Netflix Data Analysis

Presenters:

Miss. Tahseen Begum\_2010030168

Miss. Keerthana Pulugam\_2010030445

Miss. E.Pravallika\_201003004

Miss. Sowgna\_2010030344

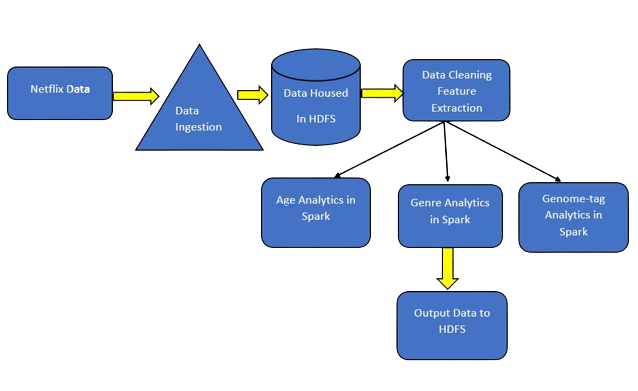
Guide: Chanda Raj Kumar Rao

**Problem Statement:**

I’ll take a look at some very important models of Netflix data to understand what’s best for their business. Some of the most important tasks that we can analyze from Netflix data are:

1. Understand what content is available.
2. Understand the similarities between the content.
3. Understand the network between actors and directors.
4. What exactly Netflix is focusing on.
5. Sentiment analysis of content available on Netflix.

