**Business Analytics project on**

**MUSIC PLAYER APPLICATION**

**Describe the business process and/or rule related with the business.**

Assume that a music app called “MUZIC” was released 5 years ago. When it was released, it attracted many peoples. The reason might be because of its huge promotion, excellent user interface, and soon. But gradually after 3 years, this app noticed that it was losing its subscribers.

The organization identifies that this problem is because of its bad music recommendations to its subscribers.

**Define the objective/goals for this business (e.g. provide decision making/related questions that you would work with through this project)**

Now the goal here is to analyze the song’s data and then identify what’s wrong with it and help the organization to get more subscribers.

**Explain data that will be collected in this business**

I have found a sample data set from Kaggle and used that for my analysis, but for the data collection I will take the data of the songs the user is listening.

Here I analyzed the songs based upon popularity, dance, energy, and bpm. Following the screenshot of the dashboard and the sheets

Chart

Description automatically generated

Link for the dashboard: <https://public.tableau.com/app/profile/balavikas.sabbineni/viz/projectfinaldashboard_16662050638170/Dashboard1?publish=yes>

Links for the sheets:

Dance:<https://public.tableau.com/app/profile/balavikas.sabbineni/viz/projectdance/songsbydance>

Bpm:<https://public.tableau.com/app/profile/balavikas.sabbineni/viz/projectbpm/songsbybpm>

Energy:<https://public.tableau.com/app/profile/balavikas.sabbineni/viz/projectenergy/songsbyenergy>

Popularity:<https://public.tableau.com/app/profile/balavikas.sabbineni/viz/projectpop/songsbypopularity>

Chart, scatter chart

Description automatically generatedA picture containing bar chart

Description automatically generatedGraphical user interface, application, table, Excel

Description automatically generatedGraphical user interface

Description automatically generated

**Identify the decision that needs to be made and select model(s) to help with decision making.**

The decision here we have to make is the list of songs we can recommend the user and if the user likes it then he is going to keep his subscription. If not he is going to change the platform because there is lot of competition going on in the market. **Explain how each model we learn in the class can be embedded in your project.**

**Association Analysis:**

Association analysis helps us to understand identifying the patterns, relations in existing data sets. Considering the data set, we need to figure out the relationship between each song the user is listening

To find the patterns or relations in existing songs we must define Rule. Let us consider a sample

|  |  |
| --- | --- |
| **Transaction** | **songs** |
| 1 | beliver |
| 2 | beliver, one kiss, baby |
| 3 | beliver, on the floor, burn it down |
| 4 | beliver, baby, bad romance, perfect strangers |
| 5 | Beliver, broken, friends, the monster |

A Rule can be derived from here by the user who listen {Beliver} will also listen {one kiss,baby} because based on the transaction dataset, it shows that most of the customers who listen Beliver will also listen one kiss and baby.

{one kiss, baby} 🡪 {beliver}

The strength of this rule can be measured by using the Support and Confidence. **Support** tells that for a given rule how many transactions exist in the dataset that supports the given rule. **Confidence** tells us how often the rule is found to be true.

**Text Mining:**

Unstructured text data can be transformed into useful information through a technique called text mining. In the modern corporate environment, businesses deal with a significant volume of semi-structured and unstructured textual data on a regular basis. The majority of data that a business has is not organized, making it challenging to analyse this data to obtain the meaningful and needed information. In certain business situations, text mining aids in revealing hidden value in the data.

The concept of text mining is applicable to surveys. Different kinds of unstructured information regarding products, retail experiences, and customer experiences are collected by the survey. By employing text mining to analyze this data, it is possible to discover the users' hidden realities.

We can do text mining for the music app, for that we need a different data set. Like a particular user listen history. So that we can extract which song a user particularly listened for how many number of times.

After that we still need to do k-means clustering to get the good recommendations.

**Social Networking Analysis:**

SNA is used to examine patterns of interactions between individuals inside groups. SNA aids in marketing by assisting in determining how to reach customers, who to approach who has more control over others, and who has closer access to others.

Degree centrality can be utilized in networking to find the most-liked people who can easily connect with the larger community or organizations. The concept of betweenness centrality can be utilized to identify people with influence on others. Finding the people who are in a closer position to immediately impact the entire group is known as closeness centrality.

Since they exert more influence than other individuals, we use the SNA to find a "person" with more degree centrality, greater betweenness centrality, and lower closeness centrality for marketing a product. Although hiring such a person comes at a massive cost, he or she can have a stronger impact on the product, increasing sales and profits.

We can’t apply this to our model, but we can use this algorithm for marketing of this app.

**Neural networks:**

A flexible model used to determine the impact of any input on an output is the neural network. When compared to alternative methods that combine numerous models into a single model in a hidden network, neural networks offer superior reliability estimations.

The Rsquare values for any sort of input data set are extremely accurate and show which input variables have the most influence on the output.

The neural network analysis can be applied to our app. Considering the songs data as input, calculating the values of RSquare that is, the likelihood of user listening a song ‘A’ will also listen song B, both the songs should be recommended to the user.

**Decision Tree & logistic regression:**

Decision tree is used to show the effect of categorical input on binary output. For our app, based on the dataset, the categorical inputs may be bpm, energy, popularity etc. The binary output may be user liking that song or not.

But this won’t be helpful for the music recommendation because we can only check if a particular user like a particular song or not by passing some group of songs. In our case we need a group of songs that we could recommend.

The above reason will apply for logistic regression as well.

But we can use these two models to predict the avg values for bpm, energy, and popularity(like we can set the random avg values for bpm, energy, and popularity and checks whether the calculations are reliable or not by comparing validation and training sets(in terms of reliability decision trees are the good option)).

Based on the avg values we can use k-means clustering to get a group of songs

**Finalize the models that you prefer to use for your final project, explain why you chose those models and the information provided by the model(s) that will help with decision making.**

Out of all the algorithms, I think k-means clustering is the best fit

I did k-means clustering on data set Kaggle. So it has 1000 rows in the data and based upon the formula I made it into 15 clusters. So I have 15 heterogenous groups with homogenous songs in each group.

Suppose consider a subscriber is listening to a song called “believer” (58TH row) multiple time so its in the cluster group. So at his recommendation we can keep the below 137 list of songs, and if he isn’t playing any song we can automatically add the songs which have closest similarities in the cluster (according to bpm, energy, dance, popularity).

So here we are not giving subscribers a chance to select the songs, here they are not using any efforts for selecting songs, so the simple thing they has to do is to listen the songs.

With the help of k-means we play the songs he needs like the similar songs

And we recommend song only when the subscribers are out of songs from there queue list and the playlist. If we recommend them at the middle of their playlist or queue list then they don’t get the songs that they expect.

**Graphical user interface, application, table

Description automatically generatedGraphical user interface, application, table

Description automatically generatedGraphical user interface, application, table

Description automatically generatedGraphical user interface, application, table

Description automatically generated**