MACHINE LEARNING BASED ANDROID APP RECOMMENDATION SYSTEM

Dr N. Penchalaiah[1], B. Shanthi[2], A. Sravani[3], P. Venkateswarulu[4], G. Vemalatha[5], P. Sharan Kumar Naidu[6] ,

Department of Computer Science and Engineering, Annamacharya Institute Of Technology And Sciences(Autonomous), Rajampet, Andhra Pradesh, India – 516126.

Email: bommepallishanthi2001@gmail.com

ABSTRACT:

Due to the expansion of smartphones and the App stores, the number of mobile applications is exponentially growing. Users can download a variety of Apps that offer useful services for practically every part of modern life, including socialising, listening to music, watching videos, and browsing the web, to name a few. The current Google and Google Play store recommendation system is said to make suggestions for apps that are similar to the target application while also taking into account the popularity of each app.

However, it does not account for the security features of each programme or the user's preferences. End users can access a wide variety of mobile applications (or apps) through app stores. These apps typically produce network traffic, which uses up users' mobile data plans and could potentially pose a security risk. Due to the lack of a standardised measuring methodology, it is currently difficult to understand how much and what kind of network traffic a mobile app produces in the real world.

In this paper, we quantify and examine the network traffic costs associated with Android apps available in the official Android stores. Our analysis of the data reveals that the traffic costs for apps in various categories vary. Regarding the cost of network traffic, there is a notable variation among the apps with comparable functionality. Then, in contrast to traditional app recommendation methods, we incorporate measurements of traffic cost into our algorithm for app recommendation. According to experimental findings, the recommended recommendation algorithm can successfully guide mobile app users away from a number of potential security and privacy problems brought on by the unneeded network traffic consumption.

Keywords: Content Based Filtering, Recommendation Systems, Unsupervised Machine Learning, Android Applications.

INTRODUCTION:

Nowadays, the industry of app development is flourishing. You may think of it, and there is an Android app for everything. This includes entertainment, education, learning, navigation, monitoring, and even health and fitness. It has undoubtedly gotten difficult for users to choose apps they want to use and instal on their smartphones because there is such a large variety of Android apps available and they are being added to the list at such a rapid rate. [1]

The utilization of user data and intent to generate pertinent recommendations for users is the core of the newly growing field of research known as recommendation systems. using the identical strategy. The goal is to provide suggestions that are in line with the tastes and preferences of the consumers. Content-Based Filtering and Collaborative Filtering are two popular strategies in this area. [2] The difference between collaborative filtering and content-based filtering is that collaborative filtering filters and suggests content based on user activity and content description as opposed to using the user's preferences along with the content description. Hybrid Content Filtering, a different strategy, is also employed for this goal.

In order to address the sparsity and cold start difficulties that are related to the earlier approaches, Hybrid Content Filtering incorporates both Content and Collaborative filtering techniques [3]. Although there has been a lot of work done in the field of recommendation systems, the idea of using machine learning algorithms to improve the outcomes is still relatively new. Making autonomous models that can learn from available data and automate operations is at the heart of machine learning. Similar outcomes are anticipated when applied for content suggestion purposes. The machine learning-based app recommendation system is a field with plenty of development opportunities. It is well recognised that recommendation systems enhance the decision-making process and decision-making quality [4]. The foundation of this project is the usage of a machine learning algorithm to propose apps to users. Users were shown screenshots of several apps that had been gathered and presented. Features are extracted and given to the machine learning model based on user preferences and screenshots of the apps. The algorithm picks up knowledge from these extracted properties and applies it to future app recommendations for consumers. The level of similarity between the apps enjoyed by users and the recommended apps serves as the foundation for app recommendations.

LITERATURE SURVEY:

P. P. P. D. A., P. Singh, "Recommender systems: an overview, research trends, and future directions," Int. J. Business and Systems Research, vol. 15, pp. 14-52, 2021.

Recommender systems (RS) have become a focus of significant research because they aim to make it easier for users to find products online by making recommendations that accurately match their preferences. The many recommendation approaches, related problems, and information retrieval strategies are all covered in-depth in this paper's analysis of the RS. It has generated study interest among a sizable number of scholars worldwide because of its extensive uses. This paper's primary goal is to identify the RS research trend. Since 2011 through the first quarter of 2017, more than 1,000 research publications from ACM, IEEE, Springer, and Elsevier have been taken into consideration. This work has produced a number of intriguing discoveries that will aid RS researchers—both present and future—in evaluating and planning their future directions. This work also looks ahead to RS's future, which could lead to new areas of study in this area.

