Predicting a Hit Song with Machine Learning: Is there an apriori secret formula?

Conference Paper · July 2020

DOI: 10.1109/DATABIA50434.2020.9190613

2 authors, including:

Krishnadas Nanath
Middlesex University Dubai
26 PUBLICATIONS **133** CITATIONS

SEE PROFILE

2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA)

Predicting a Hit Song with Machine Learning: Is there an apriori secret formula?

Agha Haider Raza
Middlesex University
Dubai, United Arab Emirates
AR1416@live.mdx.ac.uk

Krishnadas Nanath
Middlesex University
Dubai, United Arab Emirates
username.krishna@gmail.co

***Abstract*— Thought to be an ever-changing art form, music has been a form of recreational entertainment for ages. The music industry is constantly making efforts for songs to be a hit and earn considerable revenues. It could be an interesting exercise to predict a song making it to top charts from a mathematical perspective. While several studies have looked into factors after a song is released, this research looks at *apriori* parameters of a song to predict the success of a song. Data sources available from multiple platforms are combined to create a dataset that has technical parameters of a song and sentimental analysis of the lyrics. Four machine learning algorithms (Logistic Regression, Decision Trees, Naïve Bayes and Random Forests) to answer the question-Is there a magical formula for the prediction of hit songs? It was found that there are elements beyond technical data points that could predict a song being hit or not. This paper takes a stand that music prediction is yet not a data science activity.**

Keywords— Machine learning, Supervised learning, API, Spotify, Audio Features, Sentiment Analysis, Hit Song, Billboards, Data Science, Android App development

I. INTRODUCTION

Music is an art form that has been around for centuries. Constantly in flux, it has proved to be a very diverse field. With revenue numbers in millions of dollars each year, the music industry is growing rapidly. In 2018 the recorded music industry was worth $19.1 Million [1]. With the emergence of computer technology and storage becoming affordable, an unprecedented amount of electronic data collection has been observed [2]. The data collection could vary from medical records, financial records, consumer data to fields more artistic fields like movies and songs.

Music trends undergo rapid changes in short periods. The success of any track released depends on various factors- both *a priori* and *posteriori*. The parameters involved in *posteriori* could include play-count, user downloads, featuring in top charts and others. These parameters are interesting from the recommender systems point of view and used by several platforms to suggest songs for their users. However, from a music producer, artist, and studio perspective, it would be interesting to analyze the success of a song *a priori*. A system that predicts the success of a song based on its technical properties would be of great benefit to music producers. It can decipher the expected response to a song from the listener's point of view, by not only considering the audio features as well as the sentiment of the lyrics. While there are several studies available to understand and develop a model based on *posteriori properties,* this research focuses on a priori parameters.

A system may be able to detect a hit song as well as a flop song with impressive accuracy levels [3]. The Artificial Intelligence would extensively aid users to get a reasonable estimate on the appeal of a track. Breaking down the features of a song to predict whether the song will be a hit before it has been released would significantly enable producers to maximize their profits and diminish risks that they take with producing songs. The features can be processed using machine learning, and patterns may be identified in the processed data, which can finally be employed effectively to predict hit songs. Effectively, this research attempts to answer the question – Is there a secret formula to predict a hit song just by looking at its properties *apriori*.

While a few studies have looked into Billboard top 100 for the prediction of hit songs, this research goes beyond the data available and uses other platforms like Spotify to collect more information on parameters of top songs. These parameters include technical properties of music and are restricted to a priori properties as it is the focus of this study. While great accuracies are reported with *posteriori* properties like user perception, twitter response, and others [4], it would be interesting to see if similar accuracies can be obtained with a priori properties.

This research uses two types of data points related to a song -audio features of a track and the sentimental analysis of the lyrics. Data from various sources like Spotify, Billboard, open data lyrics were aggregated and subjected to machine learning algorithms to identify if a formula for hit song exists. It was found that the accuracies reported in previous studies with posteriori properties cannot be attained with *aprioi*, and it might be a tough task to get a magic formula when the prediction of a hit song is considered.

The paper proceeds are as follows. Section-2 looks at the literature review to highlight similar work that has been executed in the past. Section-3 elaborates on the process of data collection and processing with a detailed explanation of various data sources and preparation. Section-4 explains the various algorithms of machine learning used for this research and also compares the results for evaluation. The paper then concludes with future research and summary in Section-5.

