

Design an Approach for Prediction of Music Prediction System by XGBClassifier & Random Forest

Neelu Sharma¹, Rajesh Boghey², Sandeep Rai³

¹M.Tech Scholar, Department of CSE, TIT Excellence, Bhopal (M.P), India

²Assistant Professor, Department of IT, TIT, Excellence, Bhopal (M.P), India

³Professor & Head, Department of IT, TIT, Excellence, Bhopal (M.P), India

ABSTRACT: Data mining played very important roles from last decade now this data mining is migrated to Machine Learning. Machine Learning comes by Artificial Intelligence and Mathematical Statistics. Machine learning is categorized into supervised, unsupervised and reinforcement. Supervised machine is using the learning algorithm to detect the discovery that is clearly due to the examples supplied to produce general interpretation, which then predicts future scenarios or events. In this Paper Authors explains with the advancement of technology in music players, especially in intelligent cell phones, users have access to large archives. Here we analyse the music portal KKBOX in this work we analyse the different dependent attributes their behaviour and others. Finally, we apply two algorithms Random Forest & XGBClassifier.

Keywords: Classification Algorithm, Machine Learning, Deep Learning, Decision Tree & Random Forest, ensemble Techniques, CNN.

I. INTRODUCTION

In some previous years we all are talking about machine intelligent programs, and what this factor for our health issues, richness. Now days , machine related intelligent learnt program possibly pledge to protect growth of population , address worldwide challenges such as environment change, and add much of income to the world's economy through expanding potency; so that it also essentially make changes in the nature of work, and Size, or define, do the choices in everyday life. Between these extremes, there is a potentially variable technology, which brings both its opportunities and challenges, and whose risks and benefits are needed to navigate because it is used to become more central to everyday activities.

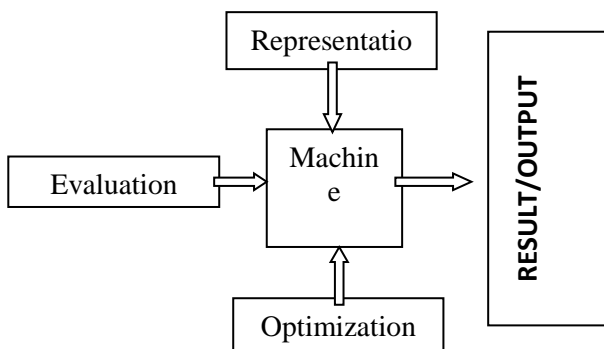


Figure 1: Major Component in Machine Learning

1.1 Data Mining Process

Data mining process contains many steps like Data cleaning, Data integration, Data selection, Data transformation, Data mining, Pattern evaluation, Knowledge presentation.

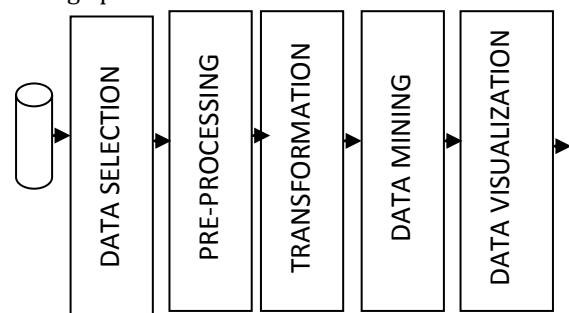


Figure 2: Process of Data Mining

Since we know that data mining process is complex process where we have to apply number of intermediate process. During Intermediate process we concentrate that how data can be filter or refine to the pattern generation. We know that every intermediate process plays very important roles during pattern finding. When we reach to data mining process, we have numbers of Data mining algorithms for solving our problem.

We know that Data mining is very growing field now days. Every Industries is trying to explore their previous Data for Next Upcoming Business Year. In previous some years it has shown that with the exploring previous Data you can improve your business.

1.2 Data Mining Techniques

Data mining Algorithms is categorized into different which is given below:

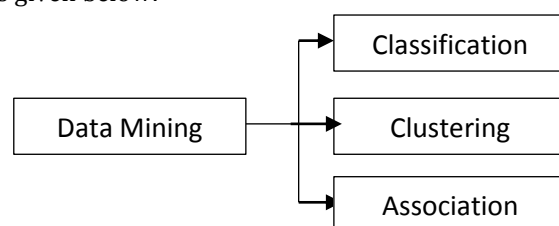


Figure 3: Types of Data Mining

Data mining is categorized in three parts i.e. classification, Clustering & Association. Classification deals with supervision of some level where we classify on the basis of some parameters. Clustering deals with unsupervised mechanism. Finally, Association deals with how one data create penetration over another data.

