Research Result

Efficient Machine Learning Technique for Rating Prediction of Google Play Store Apps

Dimple Ghole[1] and Mr. Sarwesh Site[2]

[1]MTech Scholar, Department of Computer Science Engineering, All Saint's College of Technology, Bhopal (M.P.), INDIA
[2]Assistant Professor, Department of Computer Science Engineering, All Saint's College of Technology, Bhopal (M.P.), INDIA

ABSTRACT

Online reviews play a significant role in influencing the decision of consumers. Consumers show their experience and information about quality in their reviews. In the case of Google Play store fake numeric ratings can play a big role in the success of apps. People tend to believe that a high-star rating may be significantly attached with a good review. However, user star level rating information does not usually match with text format of review. This paper presents efficient machine learning technique for rating prediction of google play store apps.

KEYWORDS

Machine Learning, Google,Play Store, ,Online, Rating

1. INTRODUCTION

Machine learning approaches are essential for us to take care of numerous issues. Machine learning has numerous applications in numerous perspectives and has incredible advancement potential.

It is unsurprising that AI could set up ideal hypotheses to explain its presentations. Meanwhile, its abilities of unaided learning will be improved since there is a lot of data in the world anyway adding names to all of them isn't important. It is also guessed that brain framework designs will end up being progressively unusual with the objective that they can isolate every one of the more semantically significant features. Furthermore, significant learning will combine with help adjusting better and we can use this focal points to accomplish more tasks.

In the present situation we can see that portable applications assuming a significant part in any singular's life. It has been seen that the improvement of the versatile application publicize unquestionably affects progressed advancement. Having said that, with the reliably creating flexible application feature there is also a prominent rising of convenient application originators definitely achieving high as can be pay by the overall compact application industry.

With huge test from wherever all through the globe, it is fundamental for an originator to understand that he is going on in the right heading. To hold this pay and their spot in the market the application fashioners might have to sort out some way to stick into their current position. The Google Play Store is seen to be the greatest application stage. It has been seen that notwithstanding the way that it makes in excess of two crease the downloads than the Apple Application Store yet makes simply an enormous piece of the money diverged from the Application Store. Along these lines, I scratched data from the Play Store to coordinate our assessment on it.

Figure 1: Mobile App

With the quick improvement of cutting-edge cells, versatile applications (MobileApps) have ended up being essential bits of our lives. In any case, it is problematic as far as we're concerned to track with the reality and to comprehend everything about the applications as new applications are entering market every day. It is represented that Android1market accomplished an enormous piece of 1,000,000 applications in September 2011. Beginning at now, 0.675 million Android applications are open on Google Play Application Store. Such a ton of uses are apparently an unprecedented entryway for clients to buy from a wide assurance broaden. We trust flexible application clients consider online application studies as an

important effect for paid applications. It is pursuing for an expected client to examine every one of the scholarly comments and rating to make a decision. Moreover, application engineers experience issues in finding how to further develop the application execution subject to by and large alone and would benefit by figuring out an immense number of printed comments.

2. BACKGROUND

R. Gomes et al., work expects to make induction motors, permitting the expectation of use appraisals, utilizing the KNN and Arbitrary Woods relapse strategies. The Arbitrary Timberland showed improved results than the KNN [1]. C. Zhu et al., present the disconnected examinations for three huge scope datasets approve the predominant exhibition of Point. A three-week online A/B test in the proposal administration of a standard application market shows that Point further develops DeepFM model by 4.4% as far as CTR [2].

G. S. Bhat et al., track down a relationship between's the particular matter (PM) found inside and the external climate with the PEFR. The PEFR results are ordered into three classifications, for example, 'Green' (Safe), 'Yellow' (Moderate Gamble) and 'Red' (High Gamble) conditions in contrast with the best pinnacle stream esteem acquired by every person. Convolutional brain organization (CNN) design is utilized to plan the connection between the indoor PM and climate information to the PEFR values [3]. Z. Wu et al., found three variables making the irregularity among portrayals and consent uses be: 1) human mediations recorded as a hard copy depiction; 2) terrible practices on authorization utilizations; and 3) productive engineers. These discoveries will work with engineers to refine application depictions and upgrade authorization utilizations in the applications [4].

