Asia Pacific Journal of Contemporary Education and Communication Technology (APJCECT)
ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

APIAR

Asia Pacific Journal of Contemporary Education and Communication Technology
ISBN (eBook): 978 0 9943656 8 2 | ISSN : 2205-6181
Year: 2017 , Volume: 3, Issue: 1

ADSEEKER - PERSONALIZED ADVERTISEMENT ENGINE BASED ON SOCIAL MEDIA

W.P.A.C.Perera [a], P.S.Jayasinghe [b], V.G.S.Pramudith [c], D.H.A.De Silva [d], D.Kasthurirathna [e], T.D. Perera [f]
[abcdef] Sri Lanka Institute of Information Technology, Malabe, Sri Lanka
Corresponding email: hansi.desilva1129@gmail.com

Abstract

This paper reports evidence that AdSeeker, user preference based advertisement engine based on social media is the most applicable solution to improve the business value of the marketing and advertising. In the modern world, people tend to post their advertisements via internet, television and so on. Since the internet is using vast no of people, it essentially needed a comprehensive method to push right advertisements to the right people. To improve the accuracy of advertisement pushing mechanism it is used ontology system, advertisement classification and data mining and warehousing with machine learning concept. By identifying personal relationship hierarchy, Ontology-based product amount mentioned algorithm are most applicable algorithms which are used to identify user preferences hierarchy. According to the tweets, people who posted on Twitter get the actual preferences of each and every user separately. According to their preferences push advertisements to them. This mechanism is most applicable to identify users and efficient way to advertisers to push advertisements to the right people.

Keywords: Ontology; Data Mining; Data Warehousing; RelationshipH; Weighted Scoring

1. Introduction

Social media are computer-mediated tools that allow people and organizations to create, share, or exchange information, career interests, ideas, pictures/videos in virtual communities and networks (Authority, 2016). It facilitates the development of online social networks by connecting a user's profile of other individuals and/or groups.

Since people posts everything in social media, it becomes a practical way of promoting advertisements via internet. By using users' profile easy to gather personal information such as interests, likes, hobbies, thoughts and so on. Therefore, social media gives a huge impact to the marketing and advertising fields. Most of the commercial advertisements web applications do not based on users' actual preferences. They do not have social media related advertisements classification system, self-updating character profile more personalization and preference based advertisements pushing mechanism. Therefore users, advertisers and corporations cannot get the comprehensive advantage. In this paper, we present the most efficient way to promote your advertisements to the right people, at the right time as well as advertisements pushing mechanism based on user preferences. We used ontology classification, data mining and

ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

warehousing and advertisements classification mechanism to achieve this target in our AdSeeker advertisement engine.

2. User Influence in Twitter

Twitter is one of the most famous social media around the world. Users in Twitter have different influence level on the users about various topics. Normally, Twitter users follow each other to getting information of what they posted on their accounts. When it comes to the highly followed users' profiles and things they post about the effect the massive influence on their followers. There is a huge advantage you can gain through social media-based advertising. A trend can be initiated by anyone, and if the environment is right, it will spread. The following three activities represent the different types of influences of people.

1. In-degree influence, the number of followers of the user. By using this you can directly identify the audience of the particular user.
2. Retweet influence, which we measure through the number of retweets containing one's name.
3. Mention influence, which we measure through the number of mentions containing one's name, indicates the ability of that user to engage others in a conversation (Tao, Abel, Hauff & Houben, 2015; Cha et al., 2010; De Silva & Riloff, 2014).

3. Ontology

Ontology is a set of concepts and categories in a subject area or domain that shows their properties and the relations between them (wikipedia, 2016). When it comes to advertising, the system will need domain knowledge (Technology, Automobile, and Fashion) of ontology to identify the contexts of the social media updates. Ontology is very useful for identifying different categories of a domain knowledge. We can use open source ontology building software like Protégé to build massive ontologies with a relationship that is suitable to implement the analyzing system and context identifying modules. After building ontologies when analyzing social media updates the system can correctly identify the contexts of the social media updates and build the user character (preference) profile according to those social media updates. User character profile means a profile that contains the products and the brands that user has explicitly mentioned in their tweets. We built this, after analyzing all the tweets (Malesh & Nirenburg, 1995; Madubashana, 2016). The following figures explains ontology and onto graph using Protege tool.