1.3 Machine Learning

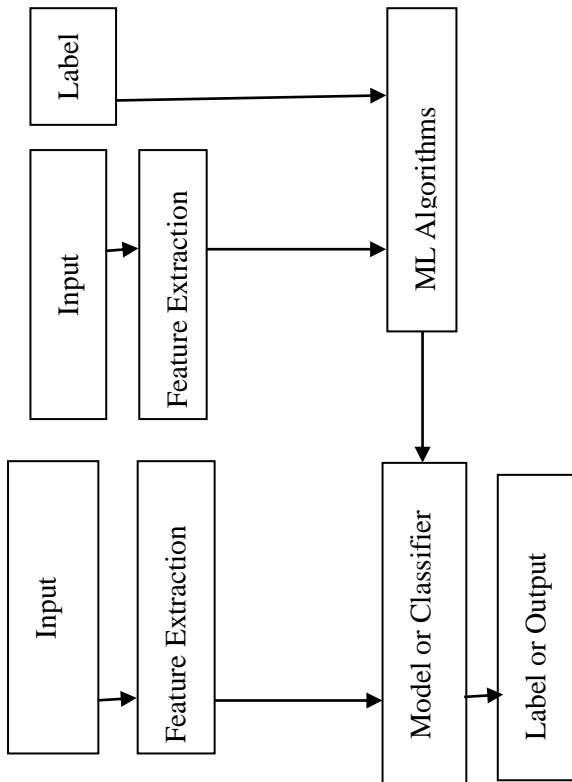


Figure 4: Machine Learning Sketch Diagram

Data In the above figure we explained about machine learning major components available. Here major work done by ML Algorithms.

II. LITERATURE SURVEY

According to the paper, Nowadays the recommender system is a software used as an important tool of e-commerce, which helps in analyzing the tastes of users and providing them with a list of products that they want to enjoy. This paper is an investigation of using collaborative filtering techniques for a music recommending system. Collaborative filtering is a technique that focuses on making relationships between users and making a prediction between objects. For experimental purpose we explore different matrices to measure the similarity of users and objects such as Euclidean distance, cosine metric, Pearson correlation and others. Finally, we compare various evaluation metrics that represent the effectiveness of the recommender system. [1].

With the advancement of technology in music players, especially in intelligent cell phones, users have access to

large archives. The quick and easy selection of favorite music in these large archives becomes one of the biggest problems for users. In this paper, a system is designed such that it collects users' reference information such as weather, temperature, geographical location, etc., and according to their weighted combination, it recommends a suitable music that is currently the user's favorite is. . Thus, the system incorporates a rating method that determines how close the musical references are to those that have been played in the earlier context and recommends the music that has the most proximity. The results of this research suggest that the recommendations that this system makes under different circumstances are closed to the user's liking [2].

III. PROBLEM IDENTIFICATION

We know that all music giant available online like songs.com, gana.com and many more. We know that we are living in 21st century where all customers like online availability of products so that they can get their products easily. When we studied many Product Recommendation System related papers, we found that many techniques are available to find the prediction. We read number of research papers and find that every algorithm has some specification in terms of their performance. Author decided that he will give an algorithm which has better Accuracy. Since we know that deep learning or neural network gives better result in terms of performance so we will apply Random Forest & SGDCClassifier for better performance.

IV. Block Diagram & Methodology

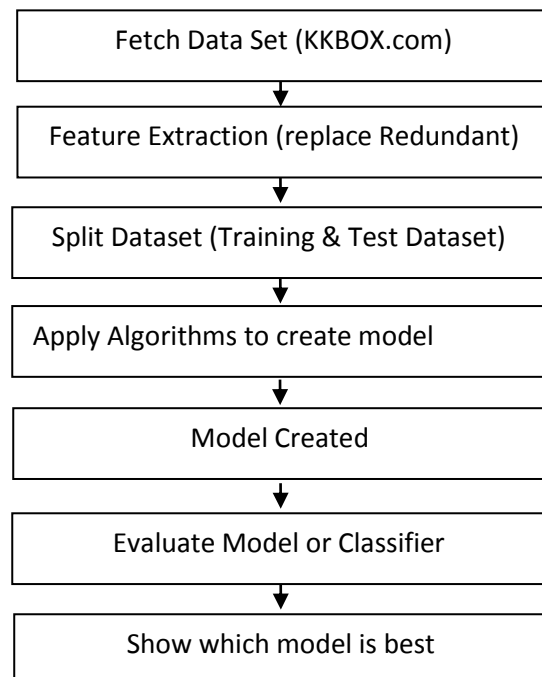


Figure 5: Block Diagram of Flow of Operation

V. Algorithms Used

Step 01: Store Data from Kaggle Repository

Step 02: Import Prior Libraries:

Step03: Now Import our Required Dataset

Step04: Apply Feature Extraction

Step 05: Visualize Data for better understanding with the help of matplotlib

Step06: Repeat step05 Multiple Times for Better Understanding the Data

Step07: Before calling any Model Firstly call required Libraries & Train_Test_Split function

Step08: Apply Different Model

Step09: Repeat Step08 for many times with different Algorithms & with different Data split values

Step10: Finally Compare Results with performance parameters like Accuracy

VI. Experimental Result

We used python programming Language to implement our logic we used number of libraries like NumPy, pandas, tweepy, matplotlib, seaborn and many more. This project is divided into two parts in first part we fetch data in another part we will process our data our overall process is given below:

