

MUSIC STREAMING APP

NUJELLA VIJAYA LAKSHMI, K.RAMBABU

1.PG STUDENT, D.N.R. COLLEGE, P.G. COURSES (AUTONOMOUS), BHIMAVARAM-534202.

E Mail Id:nujellavijayalakshmi@gmail.com

1. Assistant Professor in DEPARTMENT OF MASTER OF COMPUTER SCIENCE, BHIMAVARAM-534202.

E Mail Id:kattarambabudnr@gmail.com

ABSTRACT

Recommendation systems have emerged as a result of the large amount of data available on the Internet. Many firms, such as Amazon and Flipkart for e-commerce, wynk music and ganna.com for music streaming, are now employing recommender systems to their advantage. We provide a framework in this particular situation that can then recommend new melodies to clients based on their preferences. This initiative primarily focuses on providing music recommendations to music fans in order to assist them in listening to tracks that they may enjoy. Clients can use this framework to identify new collections of tunes, making the melodic list available for tuning in.

Music is life for music fans, and it has become a larger part of everyone's lives. Music helps us tune in to the cosmos, and the best part about music is that nothing can soothe you like a soothing melody. We chose to do this project because of all the positive aspects of music and the increasing demand for recommender systems on the market. The report comprises a topic description, and a full summary of the work completed thus far. The paper includes thorough explanations of the work completed, including snapshots of implementations, various techniques, and tools used thus far. The project schedule and deliverables are also included in the report. The major goal of music recommendation in this study is to provide strong human-computer interaction and deliver good recommendations to users. It is fluid and can be changed by variables other than the listening history of users or songs

1. INTRODUCTION

The first suggestion system was created in 1979. Elaine Rich defined her Grundy library system [1] as follows: it is used to offer books to users after a brief interview in which the user is requested to fill in his first and last name, and then Grundy asks them to define themselves in a few key terms in order to discover their preferences and classify them as a "stereotype." Grundy provides an initial suggestion by providing a summary of the book after the data has been recorded. If the user is unhappy with the option, Grundy asks questions to figure out which part of the book it made a mistake on and then proposes a fresh one.

Recommendation systems, which first appeared in the 1990s, have advanced significantly in recent years, particularly with the introduction of Machine Learning and networks. On the one hand, the expanding use of today's digital world, which is characterized by a wealth of data, has enabled us to collect massive user databases. On the other hand, when computing power increased, it became possible to handle these data, particularly using Machine Learning, when human skills were no longer capable of conducting a thorough examination of such a large amount of data.

Unlike search engines, which get queries with specific information about what the user wants, a recommendation system does not receive a direct request from the user, but instead must provide them fresh options based on their past behaviors. E-commerce sites that want to sell as many commodities or services as possible to customers (travel,

books, etc.) must swiftly recommend appropriate commodities. The purpose of services that provide streaming music and movies is to keep people on their platform for as long as possible. The recurring theme is that appropriate recommendations are required. Recent advancements in this industry have been significant, and these tips are advantageous to both businesses looking to maximize earnings and customers who are no longer overwhelmed by the quantity of options available. Making decisions is therefore made simple, and a good tip saves a lot of time.

The Recommender System is a software application and algorithm that provides suggestions for items that a user is most interested in. Recommendations are used in a variety of real-world situations, such as deciding what products to buy, listening to music, or reading the latest news. On the other side, there has been a shift in recorded commodity music, particularly after Apple acquired Beats Music in 2014 [8]. The music industry's economic model has recently shifted from commodity sales to subscriptions and streaming. In comparison to prior eras, the availability of digital music is now abundant due to the new business model in the music industry. As a result, the importance of a music recommender system for music suppliers cannot be overstated. It is foreseeing. Collaborative filtering makes suggestions based on the collaborative power of the available evaluation by users.

It is assumed that if people rate music things similarly or behave similarly, they would rate other music items similarly as well. The sparse evaluation matrix is the major issue in collaborative filtering methods since most users only see a tiny portion of all music libraries, hence most assessments are not decided. Content-based filtering, on the other hand, makes suggestions based on the characteristics of the music pieces.

We will see if we can get better recommendations by using real-time data, such as a user's heart rate and the time of day, when making recommendations in this project. The recommendations will be made by a system that employs several machine learning techniques and is accessible via a mobile application. The system uses a smart watch to recognise the user's heart rate in order to give recommendations of songs according to what kind of music is usually associated with that heart rate and time of day for that specific user.

For instance, if a user is out running, the user's heart rate is probably higher than normal. Many music firms, such as Amazon Music, Wynk Music, and Gaana.com, now use recommender algorithms, and the old technique of selling music has shifted to a cloud-based one. All of their music resources are now available in the cloud, and customers may listen to tracks directly from there. However, the problem is that the cloud system has a large amount of music. As a result, we must categorize all of the songs based on various genres, artists' regions, age groups, and languages, with the primary purpose of categorizing these songs according to the user's preferences. Because users demand a good return on their time and money, we can attract a large number of clients by offering a variety of valuable services that they are interested in. We're using a variety of machine learning methods as well as data mining techniques for this project. We tested a number of algorithms and compared the results to determine the most effective algorithm for our model.

2. LITERATURE SURVEY AND RELATED WORK

A We were awestruck with Spotify's recommendation engine. We always wondered how Spotify manages to recommend that perfect song, playlist or even that 'daily mix'. We now have more technology than ever before to ensure that if you're the smallest, strangest musician in the world, doing something that only 20

people in the world will dig, we can now find those 20 people and connect the dots between the artist and listeners. This has been the motivation for this project to use various machine learning techniques and to develop a music recommendation engine similar to that of Spotify, which takes music listening experience to another level. Music Recommendation Systems.