II. RELATED WORKS

The use of algorithms in music started with algorithmic music compositions studies back in the 1990s. As stated by Bruce Jacob [5], "Algorithmic composition is as old as music composition" and therefore are the studies associated with it. This review has its focus more on the use of machine learning and artificial intelligence to predict a song being hit or not.

A few studies in the past have looked at the *posteriori properties* of a song to predict success. These factors could include user perception, social media strategy, promotions, release platforms, and others. It was interesting to note, a song itself (internal factors, *apriori*) had a relatively minor role than the external factors like social influences to be a hit [6]. Zangerla et al [7] used Tweets to predict future charts and found that it was useful.

This research investigates a question that would help interested stakeholders to predict the success before a song is

978-1-7281-9792-0/20/$31.00 ©2020 IEEE

released. A few studies have looked at some internal factors for this purpose. Support Vector Machine was used by Dhanaraj and Logan [8] to predict if a chart would appear in music charts based on latent topic features. The literature has also used Music Information Retrieval (MIR) research and did not find it useful for the prediction of hits [9]. Some support was received, though from Ni et al. [10] where they demonstrated that some audio features could correlate with the evolution of music trends. However, they did not include the hit prediction aspect. Therefore, it would be interesting to get into this space with more advanced factors of a song (along with the lyrics) to understand if the prediction of hit songs could be a reality.

The set of articles collected for the literature review of this research resulted in the following themes: Web API (Application Programmable Interface), Dataset, Hit Song Prediction, Neural Networks, Machine Learning, Deep Learning, Artificial Intelligence, Audio Features, Sentiment Analysis. It displays the importance of machine learning algorithm in the prediction of hit songs. The review was mainly conducted from three perspectives- Data Collection, Algorithms in Analytics, and Data Representation. A summary of these domains is provided in Table-I.

Table I: AREAS OF RESEARCH

Data Collection	Data Science	Data Representation
Web API's Online Databases Global Music Charts Compiling Dataset	Audio Features Analysis Text Sentiment Analysis Algorithms Accuracy	Data Flow Evaluation Quantitative Representation

Data Collection: Data Collection is essential to any research. Many online sources and APIs can be used to collect data related to music and songs. Billboard, as a top chart, has been used in several studies [11]. Other options explored earlier include 'Jsoup' library and Echo Nest- a tool to extract audio features. In another study, The Ultra Top 50, a Belgian website and The Bubbling Under Chart was used to create a dataset from 2011-2013. This resulted in a total of 8750 songs, of which 982 were unique [12]. Other datasets are compiled by other users and are available online on websites such as Github and Kaggle. Anwuri [13] used a similar dataset to carry out research targeting the shift in popular music dating back to 1958 [14]. To get an audio feature, the Spotify API was used. Since Spotify acquired an echo nest, the audio features are the same. The study additionally used python pandas and a python package called spotipy to execute the data collection.

Another element of data collection is the lyrics for the available songs. There are several methods used in the literature that range from Genius Lyrics API [15] to Python scripts scraping lyrics from open sources available [16].

Data Science: The use of analytics in this domain is classified into two parts- algorithms used for predictions and the text analysis related to the lyrics. The algorithms used in predictions included C4.5 Tree, Naïve Bayes, Logistic Regression, and Support Vector Machines [8]. All these studies included the concept of test and train to work on the accuracy of the predictions. Text analysis included sentimental analysis on the lyrics adopting various techniques. A study used lyric.rip for the text sentiment analysis that was incorporated using a Markov chain [15]. In another study, data was analyzed from the Stanford sentiment treebank with a Recurrent Neural Network (RNN) known as LSTM (Long Short-Term Memory). RNN was used since it remembers the past sequences [17].

Data Representation and Analysis: From the collection of data, various researchers have drawn up graphs which accurately depicts the finding of their studies. By mapping out the performance of each algorithm, the accuracy percentage of each algorithm has been showcased. In research carried out by Pachet and Roi [9], the results were expressed in three tiers of popularity as opposed to being represented with accuracy percentages. They marked out the features that they had obtained to make their dataset and then used them to direct which feature made a song popular. The popularity labels were marked Low, Medium, and High. The features included generic features, Specific features, and Human features. Another study [18] looks at words that contribute to the sentiment of a song. The words that correlate were visualized for the following categories- metal, pop, rap, reggae, rock, and soul.