Fuad, Ahlam, Sahar Bayoumi, and Hessah Al-Yahya. "A Recommender System for Mobile Applications of Google Play Store." International Journal of Advanced Computer Science and Applications 11.9 (2020).

Numerous applications are available for customers to download as the market for smartphones expands. When searching for a mobile app that would satisfy their demands, users have difficulty due to the abundance of available options. Application recommendations that are specifically tailored to the user are in high demand.

A system that identifies user interest based on what the system feels the user likes through his or her profile is required due to the significant and increasing quantity of mobile applications that are now available in app stores. A user profile would help with a productive and unique application screening system. Getting a sub-collection of applications based on a chosen category is the general notion behind filtering. Information filtering can be done in a variety of ways, including through classification and suggestion.

A. S. U. Z. A. S. Q. Asifullah Khan, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," Artificial Intelligence Review, vol. 53, pp. 5455-5516, 2020.

An outstanding sort of neural network known as the Deep Convolutional Neural Network (CNN) has excelled in a number of competitions involving computer vision and image processing. Image Classification and Segmentation, Object Detection, Video Processing, Natural Language Processing, and Speech Recognition are a few of the intriguing application domains of CNN. Deep CNN uses a number of feature extraction steps that are capable of automatically learning representations from the data, which contributes significantly to its strong learning capabilities. The research on CNNs has advanced because to the availability of a lot of data and advancements in hardware technology, and interesting deep CNN architectures have recently been described. The employment of various activation and loss functions, parameter optimization, regularisation, and architectural advances are just a few of the intriguing concepts being investigated to advance CNN technology.

Fayyaz, Zeshan, et al. "Recommendation systems: Algorithms, challenges, metrics, and business opportunities." ***applied sciences*** **10.21 (2020)**

Users frequently receive recommendations from recommender systems based on their preferences. Recommender systems have shown to be an effective technique to combat information overload due to the ever-growing number of information available online. Given its potential to help resolve numerous over-choice issues, the use of recommender systems cannot be emphasised. There are many distinct kinds of recommendation systems, each with its own principles and methodology. The e-commerce, healthcare, transportation, agricultural, and media industries are just a few that have used recommendation systems. The current state of research in recommender systems is presented in this project, along with future directions in the field for diverse applications. An overview of the types, issues, restrictions, and commercial applications of recommendation systems is given in this article, which also discusses the state of the art in this area. In the study, qualitative evaluation measures are described as a way to rate a recommendation system's quality.

C. Pu, Z. Wu, H. Chen, K. Xu and J. Cao, "A Sequential Recommendation for Mobile Apps: What Will User Click Next App?" 2018 IEEE International Conference on Web Services (ICWS), 2018.

Due to the expansion of smartphones and the App stores, the number of mobile applications is exponentially growing. Users can download a variety of Apps that offer useful services for practically every part of modern life, including socialising, listening to music, watching videos, and browsing the web, to name a few. On a smartphone, there can be anywhere from 10 and 90 installed apps, with 50 installed apps being the average. Although consumers have the option to instal a variety of Apps on their smartphones to expand device functionality, it can be difficult to constantly search for and choose which App to use while also having to wait for the particular App to load. To improve selection, it is possible to group Apps into folders. Users still need to browse and sort through directories, though.

METHODOLOGY:

Due to a lack of knowledge about data visualisation, it is a bit difficult to deploy machine learning algorithms in the present system. In the current approach, constructing models is done by mathematical computations, which can be very difficult and time-consuming. We employ machine learning tools from the Scikit-Learn toolkit to get around all of this.