Recommender systems help consumers deal with the problem of information overload by providing them with individualised, unique content and service suggestions. Various methods for developing recommendation systems have recently been created, including collaborative filtering, content-based filtering, and hybrid filtering. The collaborative filtering approach is the most developed and widely used. Collaborative filtering suggests things by locating other users who have similar tastes to the current user and using their recommendations. Collaborative recommender systems have been used in a variety of settings. The common characteristics in these systems are constant when using users' preferences compared with users' context (location, mood, weather, etc.). For instance, in the library when people are sitting there maybe they need quiet and melodious music to listen according to the environment where they are in. Last.fm, All music, Spotify, Pandora and Shazam are commercial music recommendation systems which are considered to be excellent systems by focusing on the music already played in order to help the users to find more music. Users are able to connect to a web-based music streaming service to access the recommendations. All the tracks that are played on this stream are recommended.

It is Based on songs or artists which users either upload from your iTunes playlists or add as favourites on the site where users start managing their library of music with tags and keep tracking of the music the friends who listening to and getting multiple recommendations per song played. Additionally, this app filters recommendations by decade, genre, and popularity, as well as builds fabulous playlists (Song et al., 2012). It has been found that CF generally gives better recommendations than CB. However, this is only true if there is usage data available, such as the ratings given to previous tracks. If this is not the case, then it will not prove accurate results and, consequently, suffer from the Cold-Start problem, which includes two categories of problems – new items and new users. The first problem refers to the new items that are meant to be recommended, but the information that is associated with them,.

Meanwhile, many researchers have used social media (Twitter & Facebook) to identify user's mood (tension, depression, anger, vigor, fatigue, confusion) and also identify user's personality (openness, conscientiousness, extraversion, agreeableness, neuroticism) where these are very important factors which influence on user's music taste and also contextual features (location & event) can lead to different emotional effects due to objective features of the situation or subjective perceptions of the listeners (Scherer et al., 2001).

Music lyrics are also considered to be one of emotional presentation because they include some kinds of implicit thinking, thus we can fully understand emotions and their associated thinking in each song (Nunes and Jannach, 2017; Tintarev and Masthoff, 2008). Cano et al. (2017) mentioned that there is a strong relation between the user mood and listening to the music. The people may want to listen to music which has the same mood of them when they are in specific mood and in contrast the people want to listen to different kind of music which encourage them to enhance their mood and this thing depend on the psychological studies and therefore, the author produced a contextual mood-based music recommender system which is able to regulate the driver's mood and also try to put the driver in a positive mood when driving because listening to the music while driving has always been one of the most favourite activities carried out by people. Finally,

similarly, active learning approaches suffer from various limitations.

3. EXISTING SYSTEM

Music recommendation systems use various techniques to suggest songs or playlists to users based on their preferences and behavior. Here's an overview of an existing music recommendation system:

Data Collection: These systems collect vast amounts of data, including user listening history, user profiles, song metadata (genre, artist, album, release date), and user interactions (likes, skips, playlist creations).

User Profiling: User profiles are created by analyzing their listening habits, favorite genres, artists, and historical data. Machine learning algorithms are often used to cluster users with similar preferences.

Content-Based Filtering: This approach recommends music similar to what the user has previously liked. It analyzes song features like tempo, key, and genre to find similarities.

Collaborative Filtering: Collaborative filtering techniques use the behavior of users to make recommendations. There are two types:

User-Based: Recommends songs that users with similar listening histories have liked.

Item-Based: Recommends songs that are similar to songs the user has liked in the past.

Matrix Factorization: This technique decomposes the user-item interaction matrix to identify latent factors, which can be used to make personalized recommendations.

Deep Learning: Neural networks, particularly deep learning models like recurrent neural networks (RNNs) and convolutional neural networks (CNNs), are used for sequential and content-based recommendations. Recurrent models can capture sequential patterns in listening behavior.

Online Learning: These systems continuously update recommendations based on real-time user interactions, ensuring that recommendations stay current.

Evaluation: Systems use metrics like Mean Average Precision (MAP), Normalized Discounted Cumulative Gain (NDCG), and A/B testing to evaluate recommendation quality.

Privacy and Ethics: Ensuring user data privacy and addressing ethical concerns like filter bubbles and bias is a critical aspect of modern recommendation systems.

Hybrid Approaches: Many systems combine multiple recommendation techniques (e.g., content-based and collaborative filtering) to provide more accurate and diverse recommendations.

User Interface: Recommendations are presented to users through interfaces like playlists, personalized radio stations, or "Recommended for You" sections.

Companies like Spotify, Apple Music, and YouTube Music employ advanced recommendation systems to enhance user engagement and satisfaction. These systems continuously evolve through research and development to improve accuracy and user experience.

4. PROPOSED SYSTEM

A proposed music recommendation system could utilize a combination of collaborative filtering, content-based filtering, and machine learning techniques to provide personalized music recommendations to users. Here's an outline of such a system:

Data Collection:

Gather user data, including listening history, user profiles, and demographic information. Collect music data, including song metadata, genre, artist information, and user-generated content like reviews and ratings.

Data Preprocessing:

Clean and preprocess the data, handling missing values and outliers. Feature extraction from song metadata, such as tempo, mood, and key.

Collaborative Filtering:

Implement collaborative filtering algorithms like User-Based and Item-Based Collaborative Filtering to find similarities between users and songs. Use matrix factorization techniques like Singular Value Decomposition (SVD) or matrix factorization for recommendations.

Machine Learning Models:

Train machine learning models, such as neural networks or decision trees, to predict user preferences based on historical data. Incorporate hybrid models that combine collaborative and content-based approaches for better accuracy.

Real-time Updates:

Continuously update user profiles and recommendations based on their recent interactions and preferences. Implement real-time learning to adapt to changing user tastes.

