International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 11 | November - 2022 Impact Factor: 7.185 ISSN: 2582-3930

BIG DATA ANALYTICS IN SPOTIFY

Dr. C.K.Gomathy[1], N. Sri Hasitha[2], G. Sai Sampadaa[3], S. Srilalitha[4]

[1]*Assistant Professor, Department of CSE, SCSVMV Deemed to be University, Kanchipuram*
[2, 3 & 4]*Student, Department of CSE, SCSVMV Deemed to be University, Kanchipuram*

I. ABSTRACT

Millions of customers daily enjoy and use the popular music streaming service Spotify. The app is the biggest on-demand music service in the world and is extremely popular in the music industry. The most popular music streaming software, Spotify, is made possible by cutting-edge technology like Big Data. With a current net value of over $25 billion, Spotify significantly influences the music business by utilizing data in novel ways.

Keywords: Data management system, data analytics, music streaming, data service

II. INTRODUCTION

On April 23, 2006, Daniel Ek and Martin Lorentzon launched Spotify (/sptrfa/), a privately held Swedish audio streaming and media services company. With approximately 433 million monthly active users, including 188 million paid members, as of June 2022, it ranks among the top suppliers of music streaming services. In the form of American depositary receipts, Spotify is traded on the New York Stock Exchange. With Spotify, you may access millions of songs and other works from authors all around the world. Spotify is a digital music, podcast, and video service. More than 82 million tracks from record labels and media

corporations are available on Spotify as digitally copyright-restricted recorded music and podcasts. A freemium service offers certain basic functions for free with commercials and restricted control, while charging for further features like offline listening and commercial-free listening. As of October 2021, Spotify will be accessible in more than 180 countries. Users have the ability to make, edit, and share playlists as well as search for music by artist, album, or genre. Big Data Analytics can be used to enable Spotify to automatically construct playlists based on user preferences and provide a personal touch.

III. IMPORTANCE OF BIG DATA IN SPOTIFY

Big Data Analytics is the reason why Spotify is at the top. People use Spotify due to its sizable playlist and weekly discovery suggestion. The software presents users with various songs and musical genres that appeal to their likes; eventually, this encourages users to utilize the app more frequently. Data Science employs its algorithms to get data-driven insights and improve consumer experiences. It has generated over 600 terabytes of data per day. Additionally, Spotify constantly searches the web for blog entries and other relevant text and music. The app wants users to enjoy music and learn more about the songs and musicians they love. Spotify uses big data to evaluate and produce lists of the top songs, songs that are currently trending, and many other things. It keeps the audience engaged and encourages longer Spotify sessions. After taking that into consideration, Spotify made a choice and continued to inform interested users.

IV. EXISTING SYSTEMS

The current platforms, which include Google Music, JioSaavn, Amazon Music, Gaana, and Apple Music, among others, are among the most popular music apps in the world. These offer HD audio and limitless ad-free music in premium features. The user has the ability to create his own playlist and download music with paid features.

V. PROPOSED METHODOLOGY

Spotify analyses your listening habits and creates personalized recommendations using data mining, artificial intelligence, and mobile machine learning. Even just one day of access to the music can be purchased by the user. This technology is used for the vast majority of global music streaming. It provides excellent mp3 songs.

VI. METHODOLOGY

- Spotify's distinctive and potent recommendation model, which is driven by machine learning, is built using a variety of data aggregation and sorting techniques.

- The collaborative filtering method contrasts different playlists made by users that contain the music that users have listened to. The algorithm then peruses these playlists to hunt for more songs that feature on them and suggests those songs.

- The Natural Vocabulary Processing (NLP) system can categorize music based on the language used to describe them and may compare them to other songs that are discussed in a similar way. Based on the data, classification tasks for songs and artists are given, and each term is given a weighted value.

- For audio files, Spotify has configured the CNN model, which is the most widely used for facial recognition. Each music is transformed into a raw audio waveform file. These waveforms are analyzed by CNN, which assigns them important attributes like the number of beats per minute, the volume, the major or minor key, and so forth.

VI. COLLECTION OF DATA

Spotify uses a variety of scripting languages and software, including Java, HiveHive, Apache Kafka, and others, to gather data. These are the main reasons that Spotify uses big data and big data analytics to produce reliable results. User Interference in the Spotify mobile app has grown to be a critical component of machine learning. Playlists, information, and personal history are three of the main purposes for which the data is used. It aids Spotify in understanding user behavior by collecting data on everyone. They will learn the pattern and details of the music. This allows them to generate a proposal for the users. The history will indicate what they search for, what music they have been playing often, what artist or song they avoid or skip, which albums they like and detest, and much more.