III. RESEARCH METHOD

This section provides the details of data collection, the process of extracting audio features for the collected songs, lyrics collection mechanism, and performing the sentiment analysis. Furthermore, it contains details on the steps performed to pre-process the data, and, finally, the details of machine learning algorithms are provided for the prediction model.

A. Data Collection

The song titles were collected from the Billboard Hot 100 Charts, and this included both hit and non-hit songs. A song was classified as Hit when the song appeared in the top 20 at the end of each month. Similarly, for the non-hit songs, the criteria were for the titles to be amongst the bottom ten songs at the end of each month. The song titles from 3 consecutive years were collected from 2017 to 2019. These titles, along with the name of the artists and their release year, were placed into two separate CSV files. The total entries in each file after this step represented 720 hit songs and 360 Non-hit songs. After the data was compiled, the redundant entries were removed, which left the final dataset with the composition displayed in Figure-1.

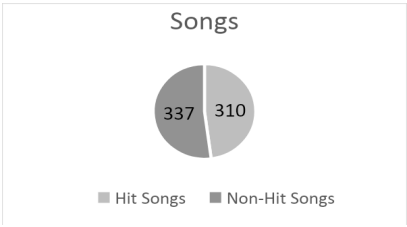

Figure 1: Hit songs composition in the edited dataset

Once the initial dataset was completed, the Spotify ID for the titles were collected and programmed from the Spotify app. The audio features were extracted from the Spotify API with the aid of an access token provided to the developers for

integrating features into applications. By using the Spotify ID previously acquired manually, the API carrying parameters were substituted that included the Spotify ID and the access token, and the features were stored in CSV files. Most audio features were extracted except for those that were not relevant for the research. These audio features, along with their data types and descriptions, are provided in Table-II.

Table II: AUDIO FEATURES

Feature	Data Type	Description
Danceability	*Float*	Danceability describes how suitable a track is for dancing based on a combination of musical elements. A value of 0.0 is least danceable, and 1.0 is most danceable.
Energy	*Float*	Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity.
Key	*Int*	Predicts whether a track contains no vocals
Loudness	*Float*	The overall loudness of a track in decibels (dB). Loudness is the quality of a sound that is the primary psychological correlate of physical strength (amplitude). Values typical range between -60 and 0 db.
Mode	*Float*	Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.
Speechiness	*Float*	Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording, the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words.
Acousticness	Float	A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
Instrumentalness	Float	Predicts whether a track contains no vocals. The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content.
Liveness	Float	Detects the presence of an audience in the recording.
Valence	Float	A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track.
Tempo	Float	The overall estimated tempo of a track in beats per minute (BPM).

B. Sentiment Analysis Method

In an attempt to strengthen the dataset, this study included the lyrics of the songs for analysis. The addition of lyrics and its associated variables makes this research unique. To the best of our knowledge, a combination of technical parameters and sentimental parameters has not been used in one model to predict the nature of hit songs. Lack of freely available APIs led to the scrapping process where open sources of the lyrical database were used and then added to the dataset. This text associated with the lyrics was then subjected to sentimental analysis to enhance the dataset further. The sentiment analysis was conducted using Microsoft Azure Machine Learning. A plugin for Microsoft Excel was used to get the sentiment for each song with a Score. This output (with scores) was added to the file consisting of the audio features and the Year of release.

Azure ML package offers advanced natural language processing over raw text. It also gives a sentiment score between 0 and 1 for each document, where 1 indicates a strong sentiment. The Azure ML training process involves a large dataset of labeled text records. It tokenizes the text into words as the first step and then applies the stem method. The features from these words are then extracted and used to train a classifier. When the classifier has been trained, it could then be used to predict the sentiment of a new piece of text.

C. Data Preparation

The first step of making the final dataset was to combine both the files together. Another variable was added in each file named Hit to type Boolean. The Hit songs were expressed with '1' and the Non-Hit songs with '0'. Both files were then combined to make the final dataset. Entries with inconsistencies (wrong data or missing values) were removed from the dataset. Songs of languages other than English were removed from the dataset as well. At this point, all the variables were checked, and irrelevant ones were removed, such as Spotify ID and Language Indicator. The final dataset after cleaning consists of 310 Hit songs and 337 Non-Hit songs, as shown in Figure-1.