Evaluation:

Use metrics like Mean Average Precision (MAP), Root Mean Square Error (RMSE), or Click-Through Rate (CTR) to evaluate the system's performance.

Conduct A/B testing to compare the proposed system against existing recommendation algorithms.

User Interface:

Develop a user-friendly interface, such as a mobile app or web platform, where users can easily access and interact with the recommendation system.

Allow users to provide feedback on recommendations to improve future suggestions.

Personalization:

Implement user-specific recommendation algorithms to provide personalized playlists, discover new music, and cater to diverse tastes.

Privacy and Security:

Ensure user data privacy by anonymizing and securely storing user information.

Comply with data protection regulations and obtain user consent for data usage.

Deployment:

Deploy the music recommendation system on scalable infrastructure to handle a large user base. Monitor system performance, scalability, and user engagement.

5. METHODOLOGIES

MODULES

Date Set Used in the Major Project

Spotify Dataset 1922-2021, ~600k Tracks Audio features of ~600k songs released in between 1922 and 2021. this dataset contains audio features and metadata of each song.

This dataset then went through different data mining techniques to make it suitable for further analysis .In simple terms data mining means firstly extracting the useful data and then making that data suitable to be used in the algorithms by cleaning the data and transforming it. This dataset then went through different data mining techniques to make it suitable for further analysis This dataset then went through different data mining techniques to make it suitable for further analysis .

source: <https://www.kaggle.com/yamaerenay/spotify-dataset-19212020-160k-tracks>

limitations :

- 1.only genre metadata was provided
- 2.Since, the dataset consists of only audio features and no data related to the user's listening history , therefore we cannot perform collaborative filtering techniques .
- 3.kaggle dataset doesn't consist of all of the songs in my playlist .

This dataset then went through different data mining techniques to make it suitable for further analysis .

Data Set Features

Types of Data Set

The dataset is in the form of csv files which have been taken from kaggle and then various data mining techniques are applied to it to extract the information.Dataset had various attributes like user id, song counts, language, artist, genre, and year. The dataset was further divided in four parts namely data by genre, data by artist, data by year and data and were stored in individual csv files to analyze and to train the model.

6. RESULTS AND DISCUSSION SCREEN SHOTS

1 Discussion on the Results Achieved

This assignment provided us with a fantastic learning opportunity. We've studied data mining and data cleansing.

The music supplier can forecast and then give acceptable songs to its consumers using a music recommender system based on the qualities of previously heard music.

for each song listened to by the user, the average vector of audio and metadata attributes Find the n-closest data points in the dataset (excluding points from the user's listening history) to this average vector.

Take these n points and come up with some tunes to go with them.

A research on the limits of an interactive music recommendation service based on artificial audio similarity calculation was provided.

A number of computer experiments, as well as a review of real download data, reveal that a large chunk of the audio collection is only never or never suggested.

A number of computer experiments are used to investigate this issue in depth, including the investigation of various audio similarity functions and comparisons with real download data.

Our music recommendation service uses Gaussian mixtures as statistical models to determine timbre similarity.

This is the de facto standard method for computing audiosimilarity, and it is recognised to produce high-quality results.

A machine learning model's first goal is to eliminate all problem-causing objects from the dataset. Data cleansing and exploration were quite beneficial in getting the dataset algorithm ready.

We learned how to design a machine learning model, train it, and then test it.

- 1) The songs which scored highest have been recommended in the result given below.

- 2) Smaller the angle, the higher the song score.

edm_top40														
	acousticness	artists	danceability	duration_ms	energy	explicit	id	instrumentalness	key	liveness	loudness	mode	name	popularity
10730	0.06430	['Valerie Broussard', 'Galantis']	0.683	184564	0.785	0	23FH9lYnG6Dr8OzombPkS	0.000013	7	0.1770	-4.879	1	Roots	
135509	0.01600	['Calvin Harris', 'Rag'n'Bone Man']	0.807	229184	0.887	0	5itOtNx0WxtJmi1TQ3RuRd	0.000503	1	0.0811	-4.311	0	Giant (with Rag'n'Bone Man)	
70513	0.08100	['Loud Luxury', 'Bryce Vine']	0.875	187797	0.858	0	7fcEMgPlcjD0LzPHwMsoic	0.000001	4	0.3810	-3.886	1	I'm Not Alright	
106357	0.02820	['Galantis']	0.674	191293	0.915	0	6M6Tk58pQvABY6ru66dY3d	0.003370	6	0.2730	-3.999	0	No Money	
106353	0.11700	['Galantis', 'Throttle']	0.762	190400	0.797	0	5kgqTe1BM720OJU78TGYDw	0.000000	5	0.2020	-2.710	1	Tell Me You Love Me	
100074	0.22900	['Gryffin', 'Katie Pearlman']	0.590	231291	0.764	0	17ejRbr688l9zdqgCZsn4m	0.000000	2	0.1920	-4.735	1	Nobody Compares To You (feat. Katie Pearlman)	

Application of the Major Project

In a later version, the goal is that the application will also be able to record an extract of a music being played. The music supplier can forecast and then give acceptable songs to its consumers using a music recommender system based on the qualities of previously heard music. For each song listened to by the user, the average vector of audio and metadata attributes. Find the n-closest data points in the dataset (excluding points from the user's listening history) to this average vector. Take these n points and come up with some tunes to go with them.

The music supplier can forecast and then give acceptable songs to its consumers using a music recommender system based on the qualities of previously heard music.

From the music or extract, the application will offer the possibility to listen to recommended songs by the algorithm developed in project.

The programme will eventually be able to record an excerpt of music being played. Using a music recommender system based on the attributes of previously heard music, the music provider may foresee and then provide suitable songs to its customers.

