IFERP

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol4, Issue 4, April 2017

Recommendation Engine in E-Commerce

[1] Shrinivas S Hegde, [2] Soumya Lakshmi G, [3] Vellaki Kiran Kumar, [4] Neelu L, [5] Pramodchakravarthy Patil
[1,2,3] UG students,Dept. CSE, RajaRajeswari College of Engineering
[4] Assistant Professor, Dept. CSE, RajaRajeswari College of Engineering
[5] Manager-Integrations, Starmark Software

***Abstract*— Recommender systems areimportant for businesses because they can help companies offer product recommendations to customers. There are many acknowledged consumer-oriented recommender systems, particularly in e-commerce. A client-product matrix is built which matches company clients to internal company products. Pairing of clients and products is based on co-clustering principles and reveals potential future purchases. Compared to other consumer-oriented recommendation systems, this approach takes into account the need for interpretability. The recipient of the generated recommendations are sales and marketing teams; hence, a detailed reasoning will be done in straightforward English that considers multiple aspects regarding why a client may be a suitable match for the particular offering.**

***Keywords*-Co-clustering, B2B(business-to-business).**

I. INTRODUCTION

Recommendation systems assist the end users with their interactions. The main advantage of recommendation systems are that the user can benefitted from the vast increasing data rather than overwhelmed by it. Within a recommendation platform, the appropriate information is reached to the user based on anticipated needs instead of the user seeking the information. In E-commerce websites, the recommendations can help point the user to the objects that they may need or want in the near future, without browsing through the long catalog of items. Recommender systems have found many useful applications in online retail shops or when purchasing video and music content. The recommendation system is not targeted for consumers, but will be used by the sales and marketing teams of a company and allows them to work more efficiently.

There are three distinct differences of B2B recommender systems, compared to consumer-oriented systems. Firstly, recommendations are not generated for individual consumers but for small or large companies and organizations, who are the clients of an organization. For client entities, a wealth of extended information can be extracted from Internet news, financial reports, etc. which creates a more holistic and accurate client model by examining not only the purchase history, but also information that spans almost every fact of the client operations. Secondly, the recommendations are created for the sales and marketing teams, not for the direct recipient of the product.

The sales or marketing person will evaluate the validity and reasoning of the recommendation and decide whether or not to contact the client about the product offering. This shows an imperative need for interpretability of recommendations. Finally, the value of the product or offering is not fixed; it varies based on the client's infrastructure and size. Product offerings in a B2B application may be a combination of hardware, software, or service offered, and typically is worth thousands, sometimes millions, of dollars. The need for a detailed and grounded reasoning becomes immediate and apparent. This manuscript, addresses the second and third point, namely interpretability and value estimation in a B2B setting. Interpretability is particularly important because it infuses trust in the recommendation process and is also important in a B2C (business-to-consumer) setting. But in a B2B scenario the provided rationale has to be substantially more detailed. A methodology is created that generates interpretable recommendations based on co-clustering principles. The offered recommendation value is determined, as it can assist a salesperson to better prioritize the generated sales opportunities. A few other things done here are as follows. An overview of our methodology is provided. Later, co-clustering technique can be used to generate candidate product recommendations for the clients. Also a rank is provided to the recommendations by buying confidence and also by the selling value estimate using Pareto principle. Subsequently, an elaboration on how to generate textually interpretable recommendations using the previous pieces is given. Empirical evidence about the usefulness of our approach compared to other recommendation techniques is also provided, and a deployment method is suggested. Finally, we examine connections and differences of our work in the context of related research.

II. LITERATURE SURVEY

Recommender systems are powerful tools that can assist end-user in their interaction with big-data. The strength of recommender systems stem from the fact that the user can benefit from the ever increasing the data rather than overwhelmed by it. This is the case, because within the recommendation platform, the appropriate information is reaching the user based on anticipated needs instead of user seeking the information.

E-commerce refers to the network as the carrier, the use of digital electronic means to carry out business activities.