D. Machine Learning Algorithms

In order to develop a prediction model on *apriori* properties of songs to predict success, a range of machine learning algorithms was used. Four algorithms were selected based on their previous application in a similar context- Logistic Regression, Decision Tree, Random Forests, and Naïve Bayes. All the algorithms were implemented in R, and the performance was also verified using RapidMiner.

Logistic Regression: For solving classification problems, Logistic Regression is the most basic and popular amongst machine learning algorithms [19]. The model is based on predicting the probabilities of success using the logistic function. Since the objective here is to predict a song being hit or not, the logit function of a hit song is used as a dependent variable, and all the parameters of the song, including the sentiment, are independent variables. The function can be represented as :

$$log\left(\frac{P(HitSong)}{1-P(Hitsong)}\right) = \beta_0 + \sum \beta_i x_i$$

The list of xi can be represented as Song Name + Artist Name + Release + Danceability + Energy +Key + Loudness + Mode + Speechiness +Acousticness + Instumentalness + Liveness + Valence + Tempo + Sentiment + Score

Decision Tree & Random Forests: The reason to include a decision tree in the suit was to offer a prediction with a case where the assumptions of logistic regression are not met. Capable of working effectively with non-linear data, it differentiates itself from most machine learning algorithms. As the algorithm uses one feature per node to split the data, the construction of a decision tree is a fast-paced process, with higher accuracy in results. Referred to as an extension of Decision Tree, Random Forests is a machine learning algorithm that can be used to solve both Classification and Regression problems. They operate like an ensemble consisting of many individual decision trees [20] that operate with consensus to get a more accurate outcome than the individual trees.

The overall problem statement remains the same as the previous one (logistic regression), where the objective of a tree is to classify Hit songs and Non-hit songs by working on the various parameters in the dataset.

Naïve Bayes: It is a Supervised Machine Learning algorithm based on the Bayes Theorem that is used to solve classification problems by following a probabilistic approach. It's called Naïve since the features used are assumed to be independent of each other [19]. This algorithm was chosen because it requires a small amount of training data to train to produce effective results. In this research, if y denotes the label "Hit Song" and x denotes the feature vector (independent variables listed in logistic regression), the conditional probability P(x |y) is indicated by:

$$P(x \vee y) = \Pi P(x_j \vee y)$$

The posterior probability can then be calculated for feature vector x belonging to a hit song by looking at P(y=1|x)

IV. RESULTS

This section is divided into two parts. The first part describes the results of the sentimental analysis that was introduced in this research. It gives an idea about the overall corpus of lyrical text that is part of the dataset. The second part describes the results of the various machine learning algorithms that were used in the research.

A. Results of Sentiment Analysis

As the first step of the analysis, the correlation of the independent variables (sentimental analysis scores) and the dependent variable was analyzed. Plots were generated in R by using a collection of libraries that were used to generate these plots. It was interesting to observe the split of positive, negative, and neutral sentiments in the two sets of Hit and Non-Hit Songs. While the composition of these sentiments within a set almost remained the same, it was an exciting finding to see the majority of the songs being tagged as negative sentimental value. A few other observations could also be important to observe a difference between the two sets. Hit songs had a relatively higher number of positive sentiment songs than the Non-Hit ones (18% in Hit compared to 12% in Non-Hit). The percentage of negative sentiments was lower in Hit Songs (80%) compared to Non-hit songs (86%). This could provide an early indication that happy songs could have a higher probability of being classified as a hit song. However, since the difference in numbers is not huge, it cannot be generalized. A simple t-test revealed that there were no significant differences between the two groups. The results are summarized in Figure-2. A scatterplot matrix illustrates the relationship between the Sentiment, Score, Release, and Hit is provided in Figure-3.

