**INTRODUCTION**

E-commerce is exploding everywhere around the world due to the increased demand for online shopping and social media platforms have become an essential part of our daily lives. The COVID-19 global pandemic has exacerbated this demand for online shopping as the government has imposed recurring national lockdowns to reduce the rate of community transmission. As a result, many companies and small businesses that lack the online infrastructure to carry out their day-to-day activities have suffered, and in extreme cases, closed due to their inability to respond effectively to this increased demand. Despite these challenges, the increased online presence of shoppers provides a significant opportunity for companies to utilize the huge inflow of data to anticipate the needs and improve the quality of service for customers. Therefore, efficient Recommender Systems are crucial for these businesses to have as the abundance of information on the internet makes it more difficult and time consuming for users to find and select what they want when they need it. As a result, the need for a personalized support system, recommender system, in filtering through sizable amounts of information on the web to get what suits our preferences is highly required as users become overwhelmed by the inflow of new information that is generated every day. How can businesses leverage a recommender system to sustain their day-to-day business activities and create a custom shopping experience for each customer?

A recommender system is a subclass of information filtering system that seeks to predict the “rating” or “preference” a user would give to an item. They improve the quality of search results by providing items that are related to the searched item or search history of a user. Researchers have distinguished four different classes of recommender systems:

* *Content-based:* The system generates recommendations from two sources: the features associated with products and the ratings that a user has given them. Content-based recommenders treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features. [7, 10, 12]
* *Collaborative:* The system generates recommendations using only information about rating profiles for different users. Collaborative systems locate peer users with a rating history similar to the current user and generate recommendations using this neighborhood. Examples include [8, 9]
* *Demographic:* A demographic recommender provides recommendations based on a demographic profile of the user. Recommended products can be produced for different demographic niches, by combining the ratings of users in those niches [10, 14]
* *Knowledge-based:* A knowledge-based recommender suggests products based on inferences about a user’s needs and preferences. This knowledge will sometimes contain explicit functional knowledge about how certain product features meet user needs. [3, 4]

**LITERATURE REVIEW**

Many researchers have tried to provide solutions to the problem of personalized search through various methods. Below are some of the solutions:

Linyuang Lu et al [13] reviewed the recent developments in recommender systems and discussed the major challenges which pose danger for the use and performance of algorithms in recommender systems. These challenges include data sparsity, scalability, cold start (insufficient information to produce recommendation for new users when they enter the system), vulnerability to attacks and the value of time as most recommendation systems neglect the time stamp of evaluations. They suggested that there is no unique best recommendation method. Rather, depending on the context and density of the available data, different methods adapting to particular applications are most likely to succeed

Robin Burke [5] surveyed the space of two-part hybrid recommendation systems, comparing four different recommendation techniques and seven 7 different hybridization strategies. He also examined and compared the implementations of 41 hybrids. The study found out that cascade and augmented hybrids worked well, especially when two components of different strengths were combined. Robin proposed four techniques as the basis for recommender systems: collaborative, content-based, knowledge-based, and demographic techniques. These techniques have well known flaws such as the cold start problem for collaborative and content-based systems where there is insufficient information to produce recommendations for new users and the limited capacity for knowledge engineering in knowledge-based systems. A hybrid system combines multiple techniques together to achieve some synergy between them. For example, a collaborative system and a knowledge-based system might be combined to so the knowledge-based system can compensate for the cold start problem.

J. Ben Schafer et al [16] introduced the core concepts of collaborative filtering, its primary uses for users of the adaptive web, the theory and practice collaborative filtering algorithms and design decisions regarding rating systems and acquisition of ratings.

Amel Ziani et al [17] proposed a multilingual recommender system based on sentiment analysis to help users decide on products, restaurant, movies, and other services using online product reviews. It combined recommendation system and sentiment analysis to generate the most accurate recommendations for users by detecting the polarity score of opinions using a semi-supervised support vector machine (SVM). The results analysis evaluation provided interesting findings on the impact of integrating sentiment analysis into a recommendation technique based on collaborative filtering.

Meghana Ashok et al [1] proposed a social framework which extracted user’s reviews, comments of restaurants and points of interest such as event and locations, to personalize and rank suggestions based on user preferences. Machine Learning and Sentiment Analysis based techniques such as Rule-based and Aspect based sentiment analysis were used to optimize search query results. This provided users with quicker and more relevant data.

Hao Ma et al [13] provided a general method for improving recommender systems by incorporating social network information to provide more personalized recommendations as traditional recommender systems ignore social relations among users. They proposed a matrix factorization framework with social regularization which involves modelling social network information as regularization terms to constrain the matrix factorization framework.

Xue Bai [6] proposed a heuristic model, the Markov Blanket model, to predict the sentiments of online users from text papers. This model was search enhanced that it could perceive the word dependencies and dispense a vocabulary that is enough for sentiment extraction. Their performance indicated the ability to identify close-fitted set of predictive features at the same time yield comparable or better predictive results about the orientations of sentiments than many state-of-the-art feature selection algorithms and sentiment prediction methods.

Hyung Jun Ahn [2] proposed a new approach to automated recommendation; popularity-based recommendation (PBR), that utilizes the popularity characteristics of products. Popularity features were considered in building recommender systems because the popularity of a product greatly influences consumer purchasing decisions and often represents important characteristics of a product. Popularity features play an important role in consumers’ purchasing decisions because most consumers are influenced by how others feel about a product or how widely a product has been exposed in the market.