The data supporting each of the aforementioned claims is content-skeptical, but by using neural networks to predict such information just from the audio signal. Later, the algorithms created by data science develop a sense of self and know what to do next and what has to be changed. Additionally, artificial intelligence aids in enhancing Spotify's functionality. It is applying a wide range of collaborative filtering approaches; using this technique, a pattern in the data has been found. Additionally, it offers potential representations of each user as well as each musician or song. As a result, Spotify has access to a wide variety of music. Users are able to listen to music exclusively in particular regions of that plan since related songs are grouped together using Big Data in this area of Spotify.

VII. IMPLEMENTATION

Producing unique material that will only be made available on the Spotify platform is the suggested strategic shift for the company. The cost structure of Spotify will be reduced, and differentiation will rise.

- ***Collaborative Filtering***: Collaborative Filtering is a well-liked method that recommender systems employ to automatically forecast users' preferences based on the preferences of other users who share those tastes.

- ***Natural Language Processing*:** Real-time understanding of speech and text by an algorithm is known as natural language processing, or NLP. To create a profile for each song, Spotify's NLP continuously searches the web for articles, blog posts, and other texts on music.
- ***Convolutional Neural Networks**:* Because less-popular songs could be overlooked by other models, convolutional neural networks are employed to fine-tune the recommendation system and to boost accuracy. The CNN algorithm ensures that both fresh and obscure music are taken into account.

The SWOT analysis:

Strengths: Spotify has many advantages over its competitors, including a strong brand name, nimble organizational structure, enormous user base, advanced algorithms, capacity for innovation, and adaptable financial situation.

Weaknesses: Lack of variety in music material and Spotify's current licensing agreements with music right holders are the service's shortcomings.

Opportunities: Spotify has the chance to set itself apart from the competition by creating original material that will only be made available on the platform. Additionally, Spotify has a chance to attract customers from both established and developing areas in the music streaming sector.

Threats: Spotify faces threats from the influence of music rights holders who may raise the cost of music licensing for streaming and from changes to data collection laws that may limit Spotify's ability to gather user data.

VIII. USE OF BIG DATA IN SPOTIFY

1. Creating Personalized Content:

Utilizing user data to create material that each user will consider to be exclusively tailored to their individual likes is a critical strategy used by Spotify to adopt user data. The aim is to guarantee people have a positive experience so they stay as clients for a long time. Various AI and machine learning algorithms have

been used to achieve this.At the moment, the "Discover Weekly" has become one of Spotify's greatest strengths. Completely created by a machine learning algorithm, it creates a customised playlist that is only relevant to the user's listening behavior

2. User preferences being digitized

The "Daily Mixes" playlists on Spotify also take the listener's daily taste preferences into account. These playlists are distinct from the musical genres that the user typically enjoys and are typically made up of songs that the user has saved or added to their playlists, as well as songs that were produced by artists whose music is already on their current playlists or by new artists or albums that the user is not familiar with.

3. To improve marketing with focused advertisements:

In addition to improving the user experience, Spotify has been able to leverage a sizable portion of the data collected by its users to update their advertising campaigns and more effectively target their audience. This is essentially accomplished by the platform looking at the information they have learned about their listeners and then using those insights to produce advertising that deftly targets the platform's target demographic.

4. Constantly updating its system:

The streaming service had announced in the beginning of 2018 that its free users will no longer be restricted to just shuffle through music on their application. Instead, their users were now free to explore 15 of the platform's well-known playlists, including "RapCaviar" and the platform's well-liked "Discover Weekly. “Data served as the driving force for the platform's decision. The platform can now provide the data of additional 100 million or more users, which is really helpful given that the company concentrating on improving its suggestion algorithms to provide its users with an enjoyable personalized experience. The platform launched the Spotify for Artists application to give access to analytics, such as which playlists have been assisting in generating new users and the overall number of streams they are receiving. This was done in an effort to make their massive amount of data available for their musicians as well as their managers.

IX. RESULTS

Due to the use of big data and the use of machine learning algorithms, Spotify now has 207 million subscribers, the most of any music app at this time. There are almost 40 million musical collections there. It is simple to access on a user's PC. Playlists are socialized through Spotify's sharing features. It's simply one more way that it distinguishes itself as the top music streaming software for iPhone and Android.Overall, Spotify offers a ton of mobile app features and capabilities to offer its users, ensuring that the experience and design are always new and intriguing. Over 100 million people use it each month, and it routinely wins awards for being the finest music streaming app thanks to its breadth, cutting-edge technology, and shareability.