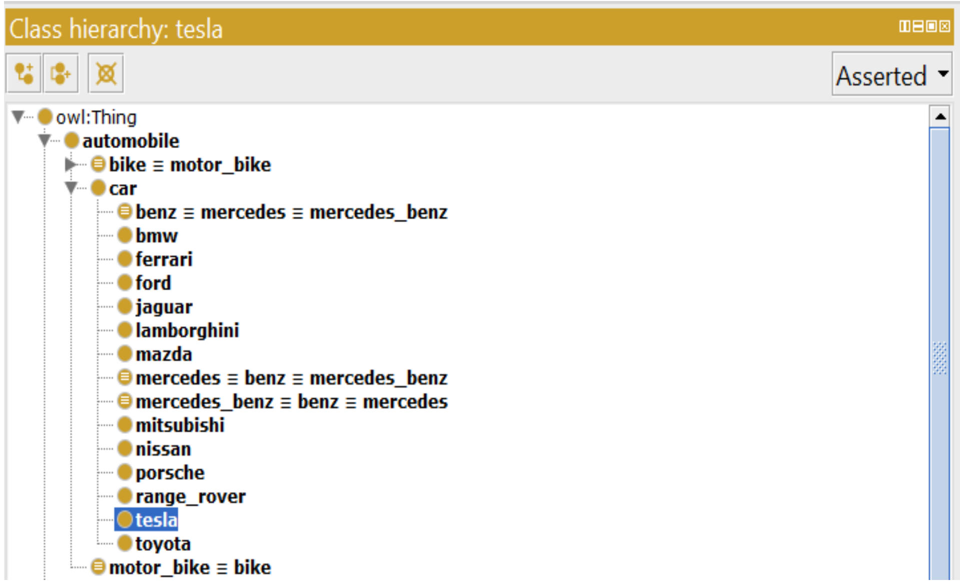


Figure 1: Ontology

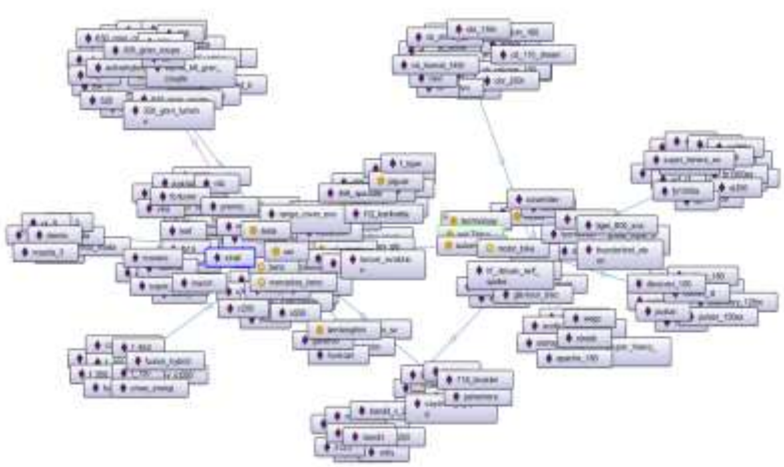


Figure 2: Onto Graph

4. Methodology

Our proposed advertisement engine built using JAVA JAX-RX Restful API web services, Apache Jena which is used to connect and query with an ontology, hibernate to deal with database, MySQL Server is used as the database, sentiment analysis (wikipedia, 2016; Apoorv et al. 2009; Adedoyin-Olowe, Gaber & Stahl, 2013) using Naïve Bayers (Apoorv et al., 2009) and used NLTK libraries. To demonstrate the advertisement engine, we used PHP web application with Codeigniter framework. Overall architecture explains by using below figure 1.