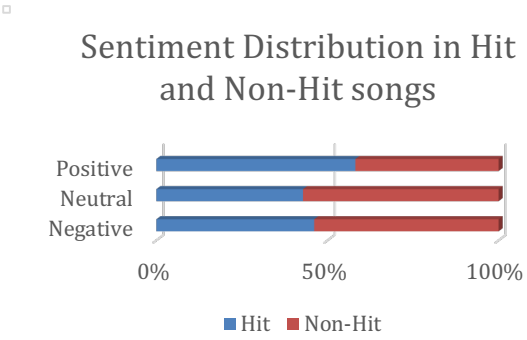

Figure 2: Sentimental Analysis

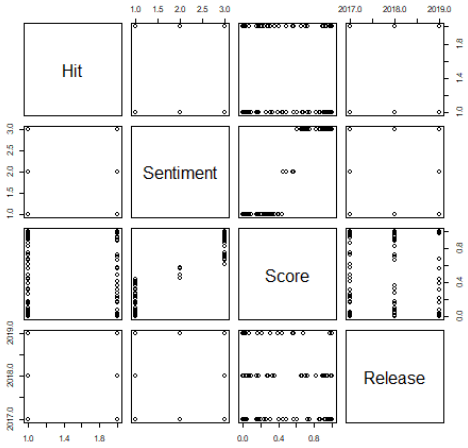

Figure 3: Scatterplot Matrix

B. Model Building and Evaluation

R studio was used to implement the machine learning models used in this research. It is an open source tool for statistical computing and graphics. Several visualizations were used first to get an early indication of the variables used to build the model. The concept of test and train was used to determine the accuracy of the model. Firstly, the data was split with 550 in the training dataset and 88 in the testing data set. Subsequent executions involved k-fold validation to eliminate any bias in the test and train dataset. It was ensured that the training model had sufficient data to learn from both the Hit and Non-Hit songs. The distribution of Hit and Non-hit songs was 48% and 52%, respectively in the train dataset. While the testing dataset had 49% Hit songs and 51% Non-Hit songs.

Evaluation Metrics: Confusion matrix was used to measure the accuracy of the models. A confusion Matrix is described as a matrix that illustrates the performance of a model with regards to the testing data [21]. There can be four different outcomes: True Negatives (TN), True Positives (TP), False Negatives (FN), and False Positives (FP). In the models made for this project, a hit may be classified as a non-hit song,

which is a False Negative or a non-hit may be classified as a hit, which is a False Positive. The accuracy of the model is determined by the following formula:

$$ACC = \frac{TP + TN}{TP + TN + FN + FP}$$

Although R Studio had several functions that could be used for the confusion matrix, for this project, a table was plotted for the predictions against the results from the test data set. For classification problems, Precision is widely used. Precision is defined as the percentage of tuples that are labeled positive and are indeed positive. It can be interpreted as a measure of exactness. Precision can be numerically expressed as:

$$Precision = \frac{TP}{TP + FP}$$

Results: A feature importance plot was plotted to understand the impact of all features in classifying a Hit song vs. a Non-hit song. It is observed that the Danceability was the most critical feature in determining the success of a song, and the least important feature was the Key. The feature importance plot is provided in Figure-4.

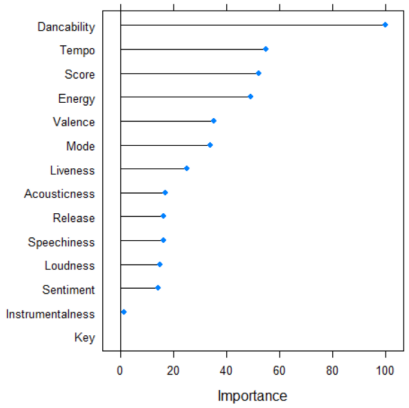

Figure 4: Feature Importance Plot

Results of Machine Learning Models: The results from all the models are presented in Table III. Accuracy and Precision are two variables that are provided for the four algorithms- Logistic Regression, Decision Tree, Random Forests, and Naïve Bayes. The findings of this project reveal that the most accurate model was trained using the Logistic Regression algorithm, with an accuracy of 52%. The least accurate model, on the other hand, was the Decision tree though the difference was not significant. In comparison, all the models performed almost similarly with no major distinction between the results of the different models. However, the features in consideration proved to be different in each of the models. It is also important to look at the weights of the parameters that helped the prediction model. These weights can be obtained from the results of the logistic regression and could be useful for remodeling in some algorithms. The weights of these parameters are provided in Table 1V.

