**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

| **Team Member’s Name, Email and Contribution:** |
| --- |
| * **ANKUSH** * **EMAIL** - [90ankushsaini@gmail.com](mailto:90ankushsaini@gmail.com) * Solo Project |
| **Please paste the GitHub Repo link.** |
| Github Link:-https://github.com/Ankushsaini90/NETFLIX-MOVIES-AND-TV-SHOWS-CLUSTERING |
| **Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)** |
| We have a dataset consisting of tv shows and movies available on Netflix as of 2019. The dataset is collected from Flexible which is a third-party Netflix search engine.  In 2018, they released an interesting report which shows that the number of TV shows on Netflix has nearly tripled since 2010. The streaming service’s number of movies has decreased by more than 2,000 titles since 2010, while its number of TV shows has nearly tripled. It will be interesting to explore what all other insights can be obtained from the same dataset.  In this project, we are required to do Exploratory Data Analysis, Understanding what type of content is available in different countries, Netflix has increasingly focused on TV rather than movies in recent years or not, and Clustering similar content by matching text-based features.  So, First we do Exploratory Data Analysis on the data set. We look for missing data values and deal with them. Then outliers and appropriately modify them. We also perform feature engineering to modify a few existing columns and drop out irrelevant ones.  We then look at several popular individual models from simple ones K-means clustering, Hierarchical clustering and the recommended system.    Finally, after facing some challenges, we found the solution for our model and we came to the conclusion that It's clear that Netflix has grown over the years. We can see from the data that the company took certain approaches in their marketing strategy to break into new markets around the world. Based on an article from Business Insider, Netflix had about 158 million subscribers worldwide with 60 million from the US and almost 98 million internationally. Netflix's original subscriber base was based solely in the United States following its IPO. A large part of its success was due to the decision to expand to international markets. The popular market prioritizes what content the company will release. In this case, we can see that a good amount of international movies and TV shows were added over the years as part of Netflix's global expansion. |