With the developing of internet technology and the Web technology, various types of e-commerce sites are coming out. It is not difficult to establish e-commerce sites, but to obtain economic and social benefits. On the e-commerce web site, there may be millions of on-line transactions, generating a large number of documents and records of the registration form every day. E-commerce companies are faced with a wealth of data, lack of knowledge of embarrassment. How to make the best use of these rich data make the e-commerce more effective? Do these data mining and analysis in order to fully understand customer preferences, buying patterns, design to meet the needs of different customer groups, personalized Web site to enhance their competitiveness, must have been. To generate interpretable recommendations, the core of our approach is based on the application of a co-clustering methodology on the client-product historical buying matrix. Through co-clustering the client-product buying matrix, we will identify groups of clients that buy sets of products. Naturally, these co-clusters may be overlapping. After these groups, or co-clusters, are identified, the products in each group not already bought by a client constitute good candidates for recommendations. The client-product pairs of candidate recommendations will be further evaluated and ranked. For example, a very big and dense co-cluster with only one product not already purchased should be ranked higher than a small co-cluster containing many non-purchased products. Similarly, if the recommendation refers to a client with a high turnover or belonging to an industry with high growth, such a recommendation merits a higher ranking than another with less favourable traits. For each client-product pair, a final purchase propensity is derived based on determining factors such as the above. Once the recommendations, the corresponding co-clusters, and the factors influencing the final purchase propensity are identified, they can be transcribed into a textual and visual format that is easily consumable by the sales representatives. To assist the representatives in better prioritizing the recommendations, in addition to the overall purchase propensity, we also provide a prediction for the value of the recommended sale opportunity. This is another difference from traditional consumer-oriented recommender systems where the value of the product is typically fixed and plays little role in the recommendation process. Given the recommendation propensity and the predicted monetary value, all recommendations are ranked using a Pareto principle to favour sales opportunities that hold both high value and high likelihood of purchase.

III. SYSTEM ARCHITECTURE

In this section, the general architecture of the recommendation is explained. In depth this design explains how the actions take place between two business organizations (B2B) i.e., the main organization entity and the other organization sales and marketing

team. The main organization entity will store the products in the cloud and that will be viewed by the various companies sales and marketing team. The main storage will be done in the cloud. Cloud computing systems are internet based, where the access is fully dependent on Internet connection The main advantage of storing data in cloud is, it is the data will be remotely maintained, managed, and backed up. Since it is B2B, it allows the users to store vast files online and it can be accessed from any location via Internet. But the main disadvantage is the downtime. No cloud provider, even the best would claim immunity to service out charges. Later while viewing the products online, recommendations will be done. It will be based on the client-product matrix of co-clustering principle. Co-clustering principle is a technique which allows simultaneous clustering of the rows and columns of a matrix. In our approach, the rows represent clients and the columns represent the products. Based on this matrix the recommendations will be done to the user according to his previous searches and historical purchase information.

***Figure: Architecture design of recommendation engine*.**

Initially the generated recommendations from the co-clustering will need to be weighted further depending on other characteristics of the client and the product, aside from historical purchase information since only the densest co-clusters are recorded and the sparse ones are discarded. So, depending on the available information about the client and the product, a recommendation will receive a final score that captures the overall confidence, or propensity, that the client will purchase the product. First we extract candidates for recommendations based on the purchase history and then re-rank the recommendations based on client and product features. When a sales representative wishes to further explore a recommendation on the confidence-value graph, then a textual interpretation of the reasoning is generated.

Some different features which can be transcribed into a textual format are:

• Co-cluster features: The co-cluster that led to the generation of the recommendation consists of clients and products.

The recommendation is created for a product not-yet-bought by a client within the co-cluster. The set of products typically bought together, as well as the clients that have already purchased the specific product, can convey useful information to sales and marketing teams. A sample output of such a textual interpretation can be: "Client X has purchased Product 1 and Product 2. Clients with similar purchases (e.g., client Y, client Z, and client W) also bought 'solution A'." Note that co-clustering is performed across all clients within a country, as well as per industry within the same country. Therefore, a rule such as the above can be generated either as a result of a country-wide or an industry-wide buying pattern.

• Client features: Financial characteristics of the client have been used as input parameters in the logistic regression that determined the confidence that a client will buy a specific product. Such client features can be transferred in a textual form as follows: "Client X belongs to the top 3% of the clients in terms of annual revenue."