Table III: Machine Learning Algorithms

Model	Accuracy	Precision
Logistic Regression	52.0%	0.5
Decision Tree	50.5%	0.428
Random Forests	51.0%	0.488
Naïve Bayes	51.1%	0.49

Table IV: Machine Learning Algorithms

Parameter	Weight	Parameter	Weight
Release	-0.12	Speechiness	1.00
Dancability	2.21	Acousticness	0.55
Energy	-0.98	Instrumentalness	-1.49
Key	-0.02	Liveness	-0.90
Loudness	0.13	Valence	-0.08
Mode	-0.13	Tempo	-0.002
Senitment Neutral	1.23	Sentiment Positive	0.77
Score	-1.05		

V. Discussion and Conclusion

In all the models, Danceability was amongst one of the most significant features, if not the most significant. Other features that proved to be significant were Energy, Valence, Tempo, Speechiness, and Loudness. Even though the Sentiment and Score were expected to be amongst the core features which determined the success of a song, this expectation was not reflected in the results. These two features were amongst the least significant variables in all the models. This possibly gives an early indication that there might not be a magic formula for the prediction of hit songs based on *apriori*. There could be an element of the year of release being a significant variable. However, the results of the predictions make it evident that even the year of release did not play a significant feature in this study. The prediction accuracies were almost the same as flipping a coin for the two outcomes of Heads and Tails. It brings back the classic debate of whether analytics in the music industry is a science that could transform into prediction experience.

Usually, tracks are not based on any algorithmic compositions and are integrated with different melodies and progressions using different instruments. Nowadays, electronic music is a prevalent method of composing unique sounds. Identifying musical instruments and popular melodies can provide more insight into this research. As Rightfully pointed out by Pachet and Roy [9] the audio features are not informative enough to predict the success of a song. Extracting significant features from the audio tracks has the potential to improve the results drastically.

Future Directions: Evaluating the music features by using algorithms can be a future endeavor. This would ensure the consistency of data and particular features to be extracted that may prove to be more significant than the features extracted from Spotify's API. An in-depth breakdown of the lyrical sentiment is something that could enhance future work. Breaking down an amalgamation of slang words, onomatopoeia, and deeper meanings to use in future models will be worked on. A research on the artist ranking and fan following that plays a vital role in the success of a song could be brought in as *apriori* parameters. By judging the success of the previously released songs, reasonable conclusions can

be drawn for future releases. Another major key feature that is excluded from this research is the genre of the song. Therefore, more parameters can be brought in to take a more durable stand on the debate- is there a magic formula? The analytics could further be scaled up

Conclusion: This research included an array of parameters that could help predict a song being hit. Based on the results, this paper adds to the stand that there is no magic formula yet that could predict a song being hit before it is released. It could be interesting to extend the research and see if perception variables (post-release) could change the stand on the debate. This research combined technical properties with sentimental analysis of lyrics to make a stronger model for the implementation of machine learning algorithms. It would be interesting to see this field advancing and probably getting to a point where hit songs can be predicted before the release.

REFERENCES

[1] H. McIntyre, "The Global Music Industry Hit $19 Billion In Sales In 2018, Rising By Almost 10%", Forbes, 2019. [Online]. Available: https://www.forbes.com/sites/hughmcintyre/2019/04/02/the-global-music-industry-hits-19-billion-in-sales-in-2018-jumping-by-almost-10/#51faf8e418a9. [Accessed: 30- Jun- 2020].

[2] K. McNulty, "What is Machine Learning?", towardsdatascience, 2018. [Online]. Available: https://towardsdatascience.com/what-is-machine-learning-891f23e848da. [Accessed: 30- Jun- 2020].

[3] M. Neal, "VICE - A Machine Successfully Predicted the Hit Dance Songs of 2015", *Vice.com*, 2015. [Online]. Available: https://www.vice.com/en_us/article/bmvxvm/a-machine-successfully-predicted-the-hit-dance-songs-of-2015. [Accessed: 30- Jun- 2020].

[4] Y. Kim, B. Suh and K. Lee, "#nowplaying the future billboard: mining music listening behaviors of twitter users for hit song prediction", SoMeRA '14: Proceedings of the first international workshop on Social media retrieval and analysis, pp. 51-56, 2014. Available: 10.1145/2632188.2632206.

