Title: Machine Learning Approach for Recommendation System Based on Online Review

Author(s): Muhammad Awais¹, Muhammad Nadeem², Muhammad Babar³, Muhammad Ali⁴, Muhammad Arslan⁵, Rida²

Affiliation(s): Three-Dimensional Lifestyle Ltd
                Sabasoft
                Heritage College
                Confiz
                We360

DOI: https://doi.org/10.32350.umt-air.21.04

History: Received: June 29, 2022, Revised: July 24, 2022, Accepted: August 29, 2022


Copyright: © The Authors

Licensing: This article is open access and is distributed under the terms of Creative Commons Attribution 4.0 International License

Conflict of Interest: Author(s) declared no conflict of interest
Analyzing the Online Review Recommendation System through Machine Learning Approach

Muhammad Awais¹,*, Muhammad Nadeem², Muhammad Babar³, Muhammad Ali⁴, Muhammad Arslan⁵, Rida²

¹ Three Dimensional Lifestyle Ltd
² Sabasoft
³ Heritage College
⁴ Confiz
⁵ We360

Abstract—Various recommendation systems have been introduced over the last few years to provide the necessary recommendations and suggestions based on online user reviews. These recommenders provide valuable information through user-generated reviews with the help of various processes. These recommendation systems follow different techniques to extract the important information from the given reviews. Furthermore, this retrieved information proves to be as valuable enough to put forward suggestions for users’ interests. In the current research, such a portable recommender system has been developed based on the three ‘machine learning approaches. Eventually, these approaches recommend its users different mobile phones available in the market and help them in selecting the best one based on the online reviews of the users.

Index Terms—Machine learning, online user review, recommendation system.

I. Introduction

Over the last few years, digital technologies have grown exponentially, and their use has become globalized. However, digital devices have become a significant part of everyone's life in modern ages. Moreover, market of smart gadgets has radically increased and it also ensures the users benefit. These digital devices are embedded with several features which are effortlessly accessible from centralized markets like Google Play [1]. Furthermore, the mammoth increase in digital information is present all over the internet. The user traffic that visits several web pages to answer the

* Corresponding author: digital@3dlifestyle.com.pk
user’s queries has generated a potential overload challenge for the data and information. Eventually, it loses the instant availability of authentic information for interested users over the web [2].

Recommender systems (RS) are the information filter systems resolving the issues of overloaded information by extracting or retrieving key data. In addition to it, the information fragments are extracted from a huge amount of information, dynamically generated as per the user's interest, preferences, or even their observed behavior regarding any item [3]. RS are useful from the user's perspective and a service provider's perspective [4]. These systems help to minimize the transaction costs of searching and opting for various items within an online shopping environment [5]. In scientific libraries, RS enhances and improves the quality and process of decision-making [5], and it also enables the users to go beyond the catalog searches [6].

Recently, several recommendation frameworks are designed to suggest users in accordance to their choice, but these designs do not completely fulfill the user's needs and desires. Besides, assembling the top-notch recommender framework for the mobile system is also important. Furthermore, the developers face some difficulties and complexities in designing RS which requires legitimate consideration. These suggestions or recommendations improve the decision-making phenomenon of the users. For instance, to make decisions regarding listening to some specific kind of music, buying an item or product, downloading or installing a particular app, or even traveling to some historical place. Also, it helps the users to make decisions of their own choice more quickly and according to their taste [7].

Another stimulus for taking help through RS is the coincidence [8]. All these predictions are computed by using some predictive models, having similar features and characteristics [9]. For instance, the user's rating of the previously bought products or items can be used for a particular scenario. RS is classified into three significant groups to generate the recommendation list, which is based on some specific prediction algorithms. These categories are as follows:

Keeping in view, this proposed novel and portable RS would be applicable to deploy an indoor environment that overcomes the shortcomings of the existing
recommendation systems. Whereas, these portable (mobile) recommender systems would have the potential to verifiably catch users' inclinations by evaluating their requirements. While making suggestions, RS processes the information without requiring users' express contributions and analyzing the logical data. Moreover, the current research paper would explore and signify the considerable complexities and evaluation openings in the growing number of recommender frameworks and systems.

The paper's sections are organized as follows: Section 2, describes the Literature Review by discussing the previous approaches that have been used in recommendation systems. Section 3, describes the Methodology to design the proposed mobile recommendation system for online reviews. Section 4, explains the results acquired by implementing the designed recommender system. Section 5, gives the conclusion and recommendation/future gap regarding the mobile recommender system.