• Product features: Clients in different industries (automotive, IT, banking, etc.) purchase different combinations of products. This is another feature that plays a significant role when building the confidence of a purchase and can be translated in textual format as follows: "Analysis Software B is the 2nd most purchased product in the automotive industry."

IV. MODULES

Thesystem mainly consists of five intermediate processes or modules. The following are the list of modules.

- Registration
- Purchase
- Capture User Behaviour
- Recommendation
- Implementation on cloud

Step by step execution of these modules produce the final encrypted file which containing the text. These modules will process the data concurrently in order to produce the output in encryption and decryption.

A. Registration

Registration is the initial stage where stakeholder or vendor has to register their account in the website. Without the registration one cannot enjoy the facility provided by the website.

In computer security, logging in or signing in is the process by which an individual gains access to a computer system by identifying and authenticating themselves. The user credentials are typically some form of 'user name' and matching 'password'. These information should be confidential in order to avoid cybercrime. The modern secure systems also often require a second factor for extra security. When access is no longer needed, the user can log out or sign out from one's account.

Logging in is usually used to enter a specific page, which trespassers cannot see. Once the user is logged in, the login token may be used to track what actions the user has while connected to site. Logging out may be performed explicitly by the user taking some actions, such as entering the appropriate command, or clicking a website link labelled as such. It can also be done implicitly, such as by the user powering off his or her workstation, closing a web browser window, leaving a website, or not refreshing a webpage within a defined period. Logging out of a computer when leaving it is a common security practice, preventing unauthorized users from tampering with it. There are also people who choose to have a password-protected screensaver set to activate after some period of inactivity. Requiring the user to re-enter his or her login credentials to unlock the screensaver and gain the access to system.

B. Purchase

Online shopping is the form of electronic commerce which allows consumers to directly buy goods or services from seller over the internet using a web browser. Consumer find a product of interest by visiting the website of the retailer directly or by searching among alternative vendors using a shopping search engine, which displays the same product's availability and pricing at different e-retailers. Online purchasing evokes the physical analogy of buying products. When an online store is set-up to enable business to buy from another business, the process is called business to business online shopping.

Once the user is successfully logged-in he may look for items to purchase. The recommendations come into picture

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol4, Issue 4, April 2017

now, where the manipulated item is recommended for unique user.
The user recommendation is based on the captured user behaviour. User selects the category and sub category until he finds his desired product. The application gives the confidence level as well as the value estimation for the recommended product which user may buy. The price is directly proportional to the time.

C.Capture User Behaviour

Enhanced Ecommerce tracking allows you to measure the number of transactions and the revenue that your website or mobile app generates.
Ecommerce tracking helps you understand user behaviour across the user's entire online shopping experience including: product impressions, product clicks, viewing product details,adding a product to shopping cart, initiating the check-out process, transactions, and refunds.

The behaviour report visualizes the paths users travelled from one screen, page or event to the next. The information is used to discover what content keeps users engaged with web application. By tracking the user behaviour can help to identify potential content or usability issues. If the numbers of users are below expectations, one has to re-evaluate the marketing efforts to see whether the application is targeting the appropriate audiences and also can focus on negative press or social content that might affect the traffic. Even if the entire marketing and social buzz are positive, the organization people may be creating the technical hurdles for users with the website or app design.
By having user behaviour information following questions can be answered:-

- Is there any event that is always triggered first? Does it lead to users to more events or other behaviours?
- Did users go right from product pages to checkout without any additional shopping?

D.Recommendation

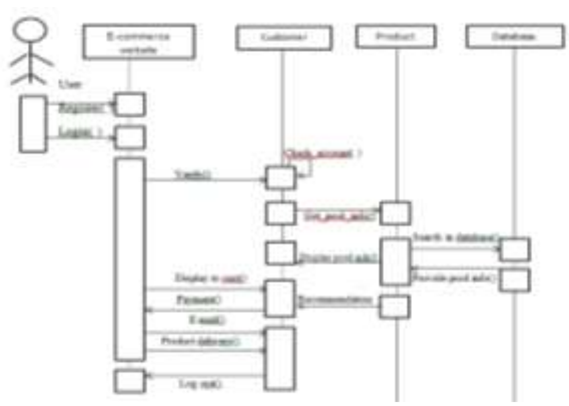
Recommendation system is subclass of information filtering system that seeks to predict the "rating" or "preferences" that a user would give to an item.
Recommender systems have become popular in recent years, and are utilized in a variety of areas including movies, music, books, search queries etc. Recommender systems typically produce a list of recommendations in one of two ways through collaborative filtering. Collaborative filtering is the approaches to build the model from a user's past behaviour as well as similar decisions made by other users. Recommender systems are a useful alternative to search algorithms, since they help users discover items they might not have found by themselves. Interestingly enough, recommender systems are often implemented using search engines are indexing non-traditional data.