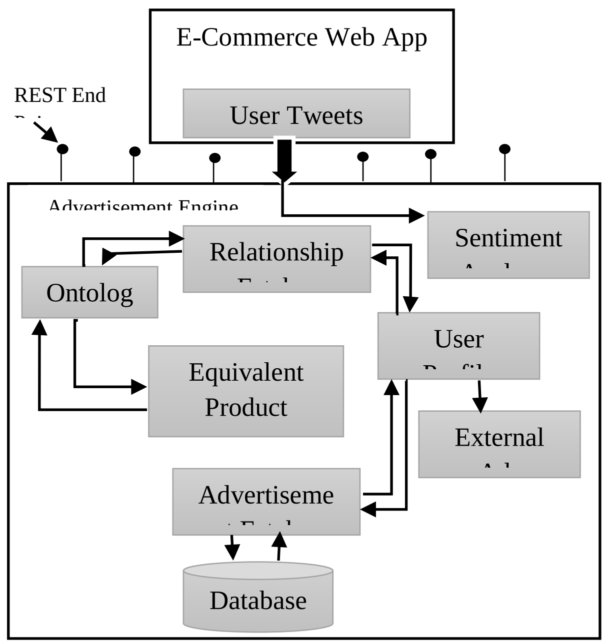
Asia Pacific Journal of Contemporary Education and Communication Technology (APJCECT)
ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

4.1 Build an Ontology Domains

Since this research is based on user preference and is based on advertisement pushing mechanism, we had to develop ontology domains using Protégé software. In this research, we mainly deal with Technology and Automobile domains. As a class hierarchy, we added accurate details to the ontology.

Ex: - technology->computer-> VGA-> titan x

As the above example added individual "Titan x" details to the relevant ontology hierarchy. Likewise, by adding relevant classes and subclasses, we built the 2 domains of ontology.



Social media is used by a vast number of people in the world. From these, Twitter can be placed in a higher place because many people are using twitter all around the world. To get twitter data, user should have been logging with twitter. By using twitter ID of a user, gather twitter to the JAX-RX Restful endpoints. After that check the user's existence in the database.

4.3 Get the relationship hierarchy (If user exists in database)

If the user exists, retrieve user's personal preference relationship hierarchy. Here, we saved the user product mention profile on the server. We save the preferences of the user in user product mention profile. In the user's mention profile, (Refer section E) it contains how many times the user has mentioned about his/her preferred products and brands.

4.4 User does not exist in database

If the user is a new user (not exist in the database), get all the tweets from users account and send those tweets to the sentiment analysis web services to get positive tweets.

Asia Pacific Journal of Contemporary Education and Communication Technology (APJCECT)
ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

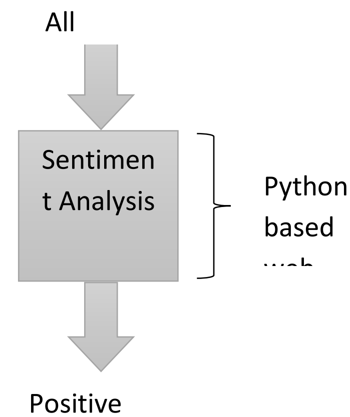


Figure 4: Sentiment Analysis

After fetching all the positive tweets as in Figure 2, send them to "FilterWordList" to filter words. Then send all remaining tweet variants to the ontology to get the relationship hierarchies. Then, send all the relationship hierarchies to get user preference mention profile algorithm as shown in step E.

Imagine a user has tweeted about GTX 1080, our system can identify GTX 1080 is a VGA card, identify its part of a computer category and its part of the technology section. Then it can get related product to the GTX 1080 like R9 490X, Titan X which are similar to (Suggesting similar products for refer section H) GTX 1080. All these details can be derived from the ontology.

Sample Relationship Hierarchy,

Technology	Automobile
Mobile Phones	Cars
Apple	Tesla
IPhone 6S	Model s

4.5 Product Amount Mentioned Algorithm

To build the product amount mentioned algorithm, we followed following steps. First of all, we change the ontology as each and every class and individual has a property called "isTypeof" and it has four types such as product, brand, product types and top category.

Example:-
Technology → Top Category
Mobile Phone → Product Type
Apple → Brand
IPhone 6s plus → Product
By analyzing positive tweets and according to the number of mentions we got a basic information profile from the tweets. Imagine a sample tweet as "I really want to buy the new iPhone 6s and new model S". After analyzing we can get a mention profile as follows.

