Recommending product based on customer review from e-commerce websites

Dr. Bhrantav Vora, Assistant Professor
Indus Institute of Information and Communication Technology,
Department of Computer Science, Indus University, Ahmedabad.

Dr.Ronak B. Patel, Associate Professor
Shrimad Rajchandra Institute of Management and Computer Application ,
UkaTarsadiaUniversity,Maliba Campus, Bardoli-Mahuva Road,
Gopal Vidyanagar, Barodli,Gujarat- 394350

Dr. Meghna Patel
Assistant Professor,
A. M. Patel Institute of Computer Studies,
GanpatUniversity, Kherava, India.

Abstract

Recommendation assumes an extremely essential part in human life. People depend a ton on suggestions from their day-to-day schedules to taking any significant choice for example buying new things, selecting an asset, purchasing furniture, or another home. Individuals' dependence on recommendation and take their choices in view of the suggestion got from different sources. In our research work, we are zeroing in on giving suggestions to the clients of online business sites. As the title proposes our suggestion framework depends on clients' review. While shopping through an internet business site, assuming the client gets befuddled in choosing an item out of numerous accessible choices then the internet business stage gives a correlation choice in light of the highlights of the item, yet imagine a scenario where the client can peruse the reviews of the item and afterward can conclude that with which item, he/she ought to go for procurement.

Further, it is particularly troublesome or we can express close to incomprehensible for the clients to stroll through a large number of accessible reviews of any single item and afterward compare with other item with conclude that with which item he/she ought to continue for procurement. To suggest an item founded on reviews we want to manage the text and henceforth further frequent pattern mining has been carried out to remove the significant substance lastly man-made consciousness applied to the separated pertinent substance. We have additionally tried our result in an AI calculation named Random Tree to approve our created calculation.

Introduction

The huge advanced in E-Commerce joins with the fame of online relational organizations that is profoundly affecting the worldwide economy. In particular, customer shopping practices have undergone changes, and an innovative kind of internet business, bloomed called "E-Commerce". E-Commerce development has the responsibility of social and economic insinuations.[1] There is a change in the habit of customers in consumptions of products and hence business houses need to acclimate to this changing environment and form their strategies of business and market. These changes in customer and business houses are being made so quickly that it could be termed as a social revolution. [2]

The concern has been shown from authors for the traditional shopping infrastructure for their existence. The proposal framework is an assistance that investigations customer information, including the customer's buy information, so as to prescribe them the most reasonable items or administrations. Most suggestion frameworks have the procedure stream.[3] By breaking down the accumulated customer information, likenesses between customers

are found, lastly things that the customer is probably going to buy are suggested. Recommender frameworks gather data dependent on the inclinations of customers (for instance— melodies, motion pictures, jokes, books, travel goal and e-learning material). Recommender frameworks work dependent on customers' data from various sources and give proposal of things. This data can be unequivocal (customer rating) and verifiable (checking customer's conduct), with a large number of customers utilizing long range informal communication administrations like Face-book, Twitter, etc.[4] The rich information that has aggregated in these person-to-person communication destinations empower an assortment of suggestion frameworks for its customers. Recommender frameworks empower customers to get to items or articles that they would other-wise not know about because of the abundance of data to be found on the Internet.

The two conventional proposal methods are collaborative filtering and content-based filtering. While the two strategies have their focal points, they likewise have certain disservices, some of which can be understood by joining the two methods to improve the nature of the proposal. The subsequent framework is known as a hybrid recommender framework. Recommendation can be divided into the following phases, which include information collection phase, learning phase, prediction or recommendation phase. [4]

Information collection phase

The framework has to know however much as could be expected from the customer so as to give sensible proposal directly from the beginning. The precision of suggestion relies upon the number of reviews gave by the customers for a product.[5]

Learning phase

In learning phase of recommendation system, we apply learning to extract and filter the customers' data collected from the information collection phase.