The application will allow users to listen to songs selected by the algorithm built in the project based on the music or extract. The music supplier can forecast and then give acceptable songs to its consumers using a music recommender system based on the qualities of previously heard music. For each song listened to by the user, the average vector of audio and metadata attributes. Find the n-

closest data points in the dataset (excluding points from the user's listening history) to this average vector. Take these n points and come up with some tunes to go with them.

For user convenience and technological simplicity, a dataset can be generated and versioned fully from a single data source. That is, data sources in a dataset cannot be mixed and matched at this time. The programme will allow users to listen to recommended tracks based on the music or extract.

Limitation of the Major Project

This project due to the nature of the dataset fails to provide accurate recommendations as the dataset does not consist of all of the songs in the playlist. It is noteworthy that a dataset can be built and versioned entirely from one data source for user convenience and technical simplicity. That is, data sources in a dataset cannot presently be mixed and matched (for example, a dataset built from a GitHub repository cannot include files uploaded from your local system). You can build numerous datasets and add them both to a Notebook if you want to leverage multiple distinct data sources in it.

Songs that are identical to a large number of other songs and hence appear unnecessarily frequently in recommendation lists prevent a big section of the

audio library from being recommended at all. A number of computer experiments are used to investigate this issue in depth, including the investigation of various audio similarity functions and comparisons with real download data.

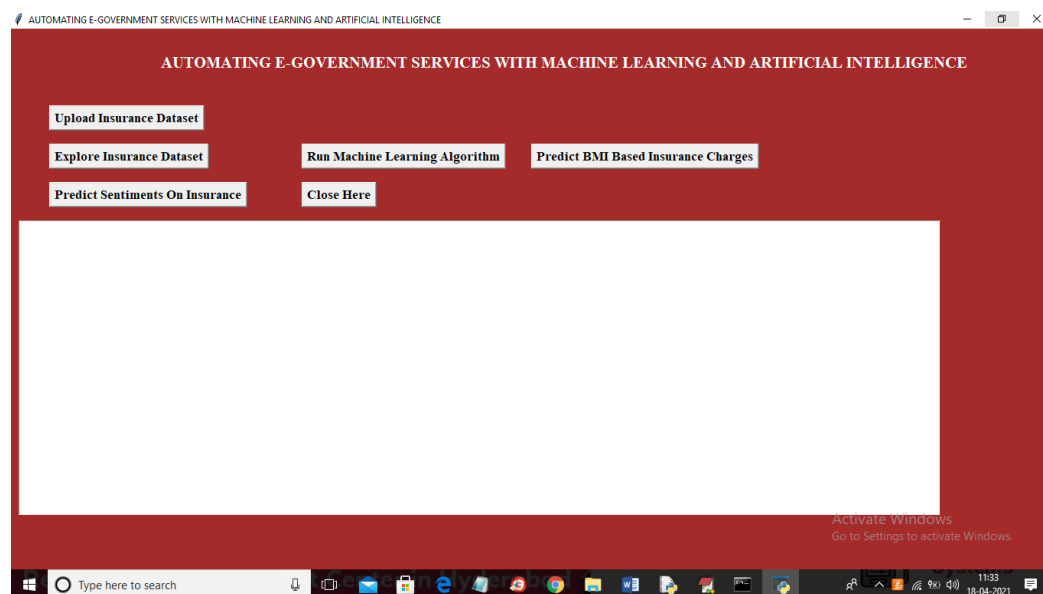
For user convenience and technological simplicity, a dataset can be generated and versioned fully from a single data source. That is, data sources in a dataset cannot be mixed and matched at this time

The programme will allow users to listen to recommended tracks based on the music or extract.

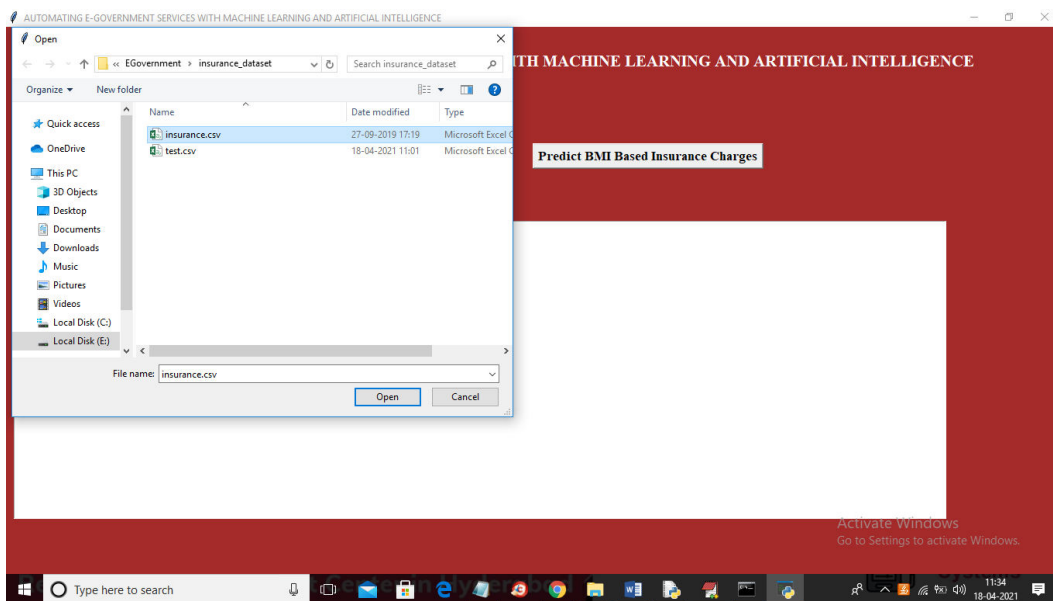
A dataset can be built and versioned entirely from one data source for user experience and technical simplicity. The music supplier can forecast and then give acceptable songs to its consumers using a music recommender system based on the qualities of previously heard music. For each song listened to by the user, the average vector of audio and metadata attributes is found. The n -closest data points in the dataset (excluding points from the user's listening history) to this average vector are found. Take these n points and come up with some tunes to go with them. That is, data sources in a dataset cannot presently be mixed and matched (for example, a dataset built from a GitHub repository cannot also include files uploaded from your local workstation). You may build numerous datasets and add them both to a Notebook if you want to leverage different data sources in it.