Table 1: User mention profile

		No of Occurrences
Product	iPhone 6s	1
	Model S	1
Brand	Apple	1
	Tesla	1
Product Type	Mobile Phone	1
	Car	1
Top Category	Technology	1
	Automobile	1

In our system, we have several number of mentions profile as above Table 1 for all the tweets. Imagine if one person has tweeted 2 times that he wants to buy Model S. Then, Model S count increases and all the hierarchy related to the Model S also increases. According to user's mention profile, we can obtain preferred products and hierarchies.

Table 2: Sample product mention profile

		No of Occurrences
Product	iPhone 6s	1
	iPhone 6s plus	2
	Model S	5
	GTX 1080	1
Brand	Apple	3
	Tesla	5
	NVidia	1
Product Type	Mobile Phone	3
	Car	5
	Computer	1
Top Category	Technology	14

	Automobile	5

4.6 Advertisement fetcher algorithm

Relationship hierarchy object which is saved from above steps C, D and E is needed to send to the advertisement fetcher algorithm in order to fetch how many advertisements we need. By using this algorithm, it is going to fetch advertisements according to the product, brand, product type, and top category. The below figure 3 represent the basics of advertisement fetcher algorithm.

Advertisements fetcher

({Relationship object}, {Total advertisements

Figure 5: Advertisement Fetcher algorithm

After analyzing the profile, (as an example table 2) identify how many advertisements we want to fetch from the database in order to display to the user. The top to bottom categories which mentioned in Table 1, we assign more priority to the Products section and then Brand, Product Type AND Main Category respectively. The reason of giving highest priority to the product section is the user has explicitly mentioned that he/she likes those products.

Imagine the situation that the home feed has requested 20 advertisements. If the product category doesn't contain at least 10% of the mentions from the requested advertisements (if the system has requested 20 advertisements, at least 2 products should be there (10%) to get advertisements related to the product), next it goes to the brand's section and fetch advertisements from it.

Example:-

Product → iPhone 6s → 1

Brand → Apple → 1

Tesla → 3

Now, the system fetches 50% of the requested advertisements which are related to the product and 50% related to the brands. According to the above scenario, system will fetch 10 advertisements related to iPhone 6s and 10 advertisements related to the tesla products. If the brand section doesn't contain 10% mentions of the requested advertisements it moves to the product type category and fetch 33% from the product section, 33% from the brand section and 33% from the product type section. If product type section doesn't contain 10% mentions of the advertisements' count, it fetches 25% percentage of advertisements from each section.

4.7 Advertisements Classifier Algorithm

Advertisement classifier algorithm gets relevant advertisements from the database according to the country and internal advertisement distribution function. Advertisement distribution function means distribute of advertisements according to the country.

ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

Example: - If you live in Sri Lanka, you mostly prefer Sri Lankan advertisements. Because marketing and business value will improve if we push advertisements to the right person at the right time. Therefore Sri Lankan people getting advertisements from Sri Lankans are more profitable.

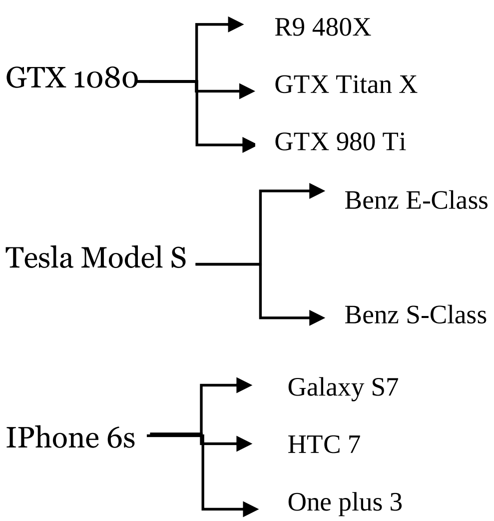
After fetching related advertisements from the database, push those from the end point to the relevant user. The above steps can continue if the user exists in the database (Step C, D, and E).

4.8 Suggesting similar products

It is more efficient if the system could suggest similar products. In our ontology system, we have mapped similar range products for each item.

As an example,

The GTX 1080 can be mapped similar price range and similar performance categories.



Imagine a user has tweeted about GTX 1080, our system can identify GTX 1080 is a VGA card, identify its part of a computer category and its part of the technology section. Then, it can get related products to the GTX 1080 like R9 490X, Titan X which are similar to GTX 1080. All these details can be derived from the ontology.