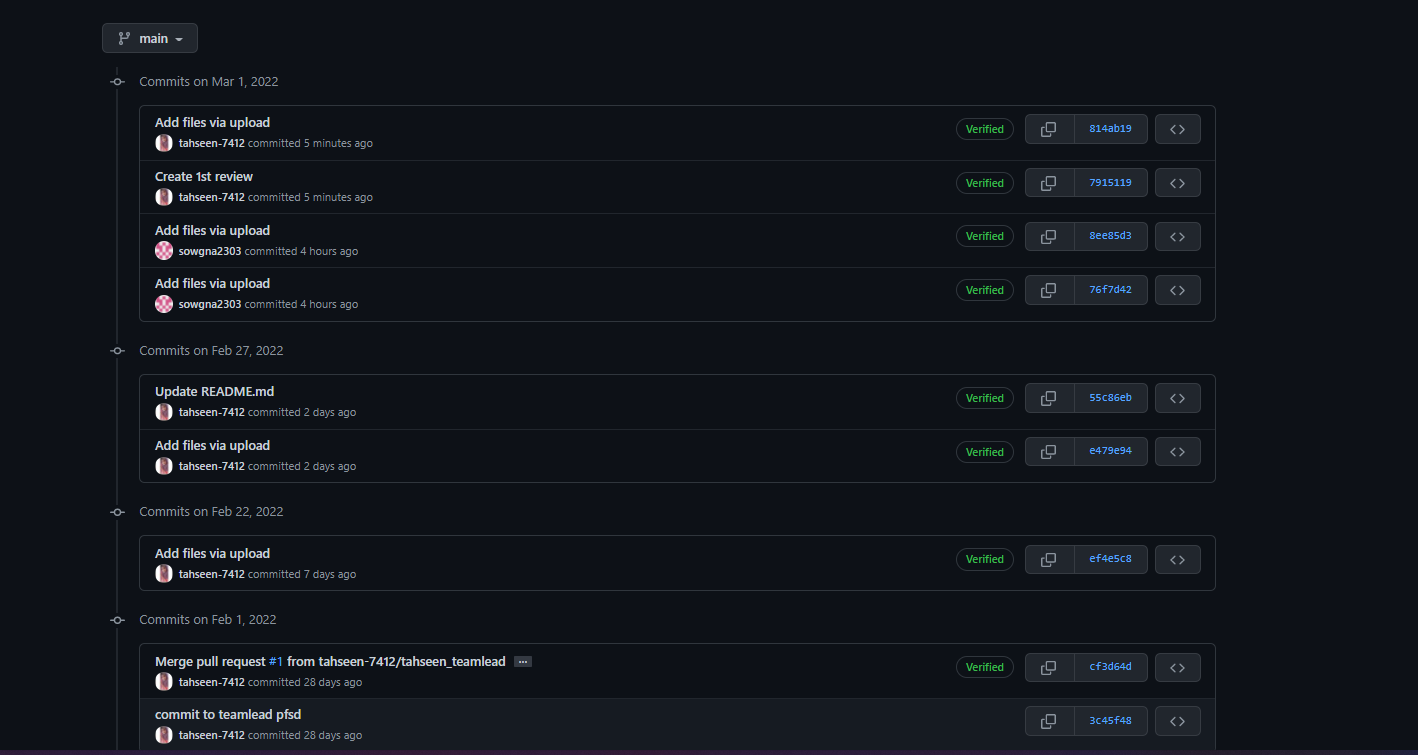
**Flowchart**

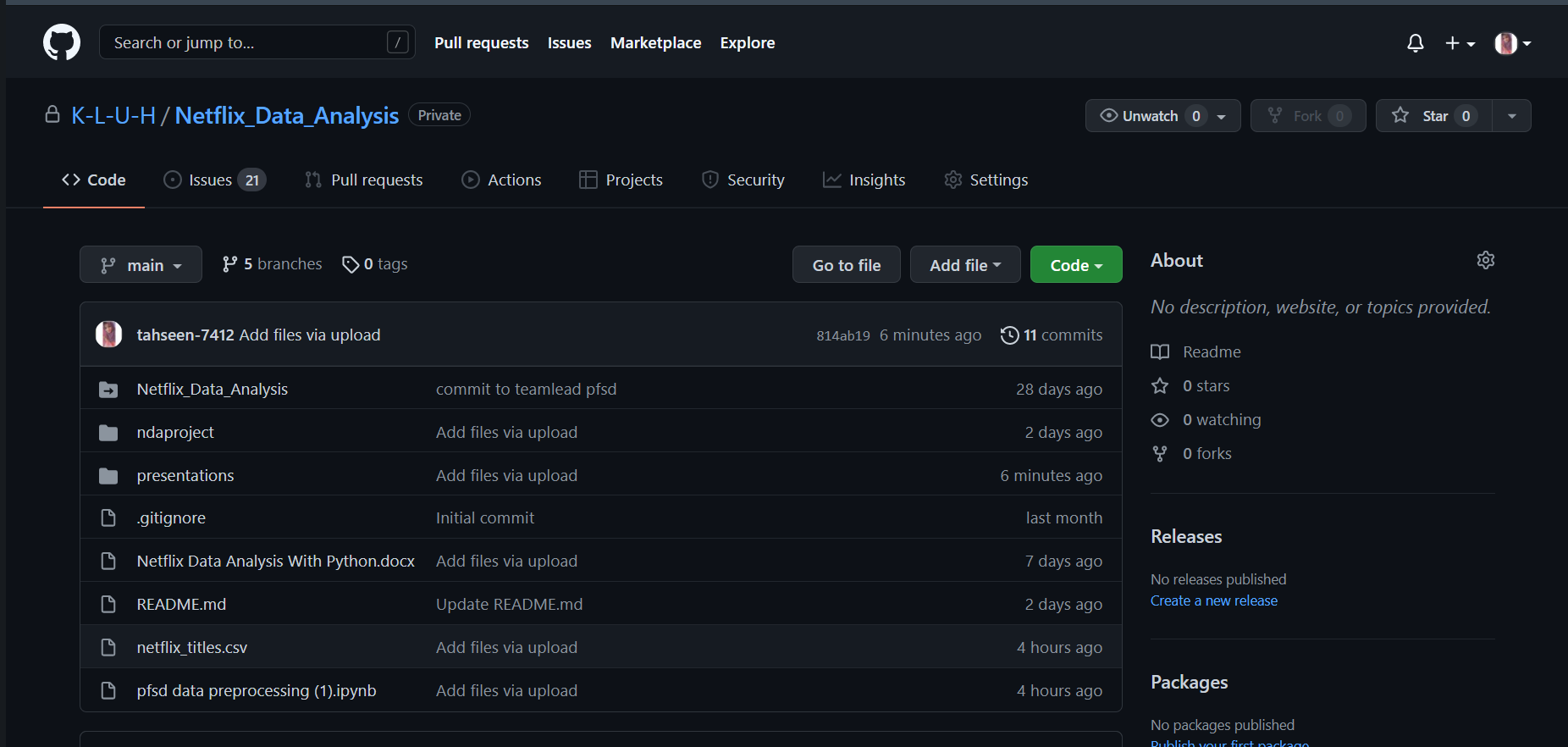
****

**Literature Survey**

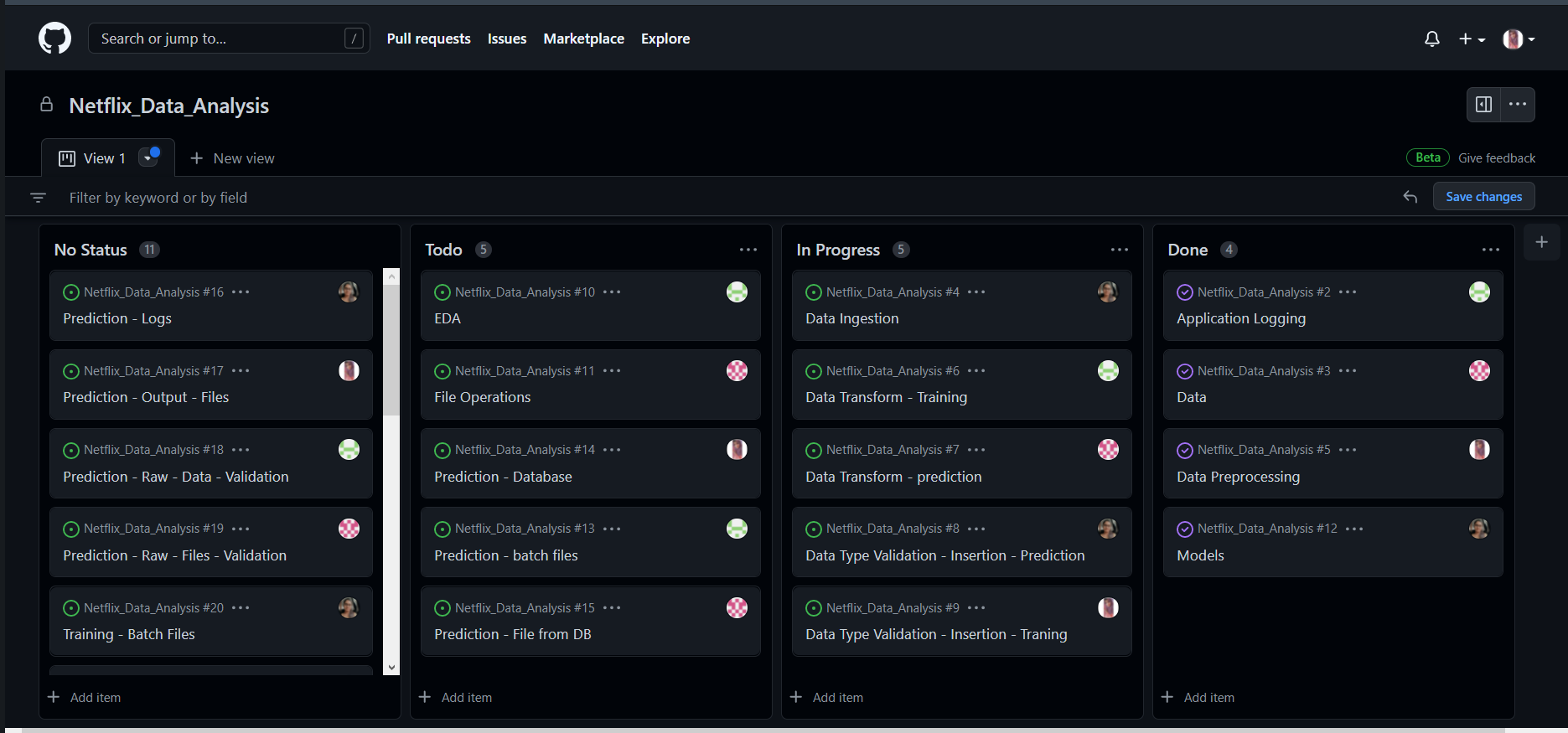
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author** | **Title** | **Publishing** | **Dataset & Techniques** | **Pros** |
| Rajeswari Nakka Dr.G.V.S.N.R.V.Prasadand R.Kiran Kumar | Offering Recommendations on Netflix dataset by Associations among Users as Trust Metric | 2021 | By this model several batches of data points are sampled among total dataset points. The results are well analysed over error rate where the proposed technique tends to have the reduced error value compared to existing Collaborative Filtering technique. In future the work can be extended on larger data points and address issues like memory issues, and utilize optimized matrix factorization technique to move a step further. | The model is evaluated using an evaluation metric Mean Squared Error (MSE) value.  The performance of the proposed algorithm on the Netflix dataset was compared with existing algorithm |