[5] B. Jacob, "Algorithmic composition as a model of creativity", Organised Sound, vol. 1, no. 3, pp. 157-165, 1996. Available: 10.1017/s1355771896000222.

[6] [6] M. Salganik, P. Dodds and D. Watts,"Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market", Science, vol. 311, no. 5762, pp. 854-856, 2006. Available: 10.1126/science.1121066.

[7] E. Zangerle, M. Pichl, B. Hupfauf and G. Specht, "Can Microblogs Predict Music Charts? An Analysis of the Relationship Between #Nowplaying Tweets and Music Charts", Proceedings of the 17th International Society for Music Information Retrieval Conference, ISMIR, New York City, United States, pp. 365-371 (978-0-692-75506-8)., 2016. Available: 10.5281/zenodo.1417881.

[8] R. Dhanaraj and B. Logan, "Automatic prediction of hit songs," in Proceedings of International Society for Music Information Retrieval, pp. 11–15, 2005.

[9] [9] F. Pachet and P. Roy, "Hit song science is not yet a science," in Proceedings of International Society for Music Information Retrieval, pp. 355–360, 2008.

[10] Y. Ni and R. Santos-Rodriguez, "Hit song science once again a science," in Proceedings of International Workshop on Machine Learning and Music, pp. 2–3, 2011.

[11] D. Herremans, D. Martens, and K. Sörensen, "Dance hit song prediction", Journal of New Music Research, 43(3), 291–302,2014.

[12] D. Herremans, T. Bergmans, "Hit song prediction based on early adopter data and audio features"in Proceedings of The 18th International Society for Music Information Retrieval Conference (ISMIR) — Late Breaking Demo. Shuzou, China , 2017

[13] R. Anwuri, "Billboard Hot 100 Analytics: Using Data to Understand The Shift in Popular Music in The Last 60…", *Medium*, 2018. [Online]. Available: https://towardsdatascience.com/billboard-hot-100-analytics-using-data-to-understand-the-shift-in-popular-music-in-the-last-60-ac3919d39b49. [Accessed: 30- Jun- 2020].

[14] "Billboard - Music Charts, News, Photos & Video | Billboard", Billboard, 2020. [Online]. Available: https://www.billboard.com/. [Accessed: 30- Jun- 2020].

[15] E. Fu, "A Teen Programmer Built A Tool To Generate Fake Lyrics For Your Favorite Artists", *Genius*, 2019. [Online]. Available: https://genius.com/a/a-teen-programmer-built-a-tool-called-lyrics-rip-to-generate-fake-lyrics-for-your-favorite-artists. [Accessed: 30- Jun- 2020].

[16] E. Çano, M. Morisio, "Moodylyrics: A sentiment annotated lyrics dataset". In Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence pp. 118-124, 2017.

[17] P. Shreyas, "Sentiment analysis for text with Deep Learning", *Medium*, 2019. [Online]. Available: https://towardsdatascience.com/sentiment-analysis-for-text-with-deep-learning-2f0a0c6472b5. [Accessed: 30- Jun- 2020].

[18] B. Toth, "From Metallica to Adele—Text Analysis of successful song lyrics with R", Medium, 2019. [Online]. Available: https://towardsdatascience.com/text-analysis-of-successful-song-lyrics-e41a4ccb26f5. [Accessed: 30- Jun- 2020].

[19] Z. Lateef, "Comprehensive Guide To Logistic Regression In R | Edureka", Edureka, 2019. [Online]. Available: https://www.edureka.co/blog/logistic-regression-in-r/. [Accessed: 30- Jun- 2020].

[20] T. Yiu, "Understanding Random Forest", Medium, 2019. [Online]. Available: https://towardsdatascience.com/understanding-random-forest-58381e0602d2. [Accessed: 30- Jun- 2020].

[21] C. Dinant, "What's so naive about naive Bayes'?", Medium, 2018. [Online]. Available: https://towardsdatascience.com/whats-so-naive-about-naive-bayes-58166a6a9eba. [Accessed: 30- Jun- 2020].

[22] K. M. Ting, Confusion Matrix, Springer US, Boston, MA,pp. 260–260,2017.