

Report

(Research on data analytics process and streaming services)

➤ Introduction

Data analytics is widely used in streaming services. Streaming services are online entertainment providers (music and videos). Example : Netflix, Spotify, Amazon prime, You tube etc. When we watching a video or listening a song on such a online platform, we are getting lots of recommendation related that song or video. The question is that how these recommendations are comes after watching that particular song or video? How streaming services know that, which recommendation they have to show to the viewers up next?

These recommendation personalise good experience to the viewers and also provides to the viewers area of interest. These streaming services ensure that viewers keep using their services long period of time.

So in this report we are going to see research made on two streaming services(music and video),through these research we have to find what type of user data is collected by streaming services and what are the various objective of collecting this data.

After that how the data analytics is helpful in music streaming services and how these streaming services comes up with bang recommendation? Also we are investigating the data analysis tools and methods used by streaming services.

At last the insights that we have founded during data analysis we have to visualize this data in data visualization tool.

• Data analytics

Data analytics is collecting, processing the raw data finding insight from it for making business decisions.

• Process of data analytics

- **Data collection** : Data collection is gathering the data from various sources as per your requirement.
- **Data cleaning** : data cleaning is removing white spaces, duplicate records, and basis errors. when you collect the data, you will find that the data is not in structured format, you have to make this data in structured format. Sometimes all collected data is not usefull for your analysis, you have to collect important data from it.
- **Data analysis** : We can analyse data using data analysis software. Data analysis is interpreting, understanding the data and come to conclusion using data analysis software.
- **Data interpretation** : You have a results from the data. Now you have to interpret them and making business decision based on your findings.
- **Data visualization** : data visualization is showing your insights in the form of charts and graph. Data visualization helps to understand data in very simple way.

• Tools and techniques used to analyse data

Tools :

- Microsoft excel
- Python
- R
- Jupyter notebook
- Apache spark

- **Sas**
- **Ms power BI**
- **Tableau**

Techniques :

- **Techques based on the mathematics and statistics.**
 1. **Descriptive analysis**
 2. **Dispersion analysis**
 3. **Regression analysis**
 4. **Factor analysis**
 5. **Discriminant analysis**
 6. **Time series analysis**
- **Techniques based on artificial intelligence and machine learning**
 1. **Artificial neural networks**
 2. **Decision trees**
 3. **Evolutionary programming**
 4. **Fuzzy logic**
- **Techniques based on the visualization and graph**
 1. **Column chart bar chart**
 2. **Line chart**
 3. **Area chart**
 4. **Pie chart**
 5. **Box plot**
 6. **Scatter plot**

7. Bubble chart

8. Gantt chart

9. Map

➤ **Barriers to effective analysis**

There are four types of barriers in data analysis

1. Challenge of data discovery

Data discovery is difficult when you have unknown data source, poor quality of data, data silos and data restrictions. This issue can trace their origin to data used or generated by specific application stored in siloed data platform. Additionally, customer incomplete views of viewers and other consent of business, duplicated data and general lack of understanding around what data is available which result in less effective services, insights and customer experience.

2. Excess cost

When your infra is unstructured for utility and elasticity, your talent is expensive and you are facing large, ongoing investments with no guaranteed return, cost can grow out of control.

3. Complexity

Picking the right mix of technologies, identifying architectural best practices for deployment and integrating cloud, on premises and edge these are all complex responsibilities. Yet they are even more difficult if your data platform mix is not optimized.

4. Lack of skills

Most of the companies do not have the skill necessary in house to optimize their data architecture for modern artificial intelligence and machine learning use cases and cloud native applications. To create a modern data fabric, you need

specialized education, training and experience is not organically available in typical IT teams.

➤ Research on two streaming services

- Music(Spotify)
- Video(Netflix)

➤ Research on music streaming services

Spotify is digital music, podcast and video service that provides their customer best music and videos content created by the creators throughout the world.

✓ What type of user data collected by spotify?

When you sign up for the Spotify services or you are creating an account on Spotify, they are asking some information about you for confirmation of viewers and to provide best service to them. The user data which they want as follows :

1. Personal data :

Personal data is used to create users spotify account and user enable to use this service.

- Name
- E-mail address
- Phone number
- Date of birth
- Gender
- Address
- Country

2. Usage data :

Data related to how you are using the Spotify service. It is users actions with Spotify services.

- **Search queries**
- **Streaming service**
- **Playlist you are create**
- **Your watching library**
- **Your browsing history**
- **Account settings**
- **Interaction with other Spotify user**
- **Connections with third party services**

3. Technical data :

- **Information of URL**
- **Cookies data and IP address such online identifiers**
- **Device Ids**
- **Type of network connection(wifi,4G)**
- **Provider**
- **Network and device performance**
- **Browse type**
- **Language**
- **Operating System**
- **Spotify application version**
- **Information enabling digital right management**

4. Data related to general (non-precise) location :

- Meet geographic requirement in your agreement with the owners of content on Spotify
- Deliver content and advertisement that is relevant to you

5. Users device sensor data :

- Eg.(accelerometer and gyroscope)

6. Voice data :

When this service is available and you have choosed it to use voice service they collect and processing voice data.

Eg. Audio recordings

7. Data related to payment and purchase :

- Type of payment method (credit or debit card)
- Users postal code/ZIP
- History related to payment and purchase

8. Data related to research and survey

9. Third party data :

- Social media
- Device including(smart watches, television, game console etc.)

✓ Why Spotify services needs user data?