X. CONCLUSION

Spotify stands out as a prominent innovation in the sector of music streaming in the Internet economy, which is undergoing rapid change. Starting out in a crowded market of rivals, Spotify soon advanced to dominate the music streaming sector and establish itself as the go-to platform for millions of customers.Big data and a user-focused business model that valued the attention of millions of users over the monetary earnings that might have been made or the economic value of the music it was giving were used to achieve this. In order to provide its users with a larger selection of music, Spotify was able to become successful and draw more listeners, which grew its network and gave it more power. However, the user base of Spotify exists in a dynamic digital environment where new competitors constantly emerge and technology develops at a rapid pace. Spotify must always improve its services and produce new products in order to stay ahead of the competition and meet the changing demands of their audience. Only because they have gotten better at offering their sizable network updated services and cutting-edge products, as well as focusing their business model on the demands of the customers, have they been able to preserve their position as the leading music streaming service. In a market where businesses can grow quickly, they also risk falling at the mercy of a user network just as soon. And it's clear that businesses like Spotify will only prosper if they keep coming up with fresh solutions to address the audience's shifting wants.

XI. REFERENCES

[1] C.K.Gomathy.(2010),"Cloud Computing: Business Management for Effective Service Oriented Architecture" International Journal of Power Control Signal and Computation (IJPCSC), Volume 1, Issue IV, Oct - Dec 2010, P.No:22-27, ISSN: 0976-268X .

[2] Dr.C K Gomathy, Article: A Study on the recent Advancements in Online Surveying , International Journal of Emerging technologies and Innovative Research (JETIR) Volume 5 | Issue 11 | ISSN : 2349-5162, P.No:327-331, Nov-2018

[3] Dr.C K Gomathy, Article: A Study on the Effect of Digital Literacy and information Management, IAETSD Journal For Advanced Research In Applied Sciences, Volume 7 Issue 3, P.No-51-57, ISSN NO: 2279-543X,Mar/2018

[4] Dr.C K Gomathy, Article: An Effective Innovation Technology In Enhancing Teaching And Learning Of Knowledge Using Ict Methods, International Journal Of Contemporary Research In Computer Science And Technology (Ijcrcst) *E*-Issn: 2395-5325 Volume3, Issue 4,P.No-10-13, April ’2017

[5] Dr.C K Gomathy, Article: Supply chain-Impact of importance and Technology in Software Release Management, International Journal of Scientific Research in Computer Science Engineering and Information Technology (IJSRCSEIT) Volume 3 | Issue 6 | ISSN : 2456-3307, P.No:1-4, July-2018.

[6] Dr.C K Gomathy, ANALYSIS OF MUSIC GENRE CLASSIFICATION, Journal Of Engineering, Computing & Architecture, Volume: 12 Issue: 03 March – 2022, Impact Factor:6.1, ISSN:1934-7197, Available at http://www.journaleca.com/

[7] Dr.C K Gomathy, THE EFFICIENT MANAGEMENT FOR MALICIOUS USER DETECTION IN BIG DATA,COLLECTION., International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 05 Issue: 10 | Oct- 2021, ISSN: 2582-3930, Impact Factor:7.185,Available at www.ijsrem.com

[8] Dr.C K Gomathy, BIG DATA ANALYSIS OF AIRLINE DATASET USING HIVE, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 10 | Oct 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072, Impact Factor value: 7.529, Available at www.irjet.net.

[9] Dr.C K Gomathy, THE CUSTOMER DATA ANALYSIS USING SEGMENTATION WITH SPECIAL Reference Mall, International Research Journal of Engineering and Technology (IRJET), Volume: 08 Issue: 9 | Sep 2021, e-ISSN: 2395-0056, p-ISSN: 2395-0072, Impact Factor value: 7.529, Available at www.irjet.net.

[10] Dr.C K Gomathy, THE EFFICIENT MANAGEMENT FOR MALICIOUS USER DETECTION IN BIG DATA,COLLECTION., International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 05 Issue: 10 | Oct- 2021, ISSN: 2582-3930, Impact Factor:7.185,Available at www.ijsrem.com

Author's Profile:

- Ms.N. Sri Hasitha, Student ,B.E.Computer Science and Engineering, Sri Chandrasekharendra Saraswati Viswa MahaVidyalaya , Enathur , Kanchipuram , Tamilnadu , India. Her Area of Interest : Data Science and Data Analytics.

- Ms.G. Sai Sampadaa, Student, B.E.Computer Science and Engineering, Sri Chandrasekharendra Saraswati Viswa MahaVidyalaya, Enathur , Kanchipuram , Tamilnadu , India. Her Area of Interest : Data Science and Data Analytics.

- Ms.S. Srilalitha, Student, B.E.Computer Science and Engineering, Sri Chandrasekharendra Saraswati Viswa MahaVidyalaya , Enathur , Kanchipuram , Tamilnadu , India. Her Area of Interest : Data Science and Data Analytics..

- Dr.C.K.Gomathy is Assistant Professor in Computer Science and Engineering at Sri Chandrasekharendra Saraswathi Viswa MahaVidyalaya, Enathur, Kanchipuram,India. Her area of interest in Software Engineering, Web Services, Knowledge Management and IOT.