Project Proposal

Team 9

Project Title

EasyShopping – Semantic Web Application

Team members and roles

1. **Pushkar Ladhe** – Researching FMCG websites for datasets and scraping
2. **Tushar Pandit** – Researching FMCG websites for datasets and scraping and UI design.
3. **Mukulsingh Jadhav** – Front End development
4. **Pranjal Karankar** – Front End Development and Crawling

Description of the problem and application that you are proposing to create

In student life, we often try for save a dollar or two while go on the weekly shopping tour. While trying to save, we usually neglect the nutrients information of the product and end up buying cheap product. However, in most cases a better product can be obtained for same or lesser price from some other store. Wallmart and safeway are two major giants in the grocery field. Hence, we are planning to develop a price comparator for daily grocery products along with the feature to display their nutrient information. This way, when next time one goes to shopping, can get a better product for lesser price.

Literature survey (include paper/resource list that you plan to read)

We plan to study ontologies, message content and message protocols for the project. We have sorted out some papers that we think would be useful in our studies. One of them is the paper by David Trastour, Claudio Bartolini and Chris Preist on Semantic web support for the business-to-business e-commerce lifecycle. We have gone through the roles of XML and RDF paper by S. Decker, F. van Harmelen, J. Broekstra, M. Erdmann, D. Fensel, I. Horrocks, M. Klein and S. Melnik which speaks about the importance of XML in semantic web design.

Project plan

We plan to follow waterfall process for development of this project. Hence, as a part of initial phases of the model, we are currently documenting requirements and performing market survey to know more about the existing applications in the same domain.

In the second half of October, we plan to complete software specification and requirement(SRS) and design document of the proposed application. By the end of the October, we are planning to crawl the required data and create RDF and ontologies based on the scraped data. With the start of the November, we plan to start developing user interface of the application.

We are planning to use JavaScript, NodeJS, HTML, CSS, SemanticUI technology stack for the development of the web application. We plan to use Python beautiful soap, Python scrapy for scraping of the data from websites.

Datasets (with source)

1. Google places API to get data about location of the Safeway, Wallmart stores
2. Scrap the website of the Wallmart to fetch grocery items using Python library- Beautiful Soap
3. Scrap the website of the Safeway to fetch grocery items using Python library- Beautiful Soap
4. USDA API to fetch information about the nutrients < https://ndb.nal.usda.gov/ndb/doc/index>