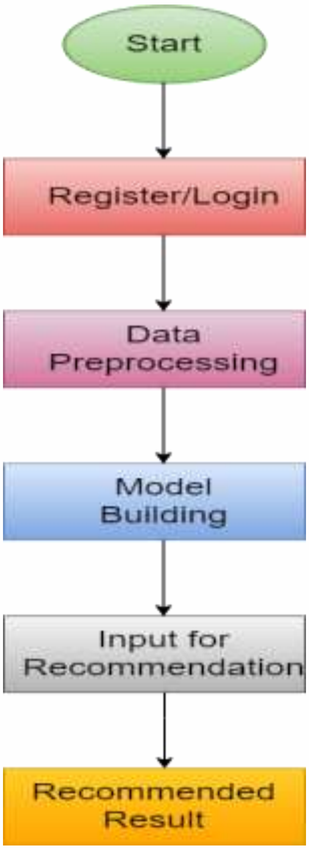

Fig 1: Block diagram for recommendation system

Content-based Filtering

They are more arbitrary and description-based, meaning that each item that needs to be categorised has a unique description, and suggestions are solely dependent on the user's profile. The type of an entity is determined by its keywords or classes, which are then added to the user's existing profile. Suggestions then include entities from the recognised type and types that are similar to it. Following the creation of a user profile, a number of weighted factors are determined, where weight denotes importance, using a variety of techniques like Bayesian Classifiers, cluster analysis, decision trees, and artificial neural networks to predict the likelihood that a user will express interest in any given entity. Additionally, since content-based filtering only considers the user and their past, it ignores the user's surroundings, peers, and future projections (i.e., if their preferences change) [5].

Collaborative Filtering

On the other hand, collaborative filtering categorises a group of users based on matching criteria and similarities to other users. One benefit is that since it types humans rather than things, it can suggest a variety of entities to a group of users once the users' types have been determined. It presupposes that a group of users who have previously expressed interest in an entity will continue to do so until the entire user type has changed. A suggestion for a particular user can therefore be viewed as something that his or her group has given the user

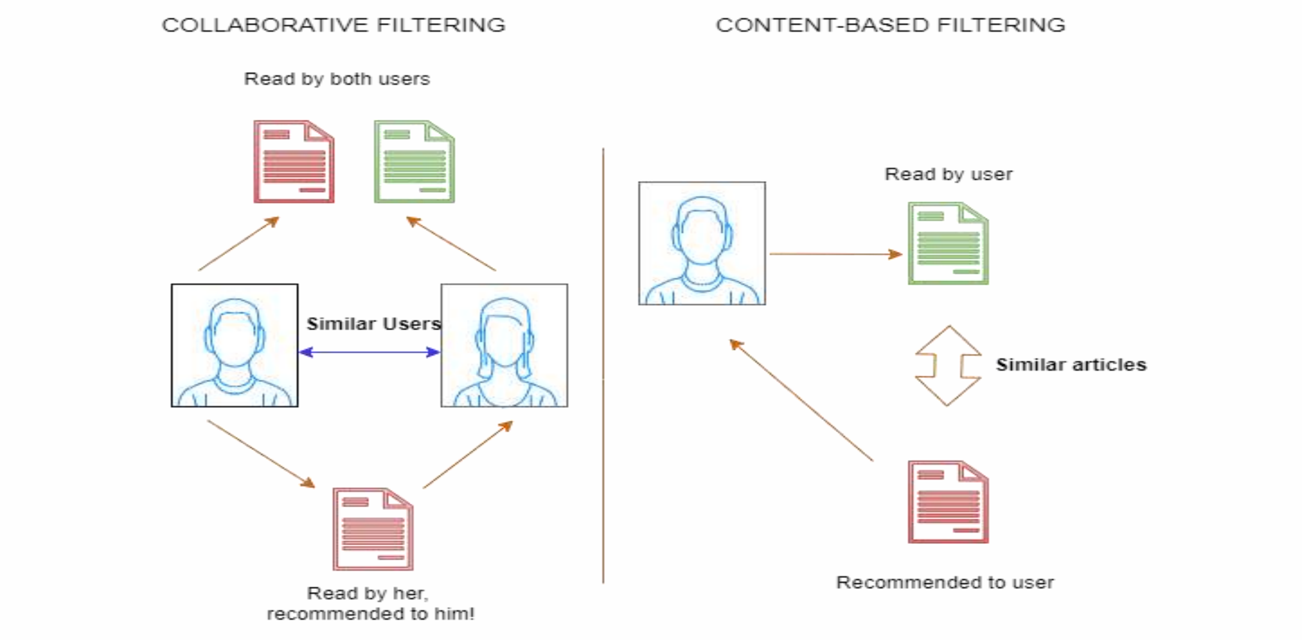

Fig 2: Collaborative vs Content-based filtering

Hybrid Recommendation System

A hybrid recommendation system combines several different kinds of recommendation systems into one. Due to its reliance on many systems and ability to calculate multiple parameters at once, it can deliver a more accurate range of suggestions. It can be done in a variety of ways, such as by building several content-based and collaborative filtering systems, merging the results, and then presenting them as one, or by converting content-based categorization into collaborative classification or the other way around.