Screens

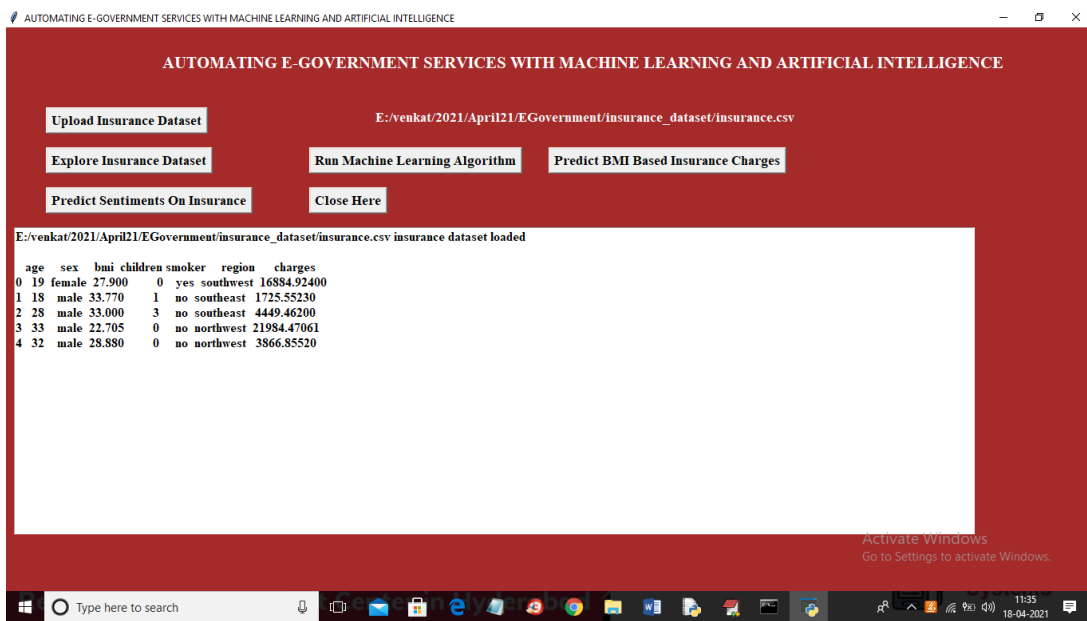
To run project double click on 'run.bat' file to get below screen



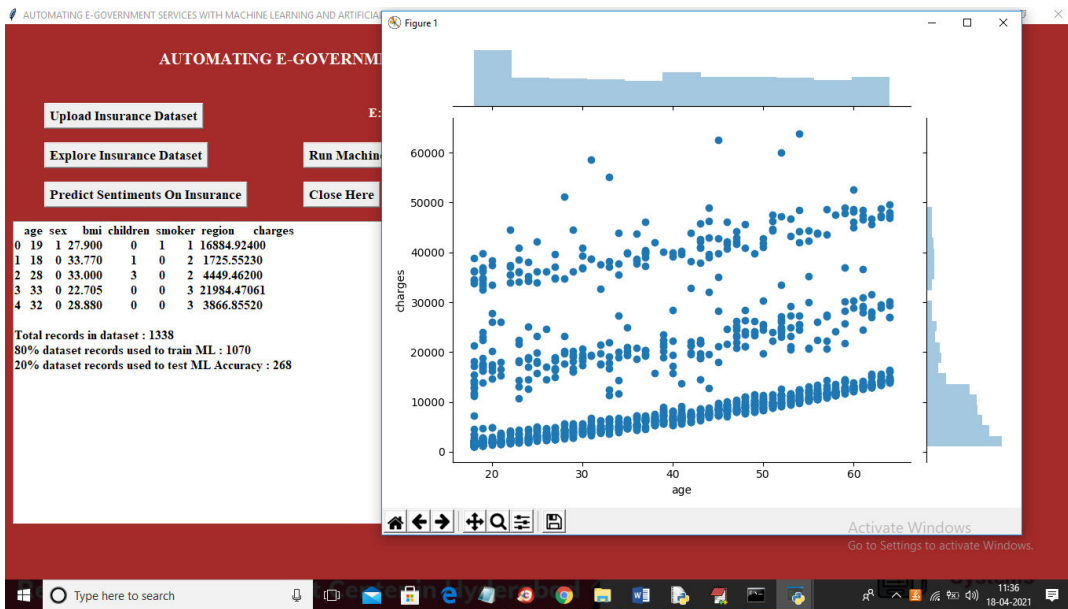
In above screen click on 'Upload Insurance Dataset' button and upload dataset