In our comprehensive system, we implemented admin panel from our web application. Because, by using this, it is easy to add individuals and classes to the ontology. Every 7 days, the user mention profile automatically updates. Old profile is moved to the different section of the database with a timestamp. Therefore. we can get reports regarding changes of preferences of the user.

Each search query the user searched (advertisements) save in the database along with the user profile and they also get archived every 7 days same as user character profile. By using this procedure we can push advertisements to the home feed of the user. Below figure explains the similar products of a specific product.

Figure 5: similar products

5. Survey

To prove the accuracy of our advertisement engine we launched a survey to compare no of relevant advertisements on eBay home feed and no of relevant advertisements on AdSeeker home feed. Therefore, we did a survey with 20 people and proved that our home feed provides around 90% relevant advertisements to the user. The following table represents the sample results of 5 people who are participated with the survey.

Table 3: Survey Results

Person No	**No of relevant advertisements in eBay home feed (out of 24)**	**No of relevant advertisements in AdSeeker home feed (out of 20)**
1	5	18
2	4	19
3	5	19
4	6	18
5	6	19

Asia Pacific Journal of Contemporary Education and Communication Technology (APJCECT)
ISBN: 978 0 9943656 82; ISSN: 2205-6181
Year: 2017, Volume: 3, Issue: 1
www.apiar.org.au

Figure.6: publish advertisements according to user's tweets

6. Conclusion

In this paper, it is being discussed the influence of the social media to improve the accuracy of our AdSeeker advertisement engine. This advertisement engine proves that user preference based advertisement engine is more efficient way to increase the business value. Users, advertisers as well as corporations have many advantages from this advertisement engine.

However, the user with a significant level of understanding about sentiment analysis, ontology, machine learning and social media can improve and develop this system and, further improve the accuracy of the AdSeeker system.

Currently, using commercial web applications and advertisement engines provide results, based on previously searched queries most of the time. Therefore, it may not be the actual preference of the user. But social media based systems provide more accurate results.

Preference results are further fine-tuned in order to increase the accuracy level of the system. Therefore it becomes a more usable and profitable product. In future, we hope to continue to increase the accuracy level of the AdSeeker system.

References

i. Adedoyin-Olowe, M., Medhat Gaber M. & Stahl, F. (2013) *A Survey of Data Mining Techniques for Social Network Analysis*. School of Computing Science and Digital Media, Robert Gordon University, 2013, pp 1-25.

ii. Apoorv, Agarwal, Boyi, Xie, Ilia, Vovsha, ... Passonneau (2009) *Sentiment Analysis of Twitter Data*. New York: Department of Computer Science Columbia University New York.

iii. Authority, B.(2016) ‘ Social and vanity’/ ‘Social media’. *Wikipedia*. [Online] Available at: https://en.wikipedia.org/wiki/Social_media Accessed: 5th August 2016.

iv. Cha, M., Haddadiy, H., Benevenuto. F., Krishna, P. & Gummadi, (2010) *Measuring User Influence in Twitter: The Million Follower Fallacy*. Germany: Max Planck Institute for Software Systems.

v. De Silva, L. & Riloff, E. (2014) *User Type Classification of Tweets with Implications for Event Recognition*. Utah: School of Computing University of Utah

vi. Madubashana, W. (2016) *NLP-Reduce: A “naive” but domain-independent natural language interface for querying Ontologies.*

vii. Mahesh, K. & Nirenburg, S. (1995) *A situated Ontology for practical NLP*. USA: Computing Research Laboratory.

viii. Tao, K. Abel, F., Hauff, C. & Houben, G-J (2015) *What makes a tweet relevant for a topic*. Delft: Web Information Systems.

ix. Wikipedia (2016) ‘Naive Bayes classifier.’ [Online] Available at: https://en.wikipedia.org/wiki/Naive_Bayes_classifier. Accessed 25th August 2016.

x. Wikipedia (2016) ‘Ontology.’ [Online] Available at: https://en.wikipedia.org/wiki/Ontology. Accessed 15th August 2016.

xi. Wikipedia (2016) ‘Sentiment analysis.’ [Online] Available at: https://en.wikipedia.org/wiki/Sentiment_analysis. Accessed: 25th September 2016