Z. Shen et al., work expects to foresee a bunch of applications a client will open on her cell phone in the following time allotment. Such a data is fundamental for the vast majority cell phone tasks, e.g., application pre-stacking and content pre-reserving, to further develop client experience [5]. Z. Xu et al., propose another cross-trio profound element inserting strategy, called CDFE, for cross-application JIT bug expectation task. The CDFE technique integrates a cutting-edge cross-trio misfortune capability into a profound brain organization to learn undeniable level element portrayal for the cross-application information [6].

K. Zhao et al., direct tests on 10 Android portable applications and exploratory outcomes show that SDF performs fundamentally better compared to relative strategies as far as 3 execution pointers [7]. G. Aceto et al.,[8] contrast the outcomes and a few ML draws near, showing execution tantamount to a best-in-class ML indicator (Irregular Timberland Regressor). Likewise, with this work we give a reasonable and hypothetically sound traffic-examination toolset to help further developing ML assessment (and perhaps its plan), and a reasonable and interpretable benchmark [8].

Y. Zhang et al., creates staggered modules in view of Repetitive Brain Organization with consideration component and produces multi-step time series forecasts by combining the results of modules. Probes a genuine

world dataset show that DeePOP beats cutting edge techniques in expectation exactness, really lessening the Root Mean Square Blunder (RMSE) to 0.088 [9].

Q. Zhu et al., present a decency mindful Application proposal strategy named Homestead. The chief investigation of this technique underscores on the reasonableness issue during the proposal interaction. In this strategy, Application up-and-comers are isolated into high perceivability and low perceivability Applications, and execute proposal calculation separately [10].

S. şahın et al., a clever way to deal with choice criticism unwavering quality assessment through web-based expectation is proposed and applied for SISO FIR DFE with either deduced likelihood or assumption engendering based delicate input. This original strategy for channel calculation is displayed to further develop recognition execution contrasted with recently known elective techniques [11]. S. Rezaei et al., this work, propose a profound learning model for versatile application recognizable proof that works even with encoded traffic. The proposed model just requirements the payload of the initial not many bundles for grouping, and, thus, appropriate in any event, for applications depend on early expectation, for example, directing and QoS provisioning [12].

3. METHODOLOGY

The methodology or the flow of the work is as followings-
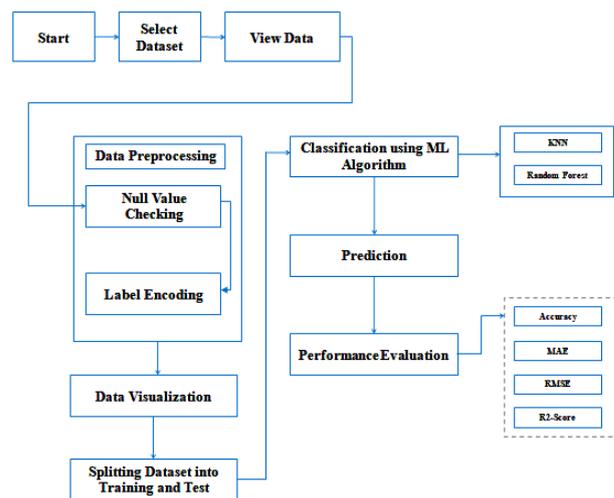

Figure 2: Flow Chart

The proposed model is introduced to overcome all the disadvantages that arise in the existing system. This system will increase the accuracy of the classification results by classifying the data based on the googleplay apps. The dataset collected from the Google Play store is semi structured or unstructured and contains significant superfluous data (defined as not contributing significantly to the prediction process). Training a supervised machine learning algorithm requires textual documents to be represented in vectorial form. For this purpose, textual data must be converted into numbers without losing information. Apply Machine Learning algorithms to predict app's rating with Regressors (KNN and Random Forest). Finally, the Performance of classifiers is evaluated on the accuracy, MAE, RMSE and R2-Score.

4. SIMULATION RESULTS

The execution of the proposed calculation is done over python spyder 3.7. The sklearn, numpy, pandas, matplotlib, pyplot, seaborn, os library assists us with utilizing the capacities accessible in spyder climate for different strategies.