Prediction/Recommendation phase

In this final phase the recommendation system provides an item in which customer most likely to purchase. It could be done by the information collected from the first phase and then applied algorithm on it into second phase. Figure illustrate the recommendation phases.[6]
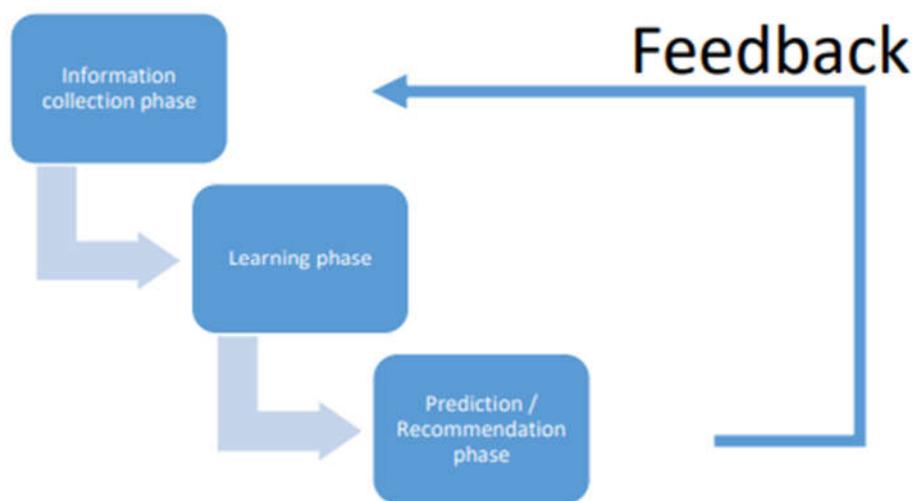
Recommendation Phases

In this research study, a recommendation model was designed, developed and implemented which analyses using patterns which is extracted from the reviews of the individual customer and rating given from the individual customer. The recommendation algorithm is used in the proposed system to recommend the product in different level to the customer. In this research our proposed model accepts users' review of more than one product and gives level of recommendation i.e., most likely to purchase, likely to purchase, recommended, less recommended and not recommended at all.[7]

Literature Review

The principle thought behind the Recommendation frameworks for E-Commerce is to fabricate a connection between the products, customers, and settle on the choice to choose the most suitable item to-a particular customer. Gone ahead time, consistently in a hurry, current customers shop in erupts from numerous gadgets, as opposed to set aside the effort to take part in a careful shopping long-distance race. [8]

Brands conveying fast, on-point offers to catch the most advantages. As indicated by the Personalization Consumer Survey did by 48% of customers go through additional with an E-Commerce organization conveying customized shopping knowledge. ML frameworks enable you to catch information from past and current shopping sessions and change them into dynamic offers. ML can pinpoint which product is in stock to feature. Keen recommender frameworks can look over your whole item index and line up the best items for individual purchasers. You can make it a stride further and show the closest blocks.[9]

E-Commerce defines the scope of products is accessible, to peruse the information, focusing on the prospect with nearby stock promotions later in the day. A few out of every unique possibility is prepared to purchase from you at this very moment. When a particular product is absent on an E-Commerce site, many customers will go looking somewhere else. Indeed, even graphic inquiries like 'a white shirt with brilliant catches' or 'nutrients for winter' may not lead customers to the items they need. On location, web search tools are somewhat inadequate when given random questions.[10]

Proposed Framework

Proposed framework can be divided into four phases, phase 1 is identifying products in the e-commerce website, phase 2 is extracting user reviews and rating of identified products in phase 1, phase 3 is implementing machine learning algorithm on collected information from phase 2 and finally in phase 4 by applying improved steepest ascent hill climbing algorithm recommending the product.
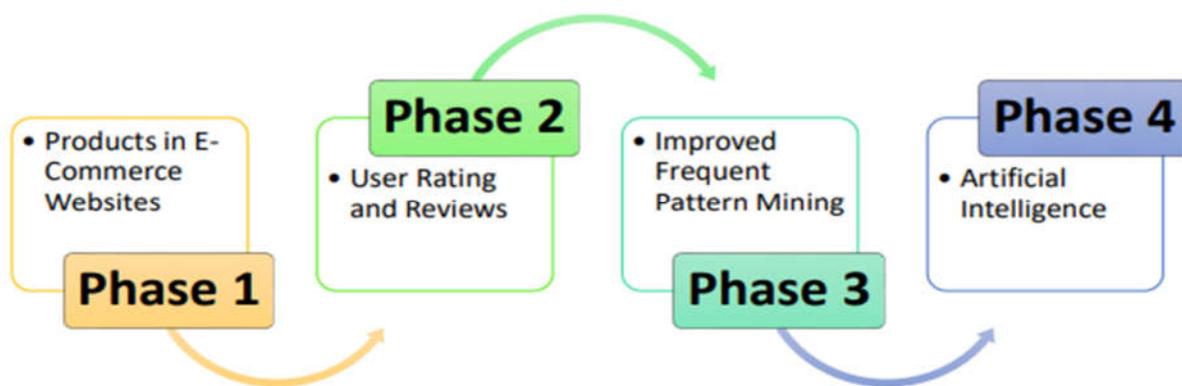
Proposed framework