| Vadloori, Karthik Babu, and Shriya Madhavi Sanghishetty. | Exploratory and Sentiment Analysis of Netflix Data | 2021 | An Open-Source Data Set obtained from Kaggle – that was wrangled and exercised to derive maximum insights using EDA – Exploratory Data Analysis and Sentiment Analysis after the amalgamation of two additional sets – Geographical Latitudes & Longitudes and Netflix Title Critics/Reviews Data Set. The project is made using different utility analytical tools present in Python Library of versatile packages | Introduces systematic and insightful usage of methods for Exploratory Data Analysis & Sentiment Analysis by utilizing various packages concerned. |
| **Author** | **Title** | **Publishing** | **Dataset & Techniques** | **Pros** |
| **Rajeswari Nakka Dr.G.V.S.N.R.V.Prasadand R.Kiran Kumar** | **Offering Recommendations on Netflix dataset by Associations among Users as Trust Metric** | **2021** | **By this model several batches of data points are sampled among total dataset points. The results are well analysed over error rate where the proposed technique tends to have the reduced error value compared to existing Collaborative Filtering technique. In future the work can be extended on larger data points and address issues like memory issues, and utilize optimized matrix factorization technique to move a step further.** | **The model is evaluated using an evaluation metric Mean Squared Error (MSE) value.**  **The performance of the proposed algorithm on the Netflix dataset was compared with existing algorithm** |
| **Vadloori, Karthik Babu, and Shriya Madhavi Sanghishetty.** | **Exploratory and Sentiment Analysis of Netflix Data** | **2021** | **An Open-Source Data Set obtained from Kaggle – that was wrangled and exercised to derive maximum insights using EDA – Exploratory Data Analysis and Sentiment Analysis after the amalgamation of two additional sets – Geographical Latitudes & Longitudes and Netflix Title Critics/Reviews Data Set. The project is made using different utility analytical tools present in Python Library of versatile packages** | **Introduces systematic and insightful usage of methods for Exploratory Data Analysis & Sentiment Analysis by utilizing various packages concerned.** |

**Git commits (Interval 11 commits after 1st review)**

****

****

**Work In Progress**

****

**Pre-processing code and techniques**

Step-1:

!pip install matplotlib

Step-2:

#importing the dataset

import numpy as np # linear algebra

import pandas as pd # data processing, CSV file I/O (e.g. pd.read\_csv)

import matplotlib.pyplot as plt

Step-3:

#to import csv file

netflix\_data = pd.read\_csv(r"C:\Users\Sowgn\OneDrive\Desktop\pfsd\netflix\_titles.csv")

Step-4:

netflix\_data

Step-5:

#head function

#to show top-5 records of the dataset

netflix\_data.head()

Step-6:

#tail function

#to show bottom-5 records of dataset

netflix\_data.tail()

Step-7:

#shape

#to show no of rows and colums

netflix\_data.shape

Step-8:

#size

#to show no of total values(elements) in the dataset

netflix\_data.size

Step-9:

netflix\_data.columns

Step-10:

netflix\_data.dtypes

Step-11:

netflix\_data.info()

Step-12:

netflix\_data.value\_counts()

Step-13:

netflix\_data.nunique()

Step-14:

netflix\_data[netflix\_data.duplicated()]

Step-15:

netflix\_data.drop\_duplicates()

Step-16:

netflix\_data.isnull()

Step-17:

netflix\_data.isnull().sum()

Step-18:

!pip install seaborn

Step-19:

import seaborn as sns

Step-20:

sns.heatmap(netflix\_data.isnull())

Step-21:

netflix\_data[netflix\_data['title'].isin(['Blood & Water’])]

Step-22:

netflix\_data['Date\_N']=pd.to\_datetime(netflix\_data['date\_added’])

Step-23:

netflix\_data.head()

Step-24:

netflix\_data['Date\_N'].dt.year.value\_counts()

Step-25:

netflix\_data['Date\_N'].dt.year.value\_counts().plot(kind='bar’)

Step-26:

netflix\_data.groupby('type').type.count()

Step-27:

sns.countplot(netflix\_data['type’])

Step-28:

netflix\_data['year']=netflix\_data['Date\_N'].dt.year

netflix\_data.head(2)

Step-29:

netflix\_data[(netflix\_data['type']=='Movie')&(netflix\_data['year']==2022)]

**Alpha Testing**

<!DOCTYPE html>

<html>

<head>

<meta name="viewport" content="with=device-width, initial-scale=1.0">

<title>Netflix Data Analysis</title>

<link rel ="stylesheet" href="style.css">

<link rel="preconnect" href="https://fonts.googleapis.com">

<link rel="preconnect" href="https://fonts.gstatic.com" crossorigin>

<link href="https://fonts.googleapis.com/css2?family=Poppins:wght@100;200;300;400;600;700&display=swap" rel="stylesheet">

<link rel="stylesheet" href="https://stackpath.bootstrapcdn.com/font-awesome/4.7.0/css/font-awesome.min.css">

</head>

<body>

<section class="header">

<nav>

<a href="index.html"><img src="images/logo.jpg"></a>

<div class="nav-links" id="navLinks">

<i class="fa fa-times-circle" onclick="hideMenu()"></i>

<ul>

<li><a href="index.html">HOME</a></li>

<li><a href="about.html">ABOUT</a></li>

<li><a href="movie.html">MOIVES</a></li>

<li><a href="data\_analysis.html">DATA ANALYSIS</a></li>

<li><a href="contact.html">CONTACT</a></li>

</ul>

</div>

<i class="fa fa-bars" onclick="showMenu()"></i>

</nav>

<div class="text-box">

<h1>Unlimited movies, TV shows and more.</h1>

<p>Enjoy on your TV.