IMPLEMENTATION AND RESULTS:

In the suggested approach, the recommendation system will be built using unsupervised machine learning, and a content-based filtering mechanism has already been put in place. Here, we must enter the index number of the particular app to quickly receive up to 10 recommended apps. . The architecture of the recommendation system is shown below:

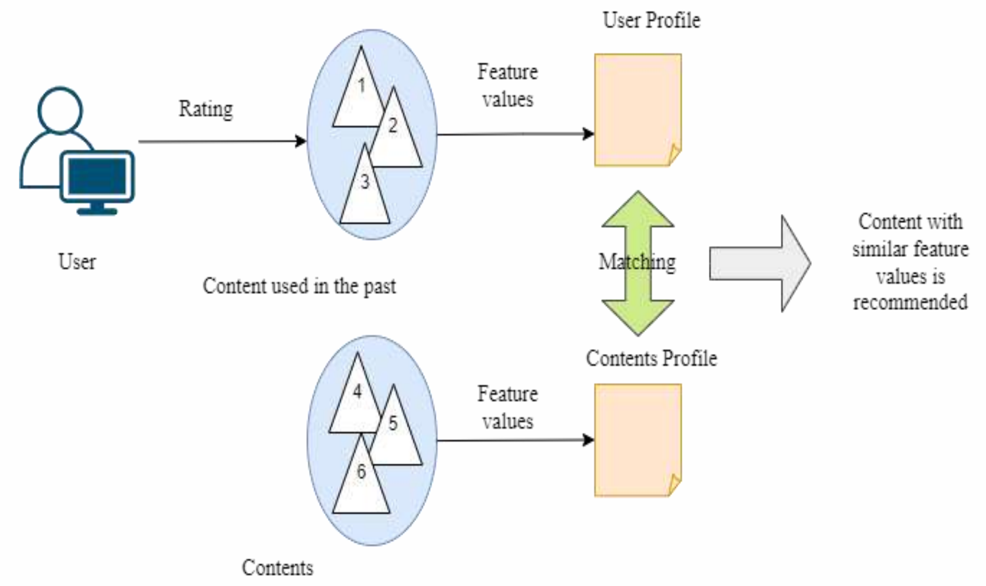

Fig 3: Architecture for recommending app to user

The first and most important phase was data collecting and data pre-processing because our recommendation system is dependent on screenshots. Through the use of Python scrapers, we first downloaded the meta data for several apps. Various categories were assigned to the screenshots based on the metadata that was scraped from each one. Business, education, entertainment, games, music, news, photography, social media, sports, and tools are some of the genres for Android apps. Then, a model for app recommendation was created using unsupervised learning. Models for machine learning translate inputs to the dataset's outputs. The dataset was split into a train set and a test set in order to both train and assess the model's performance. 1000 photographs each category made up the training set, which had a total of 10,000 images. For the test dataset, 200 photos from each category were kept, making a total of 2000 screenshots. The distribution of the dataset is summarised in Table 1.

Table 1: Dataset Distribution	
Total images	12000

Total App categories	10
Images in Training dataset	10000
Images in testing dataset	2000

Three convolutional layers, two dense layers, and a SoftMax classifier made up the architecture of the CNN model[6]. Evaluation of the first training cycle showed that we could only attain a 10% accuracy. Consequently, we moved forward with the model's fine-tuning. Following the start of the model's convolutional base, Vgg16 pre-trained ImageNet weights were loaded. A previously trained model was put on top of the layers that came before the final convolutional blocks. The model's accuracy increased because to this practise, and it now has a 48.5% accuracy. It's crucial to clarify, though, that since the project's objective is to create a recommendation system, depending just on the model's accuracy as an evaluation metric would not be a prudent course of action. How relevant the suggested apps are to the consumers' tastes serves as the true metric for measuring the model's performance. Accuracy only serves as a proxy for how well features are extracted using CNN, and is not a true indicator of its quality. The accuracy metric was preserved by the model once it reached the 48.5% accuracy target, and despite numerous attempts at retraining, it remained at this level. Our dataset was highly diversified, which is why we had such a poor level of accuracy. Using only the CNN architecture along with the Vgg16 weights would have resulted in a considerably higher accuracy score if the experiment were repeated for apps in the same category. The CNN design uses several feature extraction steps to improve the quality of the features extracted. However, feature extraction was used in order to deploy Linear Regression because the objective was to increase the model's app recommendation dependability as much as feasible.