- The Spotify using user data to provide Spotify service
- To understand, diagnose, troubleshoot and solving issue using this user data
- To analyse and create new features, technologies and improvements to the Spotify services
- It used for marketing and advertisement of Spotify services

- To provide services to user accepting legal conditions and law enforcement request
- To establish, exercise or defend legal claims
- They used to carry business planning, reporting and forecasting
- For processing users payments
- Finding and preventing frauds payments , illegal use of Spotify service and taking strict action on it.
- To conduct service related surveys and research on it

✓ **How do Spotify service end using this data to create better service?**

- Spotify is one of the best music leading streaming service in the world because they are understand how to effectively matching users music taste and suggesting the most desirable playlist.
- Music streaming platform such as Spotify services adopting or we can say acquire data when the users interacts with their services and there are many ways they can get users data like third parties. These services using user data to enhance their algorithms, boost their experiences while using this service, these services targeting potential audiences through adds and improve their business approaches and decisions.
- To provide better experience to the user spotify uses user data, usage data, technical data, users device sensor data, voice data, users payment and purchase data, survey and research data etc. These are type of data is present in spotify hands.
- After collecting various types of data, Spotify services processes these all data. While processing the data Spotify focus on following points :

- purpose of processing data
- Legal basis that permits the purpose
- categories of personal data used for purpose

1. The next step of Spotify is to sharing the users personal data.

The following personal data is always available on these services ex. Users profile name, photo, users public playlist, who is following you on the Spotify service, who you are following on the Spotify service

They share the data with following categories:

1. Categories of recipients
2. Categories of data
3. Reason for sharing

2. Data retention

They keep your data only as long as necessary to provide you with the Spotify service and for Spotify legitimate and essential business purposes, such as

- To maintain Spotify service performance
- Making data driven business decision about new featuring and offering
- Complying with our legal obligation
- Resolving dispute

Criteria used to determine the retention period contains:

- Minimizing data retention period
- Do we need to keep data to ensure the service that user expect?
- Are users able to delete or update the data themself?

- Do we need to keep the data to uphold our rules and keep our service safe?
- Is Spotify subject to legal or contractual obligation to keep or delete the data?

3. Data transfer to the other countries

Spotify provides services globally, so they have to share the personal data internationally to other Spotify branches, their subcontractor as per they are mentioned in their policies.

4. Make sure that user data safe

Spotify provide service like, you can protect your account using strong account password, you have make sure that do not share password anyone, limited access to your browser and computer etc

5. Childrens age matters

Spotify services has age limit to use. The childrens who has less than 13 years they can not use this services as per policy.

➤ Research on video streaming services

Netflix is video streaming services which provides paid online entertainment to viewers such as TV shows, movies, web-series etc.

✓ Which type of user data Netflix collects?

Netflix collects immense data from user. In addition to remembering which show or movie watched and how it was rated it. Netflix needed data as mentioned below

- Users account information : netflix collects users personal information to create an account of user. They collect personal information such as user name, e-mail id, phone number and user plan selection information.

- **Communication setting** : your indicated preferences with respect to Netflix communication you wish to receive(profile and parent control)
- **Marketing communication** : your indicated preferences with respect to Netflix promotional communications on third party services (profile and parent control)
- **Users payment and billing information** : Payment details provided by the user for subscription of plan on Netflix platform.
- **User profile** : any profile created by user on Netflix member account, including playback preferences(available by profile under profile and parent control)
- **Content history** : what type of content user watching that viewing activity and information regarding your interactions with content title on Netflix like movies and TV shows etc.
- **IP address information** : they collect information about user watching content on which recent device etc.

✓ **Why Netflix needs this user data?**

- Provide users with personalized movies and TV show related recommendation
- To recognize the user area of interest content and popular original content
- To personalize marketing content like trailer, adds, thumbnail images etc
- To optimize production planning
- Enhancing business planning and making good decision on it.

✓ **How Netflix ends using this data to create better service?**

- Personalized recommendation

When you open Netflix you see that lots of recommendation are there on the basis of your preference to watch. This is not

coincident because Netflix provide recommendation to users interested content in such a precise manner, but there exist a complex process behind the procedure. It is true that the Netflix does not collect and process user data and uses them to recommend shows and movies on this platform. But what data does it actually process? is it only watched history ? surprisingly the answer is no, watched history is only one part of data collected and processed by them.

Netflix collects data based on the searched history, the time and date show or movie, device used and show and movie skipped or paused by user etc.

2. Trends

Netflix shows trending content at top of homepage based on the what type of content that the crowd are interested to watch. Also what another content users can go through it.

3. Original content

The new concept is offered by Netflix as compared to traditional streaming platform such as “The originals”. As we know Netflix does not care of what type of content dropped by user, paused , skipped, and watched once again by user. Netflix analyses all this things and create exact model type of shows they shall work on. The collection data and perfect algorithm makes Netflix high on demand.

4. Marketing strategy

When you are on Netflix interface you will see that thumbnails or we can say the poster always keep changing in particular time. This is content related to your watching history last time you skipped. This is called A/B testing method. This method using many streaming services and companies, after some time this method is known as trial and hit method. Netflix always trying to test and tries various poster as mentioned above, they even collect data based on the users every click and then they decide what type of poster send on users Netflix interface.

References :

- How to create data science report

<https://thecleverprogrammer.com/2022/05/27/heres-how-to-create-a-data-science-project-report/>

- Research on music streaming services

<https://www.spotify.com/in-en/legal/privacy-policy/#3-personal-data-we-collect-about-you>

<https://www.spotify.com/in-en/privacy>

- Data analytics

<https://www.simplilearn.com/data-analysis-methods-process-types-article>

- Research on music streaming services

<https://help.netflix.com/en/node/100624>

<https://www.smartdatacollective.com/how-netflix-utilizes-user-data-to-create-personalized-user-experience/>