Products in E-commerce

This is the very first phase in product recommendation framework, in this phase we have identified six products (i.e., Smart Phone, Television, Refrigerator, Laptop, Washing Machine and Air Conditioner) in electronic segment

in e-commerce website to collect data. We wanted to collect ample amount of data in terms of rating and reviews of customer to provide enough amount of data to our proposed model, further we have observed that products in electronic segment are having a greater number of reviews and rating compare to others and hence we have identified six different product categories namely fridge, television, laptop, smart phones, air condition and washing machine.

User Rating and Reviews

In second phase of recommendation framework, we have developed a python script to scrap product detail from the e-commerce websites. In attempt to extract required data of the product we scraped the following product detail using python script, Brand, Product, Model, Users' Review, Users' Rating, Price. We have manually selected features like users' rating and reviews and took it forward in next phase of recommendation framework.

Improved Frequent Pattern Mining

In this phase, we apply NLP on collected reviews in previous phase to process the natural language, in order to find appropriate information from the collected reviews we have applied Bag of Words model and try to find frequent pattern in reviews and converted categorical information into numerical to process it further in machine learning algorithm.

Artificial Intelligence

In this phase of recommendation framework, Artificial intelligence is used to recommend the product to the customer based on the received information from the previous phase. Steepest ascent hill climbing algorithm is used to find the best neighboring node to meet the goal state and that product is being recommended, further in improvement we have evolved steepest ascent hill climbing algorithm to the next level and provide level of recommendation i.e. Most recommended, most likely recommended, Recommended, less recommended and not recommended at all.

Implementation

Implementation of the recommender was done in python using Google co-lab cloud platform, we have developed a python script to collect the data from e-commerce applications and then data cleaning was happened on it. Detail stapes of implementation were discussed below.

Data Collection

Information assortment is one of the most essential and important errands to be finished, to push ahead in research work, a scientist needs to gather information, and information can be accessible from the different sources. In our exploration work we have gathered information from the online business sites

Web Scraping

Web scratching is the strategy to separate enormous measure of organized information from the web. In the period of data and information on the web, it is profoundly equipped to separate information from the web. To execute information mining, design acknowledgment, estimating through data set penetrating one should require more than adequate measure of information to deal with and it ought to likewise accessible in delicate structure to handily execute unequivocally precise and result situated AI and man-made reasoning calculation and for a similar web is the most reasonable stage accessible. Separating information from the web through various sites utilizing significantly strong prearranging language for example python is web scratching.

For web scraping using python script we used Selenium Framework, Chromium Driver, Pandas Library and we have also implemented programming with Xpath to Scrap Data. Furthermore, we have exported the extracted data to collect the primary data as shown in below figure.

No	Category	Total Reviews	Total Model	Product
1	AC	2636	6	18
2	Laptop	4570	2	6
3	MobilePhone	31880	9	27
4	Refrigerator	11090	7	21
5	TV	13930	7	21
6	WashingMachine	10920	7	21
Total	6	75026	38	114

Data Collected/ Scrapped

Data Refining

Information pre-handling is vital, any progression pushing ahead without clear understanding and doing legitimate pre-handling might prompt monstrous mistakes or we may put ourselves into a pound. Subsequently, it is exceptionally expected to perform information pre-handling which incorporates information cleaning and afterward executing a bag of word model to view as key words for further use.

During data refining we have selected feature manually and worked on user reviews for and specific product, in case of features are more in number then we can go for some Feature Selection algorithms like PCA, SVD, LDA. After features selection we have clean the data by removing unwanted information, null record and duplicate records.

After data pre-processing task done, we went for Text Transformation and NLP in which we have used Bag of Words Model and Improved Frequent Pattern Mining for Selecting Key Word and Finding Rank of Key Words to work with text and recommend the product based on user reviews. Once we have transformed the data from categorical to numerical then we can start implement recommendation algorithm on it.