Features were extracted for the purpose of building the feature vector after the trained model and weights were loaded[7]. We were able to minimise the Root Mean Square and keep it within the range of 0.15 to 0.3 at the conclusion of this method. Given comparable projects with a similar goal of developing recommendation systems, this was a significant accomplishment. The user's was given the model's app recommendations. The apps that the consumers loved were ultimately displayed to them. Users' initial tastes are fairly closely mirrored by the apps that are recommended to them.

Since there were three distinct methodological phases to the project, the findings can be broadly divided into a total of three categories. The CNN was implemented in the first phase, which resulted in a 10% accuracy at best. Further advancement was made, and the accuracy rose to 48.5%, when the pre-trained Vgg16 ImageNet weights were integrated with the CNN model architecture. Then, based on the screenshots, features were derived from the model weights and the user choice. For the purposes of implementing linear regression, these two feature vectors were utilised.

The Root Mean Square Error, which was found to be in the range of 0.15-0.3, was used to evaluate the effectiveness of linear regression. Since the model's correctness and performance are highly reliable and dependable due to the Root Mean Square Error's noticeably low value.

CONCLUSION:

A recommender system was developed to assist consumers in finding and obtaining mobile applications of interest. Based on an analysis of prior research and related systems, the system is a prototype. By exploiting the user's location, it gives suggested applications from the moment a user launches the system, demonstrating the promise of application suggestions and highlighting the domain's limitations. There have been preliminary assessments. Recommender systems are becoming more necessary in this field as a result of the vast number of mobile applications that are now being created and deployed. The abundance of systems attempting to address this issue demonstrates the need. Systems like the one described in this paper demonstrate considerable promise for assisting end users, and there is also business opportunity. There is value in the data the system produces, but recommendations of applications would enhance the value of an application store or portal.

Future development of mobile application recommendation systems faces a number of difficulties. Research on recommender systems faces both general and unique obstacles, some of which are related to applications or context-awareness. This section outlines these difficulties, first in a broad sense and then by focusing on elements unique to the further development of the prototype system described in this report. When examining the issue mentioned in this study, there are a variety of perspectives that can be taken into account. One is that of the end user, who needs to quickly and easily find intriguing applications without sacrificing privacy. There are researchers who must comprehend customers in order to modify algorithms and system interactions to satisfy user wants while analysing and documenting system features. Then there is the commercial perspective, which focuses on methods to make money from the recommendations generated. This viewpoint may have an impact on research, which may have an impact on the system's general design, as well as on how useful the system is to users.

REFERENCES

[1] P. P. P. D. A., P. Singh, "Recommender systems: an overview, research trends, and future directions," Int. J. Business and Systems Research, vol. 15, pp. 14-52, 2021.

[2]Fuad, Ahlam, Sahar Bayoumi, and Hessah Al-Yahya. "A Recommender System for Mobile Applications of Google Play Store." International Journal of Advanced Computer Science and Applications 11.9 (2020).

[3] R. G. R. G. R. V. F. Y. B. Pathak, "Empirical analysis of the impact of recommender systems on sales," J Manage Inform Syst, vol. 27, no. 2, pp. 159-188, 2010.

[4] V. J. M. Neha Sharma, "An Analysis Of Convolutional Neural Networks For Image Classification," Procedia Computer Science, vol. 132, pp. 377-384, 2018.

[5] A. S. U. Z. A. S. Q. Asifullah Khan, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," Artificial Intelligence Review, vol. 53, pp. 5455-5516, 2020.

[6] Fayyaz, Zeshan, et al. "Recommendation systems: Algorithms, challenges, metrics, and business opportunities." applied sciences 10.21 (2020): 7748.

[7] L. Y. A. S. Y. T. Shuai Zhang, "Deep Learning Based Recommender System: A Survey and New Perspectives," ACM Computing Surveys, pp. 1-38, 2019.

[8] Su, Xin, et al. "An edge intelligence empowered recommender system enabling cultural heritage applications." IEEE Transactions on Industrial Informatics 15.7 (2019): 4266-4275.

[9] Raza, Shaina, and Chen Ding. "Progress in context-aware recommender systems—An overview." Computer Science Review 31 (2019): 84-97.

