instances) or other existing classes.
- Property describes binary relation. It has two types: ObjectProperty specifying the relation between individuals of the same class or different classes, and DatatypeProperty indicating the relations between individuals of classes and RDF literals and XML schema datatypes.
- Restriction and characteristic describe constraints on relations and axioms. It includes the following semantic elements: allValuesFrom, someValuesFrom, Cardinality, hasValue, TransitiveProperty, SymmetricProperty, FunctionalProperty, InverseOf, and InverseFunctionalProperty.

3.2.1 Data from The Web

There are two kind of data, primary and secondary data, primary data is data obtained directly from respondents, while the secondary data obtained from other parties. The internet can help company to get the secondary data with source from research company websites[9].

For get the data from the web, first the system from the database will send a query. Query will be send to :

- URL (Universal Resource Locator)
  Every company have their own website, with the query send as a bait, the system will get the data of customers response about the company products.
- Search Engine
  Search engine have a components use to help to get the data, the components are spider (retrieve web pages found), crawler (evaluate link), indexer(analyze elements), database(save the data), result engine (to give ratings), and web server (serve the user). Now appear meta search engine that have a same function, but it combine the index from search engine.
- Directories/Web Catalogue
  Directories or web catalogue just took the important index from a website, in it covered specific topics, so the directories more effective than search engine.
- Surfing
  the purpose of surfing in the web is to get the information we needs. In marketing world, company needs the customer data for the benefit of marketing research. Tools support for this method are : data mining (save enormous amount of data with method classification, clustering), Akamai (use to prevent disturb or crush in data mobilization), Google analytic (detect how many visitors in the web, which part or web pages they like to visited, etc.)

3.2.2 Database

Database use to store the data customer response, in this part the author use the concept of R2D (RDF to Database)[7]. R2D aimed to transform RDF data, at run-time, into an equivalent normalized relational schema, thereby bridging the gap between RDF and RDBMS concepts and make the abundance of existing relational tools available to RDF Stores.

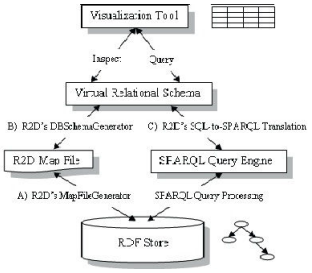

Figure 3: R2D System Architecture

3.3 Classified Data

In the process of classification, knowledge base will be implemented on the system to find the responses from customers, and also synonyms and antonyms of customers responses, and looking up words that denote the concept with the help of WordNet, semantic, RDF and OWL. after that everything will be add to the concept identifier and then become input for the classification system and the results will be stored in data mining [5].

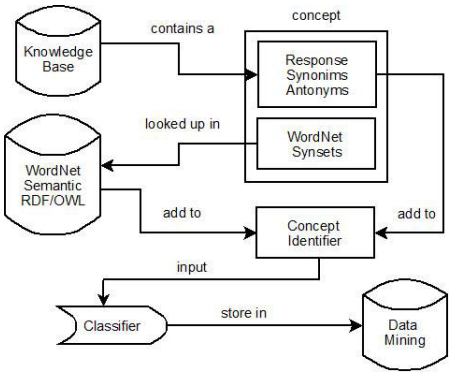

Figure 4: Overview of the classification process

3.4 Decision Support System

With the concept of a system that already described above, it is hoped the company could obtain accurate data and also real time on customer shopping habits. The data obtained will use in data mining research in the marketing research company whose purpose will use in the DSS. DSS is a system that helps in terms of decision-making, but the system is simply a tool that supports the CEO or a marketer to determine the right and better marketing strategy, so the company can always communicate a better offer to their client, can increase the level of client satisfaction provided by companies and the most important is companies can increase their profits.

4. CLOSING

To recognize macromarketing trends and real-time customer needs, a system that gives a real time, accurate and appropriate information is needed. The contribution knowledge of this paper is the system give concept and technique to decide the right marketing strategy using semantic, ontology, DSS and data mining. For now, our research has only reach the stage of conceptual design, the next problem that might occur in this system is the internet is not always present accurate and real-time data of customer response or customer shopping habits for company.
For the further work, the system will integrated with another support tool to make DSS more powerful.

5. REFERENCES

[1] S. W. Adriani Yulida Kusuma, Faiz Chandra Ardian and I. W. S. Wicaksana. Gathering flight information for decision support system using semantic-web based approach. page 2, 2009.

[2] S. M. Fabio Forno, Laura Farinetti. Can data mining techniques ease the semantic tagging burden. page 1.

[3] Y. X. HAI ZHUGE and P. SHI. Resource space model, owl and database : Mapping and integration. *ACM*, 8(4):3, 2008.

[4] P. Jens Pohl. The evolution of intelligent computer software and the semantic web. *InterSymp-2004, Baden-Baden, Germany, Jul.29-Aug.5, 2004. RESU84*, page 1, July - August 2004.

[5] L. L. Jethro Borsje and F. Frasincar. Hermes : a semantic web-based news decicion support system. *SAC'08*, page 1 and 3, March 2008. http://finance.google.com/finance.

[6] M. R. Koenraad Vandenborre, Charles Milligan and H. Tromp. Challenges in semantic web technologies : A first iteration towards a root cause analysis. *PICMET*, page 1, July 2006. PICMET 2006 Proceedings, 9-13 July, Istanbul, Turkey.

[7] L. K. S. S. Sunitha Ramanujam, Anubha Gupta and B. Thuraisingham. R2d: A bridge between the semantic web and relational visualization tools. *IEEE*, page 1 and 2, 2009. IEEE International Conference on Semantic Computing.

[8] L. Wulandari and I. W. S. Wicaksana. Addressing diversity method semantic in infomation between suppliers and manifacturing on supply chain management. *1st Seminar on Application and Research in Industrial Technology, SMART*, page 1, 2006.

[9] www.wikipedia.org. Semantic web.