II. Literature Review

Recommendation systems are based on 2D setting that provides recommendations to the users of both ends. However, while developing the mobile recommendation system, a set of components can be used or exploited to facilitate and recommend RS and its processes more effectively. A considerable number of researches exist regarding the recommendation systems which are comprising of both mobile and desktop-based recommendation systems. Besides, several other explanations of recommendation systems are also provide in the current research [10, 11].

A portable recommender system by integrated RS system in the mobile application is very essential. For this purpose, a hybrid RS has been used to deal with the additional intricacy of the context and recommend suitable mobile apps to the users [12]. This methodology helps the product or system to have a positive rating. A framework has been proposed known as "SimApp" that can detect and identify similar apps by using the ‘internet-based kernel learning techniques’ [13, 14]. Whereas, researchers collected real data from internet platforms like the Google app store and dug out heterogeneous data sets comprising of multiple modes or approaches. [15], A hybrid RS, comprised of five various filtering algorithms to assist the users in opting for the best mobile app and to download from
further platforms such as Google play store or iCloud. Furthermore, this hybrid RS could assist the user in figuring out and solving issues that were impossible to solve by collaborative RS, as collaborative RS is used to lessen the quality of the spawned predictions.

CoMeR has been presented to recommend the various media types to its users with the help of the context-aware method [16]. A mobile-based photo recommendation system has been proposed that uses contextual data which is further collected from the photos [17]. In [18], car music, RS is developed that recommends the vehicle passengers with the best music which is based on their taste by using the context-aware RS. Furthermore, in [19], the author has presented a music RS that recommends the user by depending upon the individual's daily activities. This mobile-based RS uses a probabilistic approach to recommend and propose the selected songs to the device user.

In [20], RS has been proposed to provide individuals the with selected news based on the existing context and format of the suggested news. In literature, mobile RS has also been exploited to suggest some of the best movies to mobile device users through knowledge-based and contextual data-based recommenders. Furthermore, [21] has proposed a mobile-based M-commerce recommendation system by using the collaborative context and filtering phenomenon. In [22], the author has presented a recommender system known as "Motivate" that gives suggestions and options for individuals' personalized activities to help them maintain a healthy lifestyle. Moreover, in mobile RS, various data sources can be exploited to deliver personalized suggestions or recommendations to the individuals. These comprised methods of using mobile logs as presented by [23] and also use information from different mobile channels. It has been noted that various recommender approaches and data sources can potentially be used besides contextual data to assist device users and individuals. Moreover, another mobile RS has been presented, known as "Smart Museum," which gives its users recommendations regarding heritage-based objects in outdoor and indoor scenarios [24]. Besides, various kinds of contextual information and data are used as time and location to acquire this recommendation. Following this, another RS has been presented by [25] called "iTravel" based on context information and collaborative filtering. This RS is a
peer-to-peer recommender to suggest travel places. Furthermore, this another mobile-based RS named "Tourist" has been proposed by [26] that recommends its users about the leisure or cultures of a particular place where they use this system.

Furthermore, Mobile RS can deliver some pushed recommendations without any open request from its user. These recommendations are about any services of items or objects are given to the users when the system finds anything appropriate for users. This kind of RS has been presented by [27]. In addition to it, [28] the recommendation system has proposed a context-aware collaborative filtering approach.

Another significant mobile RS is the recommender for providing data history of banking [29] that suggests users about those places which they have visited already and cleared their credit card dues. Using the same collaborative filtering approach, a mobile RS was proposed in [30] that gives recommendations for using a special set of rules and decision trees. However, this approach highlights a fundamental problem by utilizing a ‘sentiment analysis (SA) technique’ or using SA in conjunction with the collaborative approach. Hence, this study did not consider that how similar the two films were to one another.

III. Methodology

This section discusses about the detailed work of proposed methodology in terms of a mobile-based RS, further relying on online client reviews. Various steps were considered to design the mobile recommender system, such as processing, data collection, selection, and information analysis. Furthermore, several data volumes were also required from different mobile phones to collect the text information from a large pool of datasets. So, the first data was collected by using the web scrapping method. For this purpose, information and data were collected from about 4000 cell phones. Each cell phone comprised of more than 20 characteristic features that included mobile name, company name, release date, price, mobile weight, and mobile screen size. Besides, RAM, memory, OS, CPU, GPS, resolution, GPU, USB, dimensions, bluetooth, technology, battery, mobile color, sim, sensor, Wi-Fi, and many more were also included among features.