Watch on smart TVs, PlayStation, Xbox, Chromecast,<br>Apple TV, Blu-ray players and

more.</p>

<a href="" class="hero-btn">Visit Us to know More about it...</a>

</div>

</section>

<!---- Moives ---->

<section class="moives">

<h1>Enjoy on your TV.</h1>

<p> Watch on smart TVs, PlayStation, Xbox, Chromecast, Apple TV, Blu-ray players

and more.</p>

<div class="row">

<div class="moive-col">

<h3>English</h3>

<p>Red Notice · Back to the Outback · Midway · Sonic the Hedgehog · Spenser Confidential · Minions · Jumanji: Welcome to the Jungle </p>

</div>

<div class="moive-col">

<h3>Hindi</h3>

<p>Red Notice · Back to the Outback · Midway · Sonic the Hedgehog · Spenser Confidential · Minions · Jumanji: Welcome to the Jungle </p>

</div>

<div class="moive-col">

<h3>Telugu</h3>

<p>Red Notice · Back to the Outback · Midway · Sonic the Hedgehog · Spenser Confidential · Minions · Jumanji: Welcome to the Jungle </p>

</div>

<div class="moive-col">

<h3>ALL</h3>

<p>Red Notice · Back to the Outback · Midway · Sonic the Hedgehog · Spenser Confidential · Minions · Jumanji: Welcome to the Jungle </p>

</div>

</div>

</section>

<!----about--->

<section class="about">

<h1>Global Networks</h1>

<p> Watch on smart TVs, PlayStation, Xbox, Chromecast, Apple TV, Blu-ray players

and more.</p>

<div class="row">

<div class="about-col">

<img src="images/moive.jpg" >

<div class="layer">

<h3>BIGBOSS</h3>

</div>

</div>

<div class="about-col">

<img src="images/Bheemla.jpg" >

<div class="layer">

<h3>BHEEMLA NAYAK</h3>

</div>

</div>

<div class="about-col">

<img src="images/radheshyam.jpg" >

<div class="layer">

<h3>RADHE SHYAM</h3>

</div>

</div>

<div class="about-col">

<img src="images/rrr.jpg" >

<div class="layer">

<h3>RRR</h3>

</div>

</div>

</div>

</section>

<!---------- data ana ---->

<section class="data">

<h1>Data Analysis</h1>

<p>here we can see the percentage...</p>

<div class="row">

<div class="data-col">

<img src="images/netflix.png">

<h3>Data Neflix</h3>

<p>hindi movies</p>

</div>

<div class="data-col">

<img src="images/netflix.png">

<h3>Data Neflix</h3>

<p>telugu movies</p>

</div>

<div class="data-col">

<img src="images/netflix.png">

<h3>Data Neflix</h3>

<p>english movies</p>

</div>

</div>

</section>

<!---------- call to action-------->

<section class="cta">

<h1>Team Contact</h1>

<a href="" class="hero-btn">CONTACT US</a>

</section>

<!----- footer ---------->

<section class="footer">

<h4>About Us</h4>

<p>Students from KL University</p>

<div class="icons">

<i class="fa fa-facebook"></i>

<i class="fa fa-twitter"></i>

<i class="fa fa-instagram"></i>

<i class="fa fa-linkedin"></i>

<i class="fa fa-github"></i>

</section>

<!-----JavaScript for Toggle Menu----->

<script>

var navLinks = document.getElementById("navLinks");

function showMenu(){

navLinks.style.right = "0";