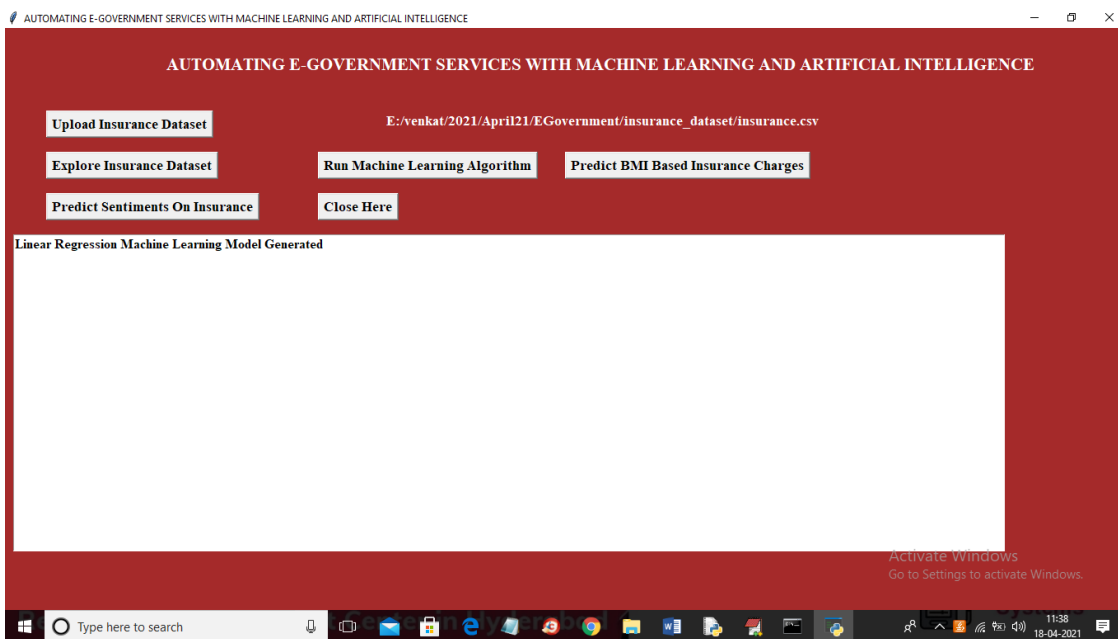
In above screen selecting and uploading “insurance.csv” file and the click on ‘Open’ button to load dataset and to get below screen



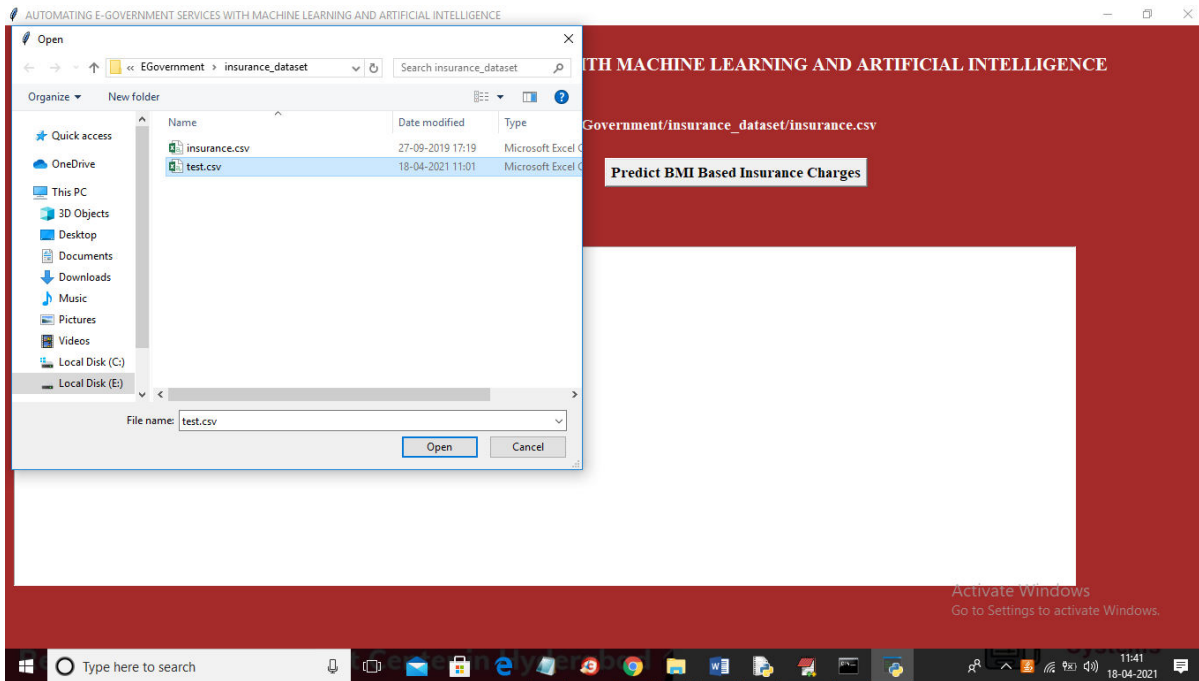
In above screen dataset loaded and we can see dataset contains some non-numeric values and ML will not take string value so we need to convert non-numeric string values to numeric by replacing male with 0 and female with 1. So, click on ‘Explore Insurance Dataset’ button to replace string with numeric values.