Figure 3: Dataset frame

Figure 3 is showing the dataset in the python environment. The dataset have various numbers of rows and column. The signal features name is also mentioned.
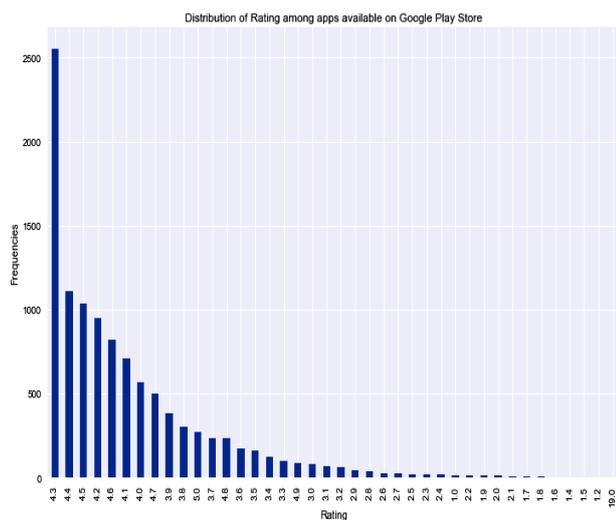

Figure 4: Distribution of rating

Figure presents the distribution of the goggle play store rating prediction. There are maximum ratings is 4.9 and minimum rating is 1.0. The maximum and minimum rating app is less, the average rating of good app lie between 3.9 to 4.5.
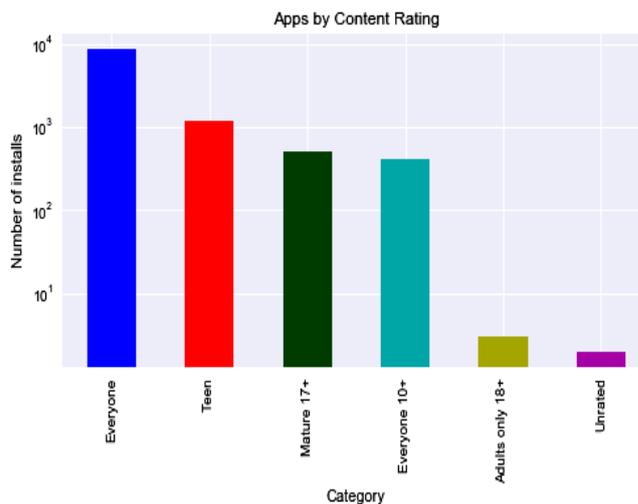

Figure 5: Apps by content rating

Figure 5 is presenting the various apps by the content rating, there are most installed apps which is used by everyone, other app downloads is less like apps for teen, mature, unrated etc.
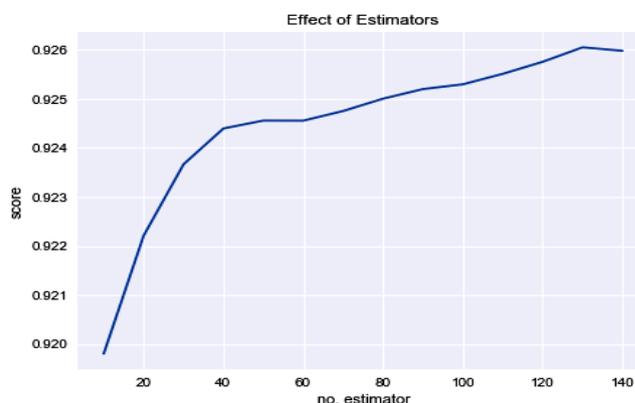

Figure 6: Effect of estimators

Figure 6 presents the effect of the estimators; the total number of the estimators is upto 140. The total score is aprox 95% for multiple check.

Table 1: Simulation Results (KNN Regression)

Sr. No.	Parameter Name	Value
1	Accuracy	95.13%
2	MAE	0.27466%
3	RMSE	0.4478 %
4	R2_Score	0.9213%
5	Error rate	4.87%

Table 2: Result Comparison

Sr. No.	Parameters	Previous Work [1]	Proposed Work
1	Accuracy	93.8%	95.41%
2	Error rate	6.2 %	4.59%

5. CONCLUSION

This paper presents efficient machine learning technique for rating prediction of google play store apps. The increasing number of Android apps available on Google Play Store with the developers' advantages has attracted many Android apps developers' attention. To benefit from developing Android apps is to know the characteristics of high rated applications on the Google Play Store. The overall accuracy is achieved by the proposed technique is 95.41% while previous it is achieved by the 93.8%. The error rate is 4.59% in the proposed work and the 6.2% by the previous work. Therefore, the proposed efficient technique archived better results than the previous.