Collaborative filtering

One approach to the design of recommender systems that has wide use is collaborative filtering.
Collaborative filtering methods are based on collecting and analysing a large amount of information on user's behaviours, activities or preferences and predicting what users will like based on their similarity to other users. A key advantage of the collaborative filtering approach is that it does not rely on machine analysable content and therefore it is capable of accurately recommending complex items such as movies without requiring an understanding of item itself. Many algorithms have been used in measuring user similarity or item similarity in recommender systems.

Sequence Diagram

The sequence diagrams are one of the interaction diagrams that depict the communication between the objects. The collaboration of objects is modelled based on a time sequence. The objects involved in the scenario pass messages between themselves. Here the return messages are not shown but are to be understood as implicit.Sequence diagrams are sometimes called event diagrams or event scenarios.
In the sequence diagram show below, first the user has to register to the website to access the data completely. Once the customer registers, then he can view the product, buy the product. Based on his views and behaviour, recommendations will be done to the user. The details of the product will be stored and retrieved from the database. Later after conforming the product, payment should be done.



E. Implementation on cloud

Cloud Computing provides us a means by which we can access the applications as utilities, over the internet. It allows us to create, configure, and customize the business applications online. Cloud computing can be explained as a variety of computing that involves sharing computing resources, rather than preparing local server machines or personal devices mend to handle applications. Implementation activities for a cloud solution are similar to that of an outsourced solution. That is, the agency will have to conduct business analysis, build a business case, source a cloud service provider (CSP), plan and implement the solution, possibly with the assistance of a third-party system integrator. The Information provides advice across the lifecycle of a cloud solution project with the aim of ensuring the cloud solution will:

- Meet business needs in terms of both functionality and performance;
- Provide the expected efficiencies and benefits;
- Adequately protect agency information;
- Comply with legislative and regulatory requirements; and
- Integrate with existing processes and systems

The work to develop an application approach to cloud-based services will provide the business context required for candidate cloud-based services. Business analysis activities will be similar to those used for an outsourced solution. Such activities include building a business model and gathering requirements to form the basis for the business case, sourcing, implementation and testing.

The business model will help the agency determine performance and resource requirements, lifecycle cost estimation, and required risk treatment measures. Agencies should consider how they would respond to business continuity and disaster recover scenarios, such as cloud service disruption or cancellation. These scenarios can later be developed into requirements and plans. The business model must have sufficient detail to estimate cost in terms which can be applied to the CSP's cost model. For infrastructure as a service (IaaS) and platform as a service (PaaS), this might be measured in resource usage per period of time, as for processing, throughput and storage. For software as a service (SaaS), service might be measured by number of transactions or number of users.

With an understanding of which resources to measure, business analysts should model expected utilization and potential surge scenarios by considering:

- User characteristics, e.g. user types/roles, number of users, usage scenarios;
- Average usage rates, e.g. transactions per second
- Data characteristics
- How usage rates will vary, e.g. upper and lower ranges;
- Where can changes to usage rates be predicted, either at planned times or based on events;
- How usage will grow or scale over time, perhaps with the number of users; and
- How usage will change for each system actor, e.g. end user, administrator, batch processes.

Where possible, agencies should validate the model either by comparison to existing systems, with a benchmarking program, or by piloting a solution. Cloud-based services may provide better value for short-term or burst use, but a non-cloud solution may provide more value over the long term, particularly for services with steady loads.