Data collected from user reviews were based on RS, it was cleaned since the collected data had some invalid characteristics. These
characteristics included incorrect records, incorrect formatting, inaccurate values, emojis, and low values. In addition, the collected data had some assorted values in the form of different currency symbols, such as Indian rupees ₹, Pakistan rupees pkr, euro €, and dollar $. Therefore, the currency symbols were further changed into dollar symbols to ease the review processing.

Other than that, segment removal information was incorrectly formatted. So, it was formatted in unique segments like PPI and angle protection. Further, all the stop words were eliminated by utilizing an ordinary articulation. Moreover, the stop words were considered those words which did not assist in figuring out the unique situation or the genuine importance of a sentence. All the <b> labels were superseded/replaced with invalid " values. Similarly, unique character and oblique punctuation lines were replaced with invalid qualities.

Feature scaling is the preprocessing method that involves scaling, which is used to standardize the scope of free factor, also known as standardization. In this procedure, component scaling was used in a dataset because it contained unprecedented scope of qualities. The screen size, camera, RAM, and thickness segments were all ranged between 1 and 120. The segment's price (50 to 4131), Battery (800 to 13000), and PPI (100 to 847) had an exceptionally huge qualities range. Since, these lower and higher qualities influenced the model exactness. Therefore, component scaling was used to standardize the dataset. ‘Sklearn preprocessing module minimax scaler’ was used to highlight the scaling, which could convert all segments to nothing and one territory.

IV. Experiment Results and Discussion

The ML algorithm training and testing were implemented on the dataset to analyze the outcome with experiments. API was also discussed and deployed within this model. The current research comprised three stages; building dataset, machine training, testing and user interface. The first dataset was collected from GSMarena.com, containing 65% of the mobile data. The second dataset targeted a website 91mobile.com, which collected about 22% of the mobile data. Similarly, bestmobile.com was the third website to target, and in this case 8% of the mobile data was collected. The last website targeted was whatmobile.com. From
whatmobile.com, 5% of mobile data was collected.

Initially, the mobile dataset consisted of 6100 mobile records. In the machine training and testing stage, six unique ML algorithms were deployed: Gaussian Naïve Bayes, Decision Tree Classifier, SVM Classifier, Ridge classifier, and K-Neighbors Classifier. The training sets and their calculation were implemented on the portable dataset. Itemized examination of each of the training algorithms was done correctly.

1) To implement this algorithm, the portable information and data set size was 5000, whereas the testing size was 1100. The time taken to train by the Bayes classifier was 17 seconds, and 75% precision could be achieved with the least exactness. Before going to the characterization report and disarray network, a few terms were observed; true positive, false positive, false negative, and true negative.

2) The decision tree classifier was implemented on an enormous dataset. The time to train by the decision tree classifier was 13.10 sec with a precision accuracy of about 76%.

3) The ridge classifier changes over the mark information into [-1, 1] and takes care of the issue with the relapse strategy because of the ridge relapse technique. When this classifier was implemented on a large dataset, the time taken for training was calculated around 24.05 sec.

4) Similar to the previous classifiers and implemented training dataset, K-NN was also applied to train and test the dataset. The time taken by this algorithm classifier was calculated around 60.02 sec, with precision accuracy of about 70%.

5) This experiment demonstrated the efficacy of each ML classifier-based prediction model for mobile phone recommendation by user review settings through experimental analysis. Figure 12, presents the experimental findings of Naive Bayes (NB), Decision Tree (DT), Ridge Classifier, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) classifier-based mobile user review prediction models, along with their respective classification models.
Figure 1. Experimental analysis of each ML Algorithm

V.Conclusion

Recommendation systems have become renowned in the world of Internet. This system delivers the accurate answers to their clients and reduce their intricacy of choice. Moreover, many scientists and experts face basic challenges by planning and assessing such frameworks. The main purpose of this study is to get client’s inclination at the end of the day. Since, cell phones have become main focus of the society in daily existence. Although, many researchers can simply recommend the mobile phones according to their prices. Still, we have collected the user opinion and reviews to design such a system that recommends a mobile phone with high accuracy.

References


recommendations of mobile applications," in *IEEE 23rd Int. Conf. Data Eng. Work.*, 2007, pp. 871-877, doi: [https://doi.org/10.1109/ICDEW.2007.4401078](https://doi.org/10.1109/ICDEW.2007.4401078)