}

function hideMenu(){

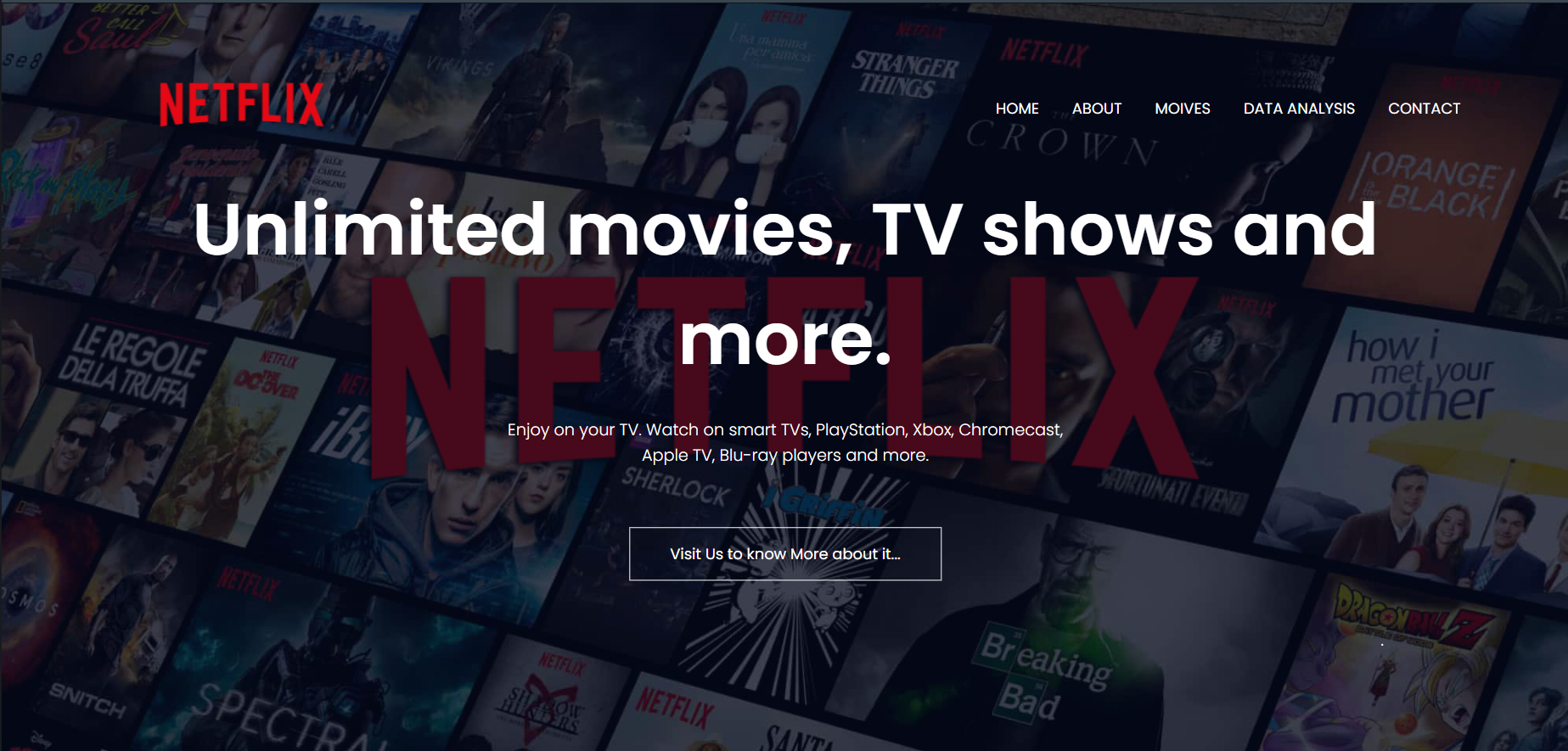
navLinks.style.right = "-200px";

}

</script>

</body>

</html>



**Division of work among the group members**

* N. Sowgna-Pre-processing code and techniques
* Tahseen Begum-Alpha Testing
* E. Pravallika – Flowchart, Doc
* Keerthana -PPT

**Conclusion**

* We have completed the front-end for this project.
* Next we will work on the back-end.

**References**

* <https://www.researchgate.net/publication/345710976_Offering_Recommendations_on_Netflix_dataset_by_Associations_among_Users_as_Trust_Metric>
* <https://www.researchgate.net/publication/354719521_Exploratory_and_Sentiment_Analysis_of_Netflix_Data>
* <https://www.researchgate.net/publication/305741976_Data_management_in_audiovisual_business_Netflix_as_a_case_study>
* <https://www.youtube.com/watch?v=b7Kd0fLwgO4>
* <https://www.kaggle.com/shivamb/netflix-shows>