V.CONCLUSION & FUTURE WORK

Predictive models that focus solely on accuracy must sacrifice interpretability. Here balance is maintained between them. Co-clustering principles can be used to drive an enterprise recommendation engine for B2B recommendations. This offers easily interpretable recommendations, which are furnished to sales and marketing teams for further evaluation. Various methodologies are used and the respective financial benefits are shown. Further, additional external indicators are used as supporting factors in the recommendation process, such as recent news or tweets that can strengthen the confidence for a particular client-product recommendation.

REFERENCES

[1]. H Chen, R H Chiang, and V C Storey, "Business intelligence and analytics From big data to big impact," MIS Quart , vol 36, no 4, pp 1165–1188, 2012

[2]. A Felfernig, G Friedrich, and L Schmidt-Thieme, "Guest Editors' introduction Recommender systems," IEEE Intell Syst , vol 22, no 3, pp 18–21, May 2007

[3]. J B Schafer, J Konstan, and J Riedl, "Recommender systems in e-commerce," in Proc ACM Conf Electron Commerce, 1999, pp 158–166

[4]. D M Fleder and K Hosanagar, "Recommender systems and their impact on sales diversity," in Proc ACM Conf Electron Commerce, 2007, pp 192–199

[5] L Barrington, R Oda, and G R Lanckriet, "Smarter than genius? Human evaluation of music recommender systems," in Proc ISMIR, 2007, vol 9, pp 357–362

[6] N Tintarev and J Masthoff, "A survey of explanations in recommender systems," in Proc ICDE Workshops, 2007, pp 801–810

[7] R Pan, Y Zhou, B Cao, N N Liu, R M Lukose, M Scholz, and Q Yang, "One-class collaborative filtering," in Proc IEEE Int Conf Data Mining, 2008, pp 502–511

[8] S Madeira and A L Oliveira, "Biclustering algorithms for biological data analysis A survey," IEEE Trans Comput Biol Bioinformat , vol 1, no 1, pp 24–45, Jan 2004

[9] A Tanay, R Sharan, and R Shamir, "Biclustering algorithms A survey," in Handbook of Computational Molecular Biology Boca Raton, FL, USA Chapman & Hall/CRC Press, 2005, ch 26

[10] M Rege, M Dong, and F Fotouhi, "Bipartite isoperimetric graph partitioning for data co-clustering," Data Mining Knowl Discov , vol 16, no 3, pp 276–312, 2008

[11] Y Zhang, G Lai, M Zhang, Y Zhang, Y Liu, and S Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in Proc SIGIR, 2014, pp 83–92

[12] D D Lee and H S Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, vol 401, no 6755, pp 788–791, 1999

[13] J Yang and J Leskovec, "Overlapping community detection at scale A nonnegative matrix factorization approach," in Proc Int Conf Web Search Data Mining, 2013, pp 587–596

[14] C -J Lin, "Projected gradient methods for nonnegative matrix factorization," Neural Comput , vol 19, no 10, pp 2756–2779, 2007

[15] R Heckel and M Vlachos, "Interpretable recommendations via overlapping co-clusters," arXiv 1604 02071

[16] W S Lee and B Liu, "Learning with positive and unlabeled examples using weighted logistic regression," in Proc Int Conf Mach Learn , 2003, pp 448–455

[17] C Elkan and K Noto, "Learning classifiers from only positive and unlabeled data," in Proc KDD, 2008, pp 213–220

[18] S Rendle, "Factorization machines," in Proc Int Conf Data Mining, 2010, pp 995–1000

[19] I Porteous, A U Asuncion, and M Welling, "Bayesian matrix factorization with side information and Dirichlet process mixtures," in Proc AAAI, 2010, pp 563–568

[20] D Agarwal and B -C Chen, "Regression-based latent factor models," in Proc KDD, 2009, pp 19–28

[21] A Jameson, B Berendt, S Gabrielli, F Cena, C Gena, F Vernero, and K Reinecke, "Choice architecture for human-computer interaction," Found Trends Hum -Comput Interact , vol 7, no 1/2, pp 1–235, 2014

[22] F M Harper and J A Konstan, "The movielens datasets History and context," ACM Trans Interact Intell Syst , vol 5, no 4, pp 1–20, 2015

[23] B M Sarwar, G Karypis, J A Konstan, and J Riedl, "Analysis of recommendation algorithms for e-commerce," in Proc Electron Commerce, 2000, pp 158–167