Recommendation Algorithm

For the final advice, we have prepared a product final matrix from the getting to know or pre-processing section and given it as an input to the steepest ascent hill climbing set of rules to expect the final product advice.[11] In our recommendation algorithm we have imported necessary library files to read data from the CSV Files and giving rank to the word after preparing product final matrix Steepest-Ascent Hill Climbing Algorithm was used to recommend the product.[12]

Result

In above mentioned experimental procedure we have identified six different products in electronic segment and we have divided products under the category of different brands, hence under each brand we have collected data for at least three different model and work on that. In final attempt we collected data i.e., customers' review and rating of different 114 products which can be divided into 38 different models and total reviews and ratings collected was 75026. After doing pre-processing on gathered data and further passed it to the recommendation or prediction phase we received result to recommend the product at different level which further passed as input to Improved Steepest Ascent Hill Climbing algorithm to find the proof of confidence.

Conclusion

These studies instigated the product advice gadget for the e-commerce internet site, which recommends the maximum in all likelihood to buy product or product at special tiers of recommendation to the purchaser using Steepest-Ascent Hill mountaineering and Improved Steepest-Ascent Hill climbing set of rules respectively and content material-based filtering, based at the clients' rating and overview. For the clients' convenience, we advanced scripts and scraped records directly from the e-commerce web sites. NLP is used to method customers' critiques and extracted applicable records from the reviews, in addition based totally on the rating and reviews together we have evolved two algorithms namely Steepest-Ascent Hill climbing and Improved Steepest-Ascent Hill mountaineering. Our algorithm has obstacles of the bloodless superstar hassle inside the advice device. The newly released product or the product which has no critiques and ratings given from the consumer isn't protected in this research. For the

destiny scope, we could awareness at the cold big-name hassle in the advice device with our proposed Improved Steepest-Ascent Hill hiking algorithm.

References

[1]   S. Escursell, P. Llorach-Massana, and M. B. Roncero, "Sustainability in e-commerce packaging: A review," *J Clean Prod*, vol. 280, p. 124314, 2021.

[2]   Y. Zhang, H. Abbas, and Y. Sun, "Smart e-commerce integration with recommender systems," *Electronic Markets*, vol. 29, no. 2, pp. 219–220, 2019, doi: 10.1007/s12525-019-00346-x.

[3]   P. Melville, R. J. Mooney, and R. Nagarajan, "Content-boosted collaborative filtering for improved recommendations," *Proceedings of the National Conference on Artificial Intelligence*, no. July, pp. 187–192, 2002.

[4]   S. Dhawan, K. Singh, and N. Kumar, "Collaborative Filtering Based Product Recommendation System for Online Social Networks," vol. 2, no. 5, pp. 126–131, 2017.

[5]   T. B. Lalitha and P. S. Sreeja, "An Analysis on E-Learning and Its Recommendations," in *New Opportunities for Sentiment Analysis and Information Processing*, IGI Global, 2021, pp. 166–187.

[6]   M. Daoud and S. K. Naqvi, "Recommendation System Techniques in E- Commerce System," vol. 4, no. 1, pp. 569–573, 2015.

[7]   Y. Koren, S. Rendle, and R. Bell, "Advances in collaborative filtering," *Recommender systems handbook*, pp. 91–142, 2022.

[8]   F. Aznoli and N. J. Navimipour, "Cloud services recommendation: Reviewing the recent advances and suggesting the future research directions," *Journal of Network and Computer Applications*, vol. 77, pp. 73–86, 2017, doi: https://doi.org/10.1016/j.jnca.2016.10.009.

[9]   P. Bertens, A. Guitart, P. P. Chen, and A. Perianez, "A Machine-Learning Item Recommendation System for Video Games," *IEEE Conference on Computatonal Intelligence and Games, CIG*, vol. 2018-Augus, 2018, doi: 10.1109/CIG.2018.8490456.

[10]   M. Patil and M. Rao, "Studying the Contribution of Machine Learning and Artificial Intelligence in the Interface Design of E-commerce Site BT - Smart Intelligent Computing and Applications," 2019, pp. 197–206.

[11]   D. Tsigkari and T. Spyropoulos, "An approximation algorithm for joint caching and recommendations in cache networks," *IEEE Transactions on Network and Service Management*, 2022.

[12]   Y. Guo, M. Wang, and X. Li, "Application of an improved Apriori algorithm in a mobile e-commerce recommendation system," *Industrial Management and Data Systems*, vol. 117, no. 2, pp. 287–303, 2017, doi: 10.1108/IMDS-03-2016-0094.