[10] Millecamp, M., Htun, N.N., Conati, C. and Verbert, K. (2019) 'To explain or not to explain: the effects of personal characteristics when explaining music recommendations', in Proceedings on Intelligent User Interface.

[11] K. Ochiai, F. Putri, and Y. Fukazawa, "Local app classification using deep neural network based on mobile app market data," 2019 IEEE Int. Conf. Pervasive Comput.

Commun. PerCom 2019, pp. 186–191, 2019.

[12] V. Viljanac, "RECOMMENDER SYSTEM FOR MOBILE APPLICATIONS," Multimed. Tools Applications., vol. 77, no. 4, pp. 4133– 4153, Feb. 2018.

[13] Alian, Shadi, Juan Li, and Vikram Pandey. "A personalized recommendation system to support diabetes self-management for American Indians." IEEE Access 6 (2018): 73041-73051

[14] Y. L. D. X. X. F. R. G. Donghui Wang, "A content-based recommender system for computer science publications," Knowledge-Based Systems, pp. 1-9, 2018.

[15] C. Pu, Z. Wu, H. Chen, K. Xu and J. Cao, "A Sequential Recommendation for Mobile Apps: What Will User Click Next App?," 2018 IEEE International Conference on Web Services (ICWS), 2018.

[16] Jisha, R. C., Ram Krishnan, and Varun Vikraman. "Mobile applications recommendation based on user ratings and permissions." 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2018.

[17] A. Arampatzis and G. Kalamatianos, "Suggesting Points-of-Interest via Content-Based, Collaborative, and Hybrid Fusion Methods in Mobile Devices," ACM Trans. Inf. Syst., vol. 36, no. 3, pp. 1–28, 2017.

[18] D. Surian, S. Seneviratne, A. Seneviratne, and S. Chawla, "App Miscategorization Detection: A Case Study on Google Play," IEEE Trans. Knowl. Data Eng., vol. 29, no. 8, pp. 1591–1604, 2017.

[19] Cao, Hong Lin, Miao. (2017): Mining smart phone data for app usage prediction and recommendation: A survey pervasive and mobile Computing. 37. 1-22.

[20] V. Radosavljevic et al., "Smartphone App Categorization for Interest Targeting in Advertising Marketplace," 2017, pp. 93–94.

[21] Y. B. F.O.Isinkaye, "Recommendation systems: Principles, methods and evaluation," Egyptian Informatics Journal, vol. 16, no. 3, pp. 261-273, 2015.

[22] Su, Xin, et al. "Android app recommendation approach based on network traffic measurement and analysis." 2015 IEEE Symposium on Computers and Communication (ISCC). IEEE, 2015.

[23] JovianLin, Kazunari Sugiyama, MinYen Kan, Tat Seng Chua: New and Improved modeling versions to improve apps recommendation - SIGIR '14: Proceedings of the 37th international ACM SIGIR conference on Research development in information retrieval,2014.

[24] V. W. X. X. Q. Y. Zheng, "Towards mobile intelligence: Learning from GPS history data

for collaborative recommendation.," Artificial Intelligence, Vols. 184-185, pp. 17-37, 2012.

[25] R. G. De Souza, R. Chiky, and Z. K. Aoul, "Open source recommendation systems for mobile application," CEUR Workshop Proc., vol. 676, pp. 55–58, 2010.

[26] Y. R. B. C. V. Koren, "Matrix factorization techniques for recommender systems," IEEE, vol. 42, no. 8, pp. 30-37, 2009.

[27] W. Woerndl, C. Schueller, and R. Wojtech, "A Hybrid Recommender System for Context-aware Recommendations of Mobile Applications," in 2007 IEEE 23rd International Conference on Data Engineering Workshop, 2007.

[28] Devika, P., R. C. Jisha, and G. P. Sajeev. "A novel approach for book recommendation systems." Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on. IEEE, 2016.

[29] M. R. Islam, "Numeric rating of Google Play Store applications by sentiment analysis based on user reviews," in 2014 International Conference on Electrical Engineering and Information & Communication Technology. Dhaka, Bangladesh: IEEE, April 2014, pp. 1-4.

[30] Hsieh, Meng-Yen, Wen-Kuang Chou, and Kuan-Ching Li. "Building a mobile movie recommendation service by user rating and APP usage with linked data on Hadoop." Multimedia Tools and Applications 76.3 (2017): 3383-3401.