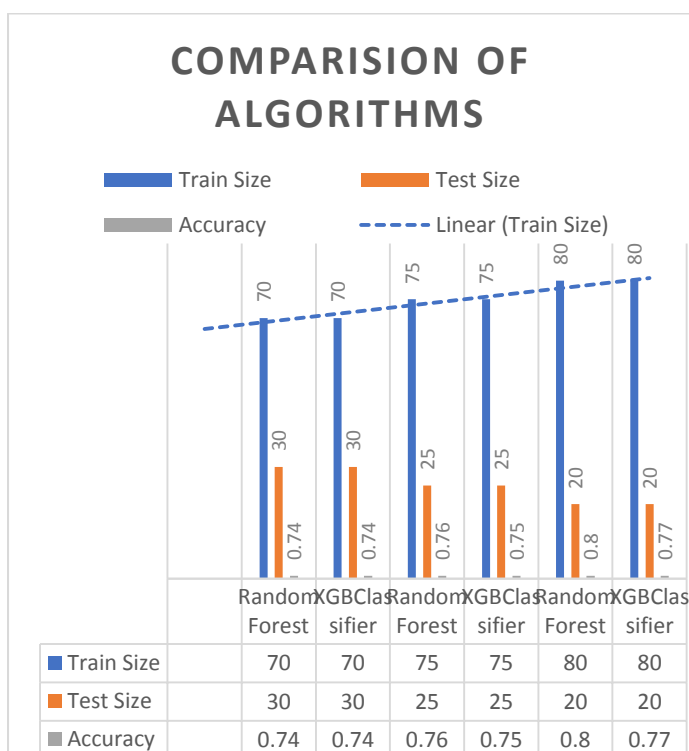


Figure 6: Graph Comparison

Explanation: By the analysis of above graph we can say that splitting of training and testing data may change the accuracy which is given above in terms of graphical values. We split our data into three different slot . In every split we found the different values.

VII. CONCLUSION

We conclude the observation about the previous technique. Our observation different terms and condition. This represents our work in a new approach. In our work we have tried to improve Accuracy we applied ensemble techniques where we applied two different algorithms i.e. Random Forest algorithm and XGBClassifier algorithm in every algorithm we got different Accuracy %. Again, we adjust training and testing splitting (% values) by this mechanism we get better result.

REFERENCES

[1] Z. Hyung, J. S. Park, and K. Lee, "Utilizing context-relevant keywords extracted from a large collection of user-generated documents for music discovery," *Info. Processing and Management*, vol. 53, no. 5, pp. 1185-1200, 2017.

[2] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," *The Adaptive Web: Methods and Strategies of Web Personalization*, pp. 291-324, 2007.

[3] M. J. Pazzani and D. Billsus, "Content based recommendation systems," *The Adaptive Web: Methods and Strategies of Web Personalization*, pp. 325-341, 2007.

[4] E. J. Humphrey, J. P. Bello, and Y. LeCun, "Moving beyond feature design: deep architectures and automatic feature learning in music informatics," in *Proc. 13th Int'l Conf. Music Info. Retrieval*, pp. 403-408, October 2012.

[5] C. W. Hsu, and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Tran. Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[6] R. E. Fan, K. W. Chang, C. J. Hsieh, X. R. Wang, and C. J. Lin, "LIBLINEAR: a library for large linear classification," *J. Machine Learning Research*, vol. 9, pp. 1871-1874, 2008.

[7] A. V. Oord, S. Dieleman, and B. Schrauwen, "Deep content-based music recommendation," in *Proc. 26th Int'l Conf. Neural Info. Process. Systems*, pp. 2643-2651, December 2013.

[8] J. Salamon, and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Sig. Processing Letters*, vol. 24, no. 3, pp. 279-283, March 2017.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Info. Processing Sys.*, pp. 1097-1105, 2012.

- [10] R. Jozefowicz, W. Zaremba, and I. Sutskever, "An empirical exploration of recurrent network architectures," in *Int'l Conf. Machine Learning*, pp. 2342-2350, June 2015.
- [11] J. S. Downie, "Music information retrieval," *Annual Review of Information Science and Technology*, 37:295-340, 2003.
- [12] R. Typke, F. Wiering, and R. Veltkamp, "A survey of music information retrieval systems," *Proc. ISMIR*, pp. 153-160, 2005
- [13] D. Jang, C.-J. Song, S. Shin, S.-J. Park, S.-J. Jang, S.-P. Lee, "Implementation of a matching engine for a practical query-by-singing/humming system," *Proc. ISSPIT*, pp. 258-263, 2011
- [14] S. W. Hainsworth and M. D. Macleod, "Particle filtering applied to musical tempo tracking," *EURASIP J. Applied Signal Processing*, vol. 15, pp. 2385-2395, 2004
- [15] S. Jo and C. D. Yoo, "Melody extraction from polyphonic audio based on particle filter," *Proc. ISMIR*, pp. 357-362, 2010