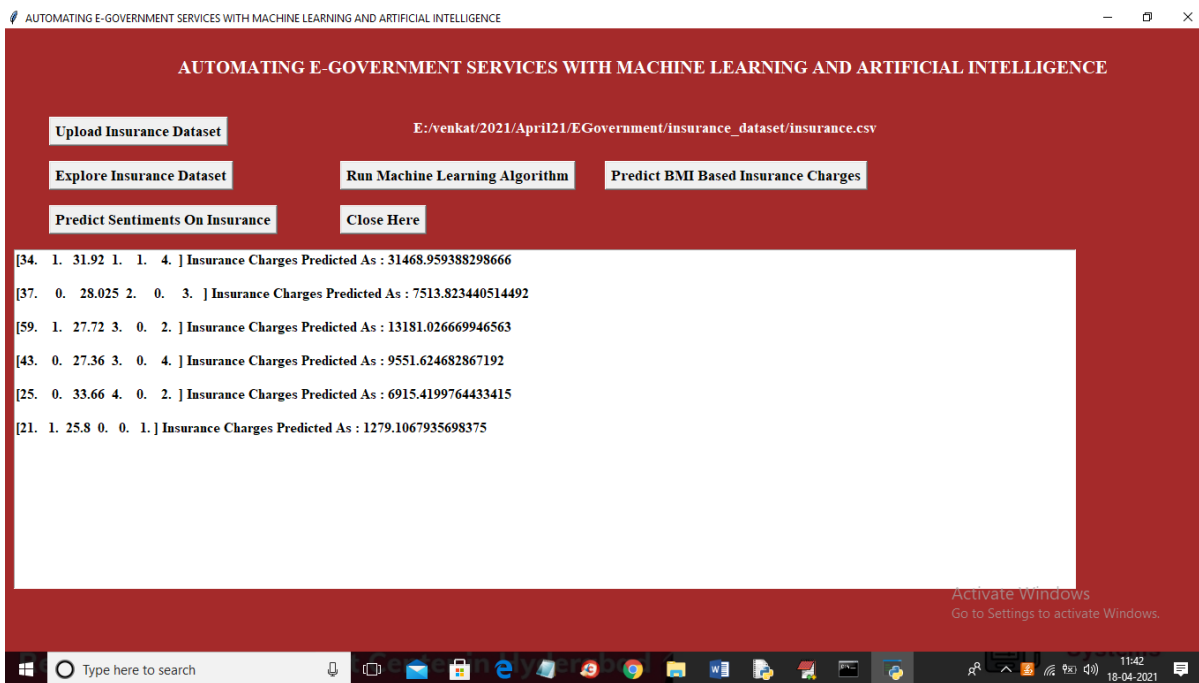
In above screen we can see all string values replace with numeric data and we can see dataset contains total 1338 records and application using 1070 records to train ML and 268 records to test ML performance and in above graph x-axis represents AGE and y-axis represents insurance charges and we can see in above graph when AGE increasing then insurance charges also increasing. Now dataset is ready and now click on 'Run Machine Learning Algorithm' button to build ML model



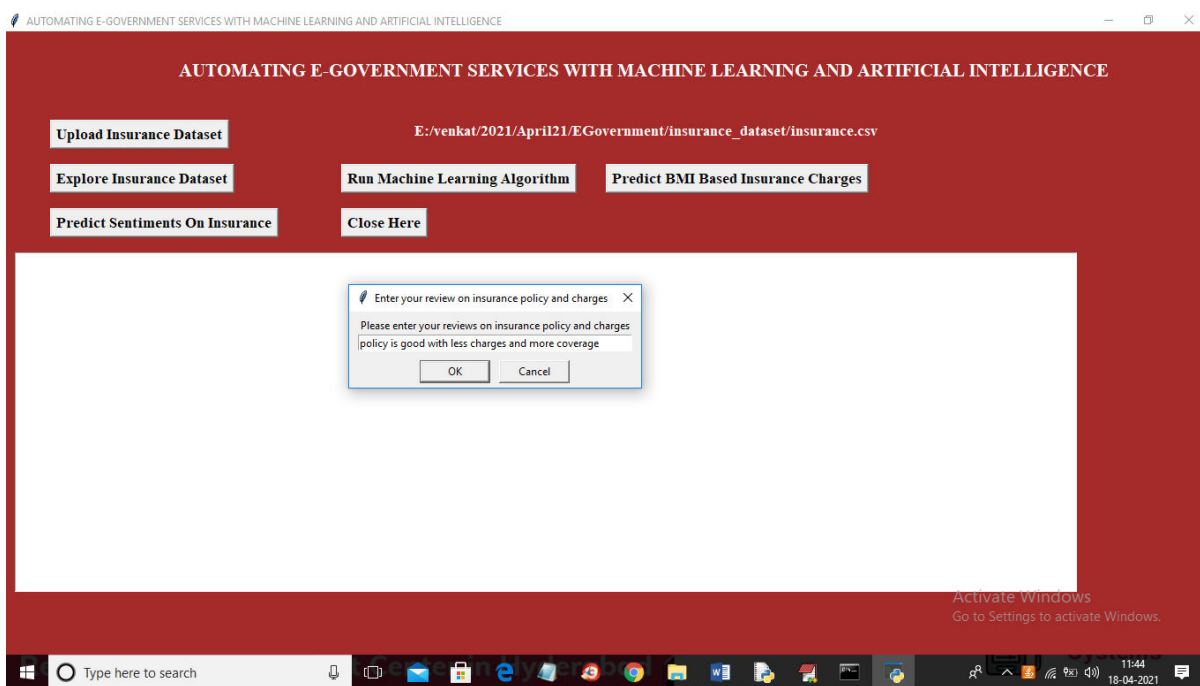
In above screen ML model generated and now click on 'Predict BMI Based Insurance Charges' button to upload test data and then ML will predict insurance policy and charges.