The application of the popularity concept in building recommender systems requires an understanding of its three key dimensions. The first dimension represents whether consumers perceive the product to be of high value. When used alone, this dimension represents a product’s quality or level of satisfaction rather than its popularity, but when combined with the other two dimensions, it plays an important role in explaining the popularity features of a product. The second dimension represents the frequency of a product’s being purchased regardless of its perceived value. It differs from the first dimension in that many products are mass-marketed and widely exposed but are not rated highly by consumers. The third dimension of product popularity is the size of the strong-support group for a product regardless of its average rating or frequency of purchase.

When compared with the widely used collaborative filtering method, the PBR system showed significant improvement under data sparsity and cold-starting conditions. This outcome demonstrates that PBR would be of great practical value for recommender systems.

**DATASET**

Dataset: Amazon Product Reviews – Electronics\_products user ratings.

**Web link:** <https://nijianmo.github.io/amazon/index.html>

**Citation**

**Justifying recommendations using distantly-labeled reviews and fined-grained aspects**

Jianmo Ni, Jiacheng Li, Julian McAuley

*Empirical Methods in Natural Language Processing (EMNLP), 2019*

**Range index:** 7,824,481 entries, from 0 to 7,824,480

**Data Colums**: 4

**Data types:** Objects (column 1 & 2), float64 (Column 3) and int64 (column 4)

**Attributes: 4 -** Only the **timestamp** attribute will not be used in this analysis.

**userId** : All users have a unique id.

**productId** : Each product has a unique id.

**Rating** : Rating of each product provided by the corresponding user.

**timestamp** : Time at which the rating was entered.

**Descriptive Statistics: Rating**

count 7.824481e+06

mean 4.012337e+00

std 1.380910e+00

min 1.000000e+00

25% 3.000000e+00

50% 5.000000e+00

75% 5.000000e+00

max 5.000000e+00

**APPROACH**

**STEP 1: Import libraries and Load Dataset**

* Load the dataset and add headers
* Display the Data
* Check data shape
* Check data types
* Take a subset of the data

**STEP 2: Analyze/Clean data**

* Get data summary
* Find min and max ratings
* Check for missing values
* Check the ratings distribution against the total number of ratings
* Check the total number of Unique users and products
* Drop the Timestamp column as it won’t be used in the recommender system model
* Analyze the user ratings to get the number of products rated by each unique user

**STEP 3: Build recommender systems**

* Popularity based recommendation

This system works with trend.

* Collaborative filtering – Item to item recommendation, using **KNNWithMeans** algorithm.

Using historical item rating of people who similar taste to predict how someone would rate an item.

* Model based collaborative filtering system using the **sklearn**

Training models to be able to make predictions

**References:**

1. Ashok, M., Rajanna, S., Joshi, P., Kamath, S. Personalized Recommender System using Machine Learning based Sentiment Analysis over Social Data. *2016 IEEE Students’ Conference on Electrical, Electronics and Computer Science*
2. Ahn, h. *Utilizing Popularity Characteristics for Product Recommendation, 2006.*
3. Burke, R.: The Wasabi Personal Shopper: A Case-Based Recommender System. In: Proc. of the 11th Nat. Conf. on Innovative Appl. of AI, 844-849. (1999)
4. Burke, R.: Knowledge-based Recommender Systems. In: A. Kent (ed.): Encyclopedia of Library and Information Systems, 69, Sup. 32. (2000)
5. Burke R. Hybrid Web Recommender Systems. *School of Computer Science, Telecommunications, and Information Systems DePaul University, 243 S. Wabash Ave. Chicago, Illinois, USA.*
6. Bai, X. Predicting consumer sentiments from online text. Decision Support Systems, 50(4):732–742, 2011.
7. Chen, L. and Sycara, K.: WebMate: A personal agent for browsing and searching. In Proc. of the 2nd Intl Conf. on Autonomous Agents (Agents'98), 132-139, New York. ACM Press. (1998)
8. Goldberg, D., Nichols, D., Oki, B, & Terry, D. Using collaborative filtering to weave an information tapestry. 35(12):61–70. CACM. (1992)
9. Hill, W., Stead, L., Rosenstein, M. and Furnas, G.: Recommending and evaluating choices in a virtual community of use. In: CHI ’95: Conf. Proc. on Human Factors in Computing Sys., Denver, CO, 194-201. (1995)
10. Jennings, A. and Higuchi, H.: A User Model Neural Network for a Personal News Service. UMUAI 3, 1-25. (1993)
11. Krulwich, B.: Lifestyle Finder: Intelligent User Profiling Using Large-Scale Demographic Data. AI Magazine 18 (2), 37-45. (1997)
12. Lang, K.: Newsweeder: Learning to filter news. In: Proc. of the 12th Intl Conf. on Machine Learning, Lake Tahoe, CA, 331-339. (1995)
13. Lu, L., Medo, M., Zhang, C., Zhang, Z., Zhou. Recommendation Systems. *Beijing Computational Science Research Center, Beijing, 100084, PR China, 2012*
14. Ma, H., Zhou, D., Liu, C., Lyu, M., King, I. Recommender Systems with Social Regularization. *Chinese University of Honkong, 2011*.
15. Pazzani, M. J.: A Framework for Collaborative, Content-Based and Demographic Filtering. AI Review 13 (5/6), 393-408. (1999)
16. Schafer, J., Frankowski, D., Herlocker, J., and Sen, S. Collaborative Filtering Recommendation Systems*. School of Electrical Engineering and Computer Science Oregon State University*
17. Ziani, A., Azizi, N., Schwab, D., Aldwairi, M., Chekkai, N., Zenakhra, D., Cheriguene, S. Recommender System Through Sentiment Analysis. *2nd International Conference on Automatic Control, Telecommunications and Signals, 2017.*