REFERENCES

[1] R. Gomes da Silva, J. de Oliveira Liberato Magalhães, I. R. Rodrigues Silva, R. Fagundes, E. Lima and A. Maciel, "Rating Expectation of Google Play Store applications with use of information mining strategies," in IEEE Latin America Exchanges, vol. 19, no. 01, pp. 26-32, January 2021, doi: 10.1109/TLA.2021.9423823.

[2] C. Zhu et al., "Point: Programmed Connection Machine for Active visitor clicking percentage Expectation," in IEEE Exchanges on Information and Information Designing, doi: 10.1109/TKDE.2021.3134985.

[3] G. S. Bhat et al., "AI Based Asthma Chance Forecast Utilizing IoT and Cell phone Applications," in IEEE Access, vol. 9, pp. 118708-118715, 2021, doi: 10.1109/ACCESS.2021.3103897.

[4] Z. Wu, X. Chen, M. U. Khan and S. U. - J. Lee, "Improving Loyalty of Portrayal in Android Applications With Class Based Normal Consents," in IEEE Access, vol. 9, pp. 105493-105505, 2021, doi: 10.1109/ACCESS.2021.3100118.

[5] Z. Shen, K. Yang, Z. Xi, J. Zou and W. Du, "DeepAPP: A Profound Support Learning Structure for Versatile Application Use Expectation," in IEEE Exchanges on Portable Figuring, doi: 10.1109/TMC.2021.3093619.

[6] Z. Xu et al., "Exertion Mindful In the nick of time Bug Expectation for Portable Applications By means of Cross-Trio Profound Element Implanting," in IEEE Exchanges on Unwavering quality, doi: 10.1109/TR.2021.3066170.

[7] K. Zhao, Z. Xu, T. Zhang, Y. Tang and M. Yan, "Worked on Profound Woods Model Based In the nick of time Imperfection Expectation for Android Portable Applications," in IEEE Exchanges on Dependability, vol. 70, no. 2, pp. 848-859, June 2021, doi: 10.1109/TR.2021.3060937.

[8] G. Aceto, G. Bovenzi, D. Ciuonzo, A. Montieri, V. Persico and A. Pescapé, "Portrayal and Expectation of Versatile Application Traffic Utilizing Markov Displaying," in IEEE Exchanges on Organization and Administration The board, vol. 18, no. 1, pp. 907-925, Walk 2021, doi: 10.1109/TNSM.2021.3051381.

[9] Y. Zhang, J. Liu, B. Guo, Z. Wang, Y. Liang and Z. Yu, "Application Prevalence Forecast by Consolidating Time-Fluctuating Progressive Communications," in IEEE Exchanges on Versatile Registering, doi: 10.1109/TMC.2020.3029718.

[10] Q. Zhu, Q. Sun, Z. Li and S. Wang, "Ranch: A Reasonableness Mindful Suggestion Strategy for High Perceivability and Low Perceivability Portable Applications," in IEEE Access, vol. 8, pp. 122747-122756, 2020, doi: 10.1109/ACCESS.2020.3007617.

[11] S. şahın, A. M. Cipriano, C. Poulliat and M. Boucheret, "Iterative Choice Criticism Evening out Utilizing On the web Expectation," in IEEE Access, vol. 8, pp. 23638-23649, 2020, doi: 10.1109/ACCESS.2020.2970340.

[12] S. Rezaei, B. Kroencke and X. Liu, "Enormous Scope Versatile Application Recognizable Proof Utilizing Profound Learning," in IEEE Access, vol. 8, pp. 348-362, 2020, doi: 10.1109/ACCESS.2019.2962018.

[13] S. Zhao et al., "Orientation Profiling From a Solitary Preview of Applications Introduced on a Cell phone: An Observational Review," in IEEE Exchanges on Modern Informatics, vol. 16, no. 2, pp. 1330-1342, Feb. 2020, doi: 10.1109/TII.2019.2938248.

[14] C. Min et al., "Adaptable Power Effect Forecast of Portable Detecting Applications at Pre-Establishment Time," in IEEE Exchanges on Versatile Figuring, vol. 19, no. 6, pp. 1448-1464, 1 June 2020, doi: 10.1109/TMC.2019.2909897.