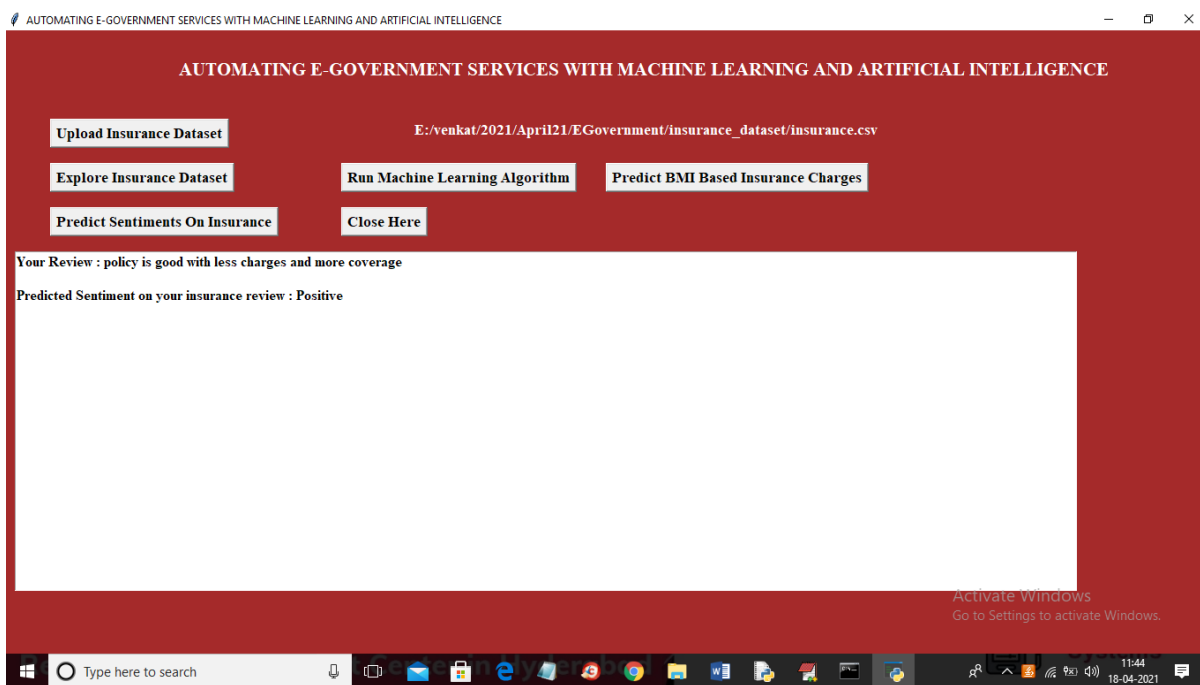
In above screen selecting and uploading 'test.csv' file and then click on 'Open' button to load test records and then ML and AI will analyze gender, smoking, BMI and then predict policy charges and will give below output



In above screen in square, bracket we can see test values and after square bracket we can see predicted policy charges. After seeing charges and policy user may express some opinion and then ML will predict sentiment from that opinion. So, click on 'Predict Sentiments on Insurance' button to express some reviews



In above screen in dialog, box I entered some review and then click on 'OK' button to get below result



In above screen for given review AI and ML predict sentiment as POSITIVE. Similarly, you can enter any opinion and then ML will predict sentiment.

7. CONCLUSION AND FUTURE SCOPE

The range of characteristics covered by the recommender system is extensive. In today's generation of e-services and commerce, it is growing and evolving. However, there is a requirement to create and optimise the working and output of the recommender system at the same time. For user convenience and technological simplicity, a dataset can be generated and versioned fully from a single data source. That is, data sources in a dataset cannot be mixed and matched at this time (for example, a dataset built from a GitHub repository cannot include files uploaded from your local system).

The programme will allow users to listen to recommended tracks based on the music or extract.

Several service providers provide consumers with a shopping list. However, this is insufficient since consumers have varying preferences and decisions that are influenced by a variety of circumstances and restrictions. It may also be impossible to propose specific things to individual users in many circumstances. As a result, there is potential for combining several dimensions into music recommender systems in particular.

We were unable to create a model utilising singular value decomposition and support vector machines due to a lack of time. Because popularity-based models are adept at making suggestions, we'll aim to utilise it to forecast the top-N songs for the users who are most popular at any given time. Indeed, during years, in order to choose music, restaurants, movies, etc.. We have been asking our friends, family, and colleagues to recommend something they liked. And it is this mechanism that is attempted to be reproduced here. Netflix was a pioneer of this method (based on stars given by other users) but it is now widely used, including for Spotify's Discover Weekly. Collaborative filtering makes suggestions based on the collaborative power of the available evaluation by users.

Customers are less likely to use the majority of the items and services offered by various e-commerce sites since they are pricey. As a result, you won't be able to accurately and properly rank an item or collection of things. As a result, typical recommender system strategies are inadequate. This paves the path for more research and development in the form of an efficient recommender system that also considers constraints. We'll aim to utilise it to forecast the top-N songs for the users who are most popular at any given time.

Discover Weekly is a 30-song playlist that includes music that are similar to what the user is listening to. This, like its daily mixes and tailored playlists, is made possible by AI and big data. The system also considers the user's streaming history and playlists, as well as their current music preferences, to improve this suggestion.

8. REFERENCES

- i. Aioli, F. (2012). A preliminary study on a recommender system for the million songs dataset challenge. PREFERENCE LEARNING: PROBLEMS AND APPLICATIONS IN AI
- ii. Spotify. (2013, March) Spotify Music for everyone. [Online].

- iii. <https://www.spotify.com/us/> [10] (2013, March) Wikipedia. [Online].
- iv. Peter Knees and Markus Schedl. Music retrieval and recommendation – a tutorial overview. In Proceedings of the 38th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), Santiago, Chile, August 2015.
- v. Peter Knees and Markus Schedl. Music Similarity and Retrieval: An Introduction to Audio-and Web-based Strategies. 2016.
- vi. Chris Johnson. From idea to execution: Spotify's discover weekly. November 2015.
- vii. Zhiwei Gu Li Guo and Tianchi Liu. Music genre classification via machine learning.
- viii. [11] Vikas Sindhwani Prem Melville, "Recommender Systems," in Encyclopedia of Machine Learning. Yorktown Heights, USA: IBM T. J. Watson Research Cen-ter, 2010, ch. 00338.
- ix. Beatriz C. F. de Azevedo Glaucia M. Bressan and Elisangela Ap. S. Lizzi.