JOURNAL OF
GENERAL MANAGEMENT RESEARCH

ANTECEDENTS OF RECOMMENDATION FRAMEWORK ENABLED THROUGH ARTIFICIAL INTELLIGENCE ON NETFLIX PLATFORM

Aman Tibrewal*
Rupashi Sehgal*
Dr. Shikha Singh*

*Symbiosis Centre for Management Studies, NOIDA, Symbiosis International (Deemed University), Pune India

Abstract

With the rising technological advancements, various tools like Artificial Intelligence and recommendation are utilized by companies to gain an upper edge in the markets. This paper aims to gain a deeper understanding of how Netflix uses Artificial Intelligence to enable effective usage of the recommendation engines in their business model. To conduct an effective study, this paper follows a triangulation method, under which a literature review was conducted initially to gain a conceptual understanding of the topic. Based on it, structured personal interviews and a survey was conducted. Lastly, a quantitative analysis was executed to gain a deeper insight into the data as collected. This paper aims to gain a deeper insight into the functioning of recommendation systems for Netflix's Over-the-top (OTT) media services. The paper also focuses upon the usage of Artificial Intelligence technology which acts as a key enabler for Recommendation Engines for efficient and effective utilization to gain a competitive advantage for Netflix. This paper is of value to each individual who seeks to understand the mind behind the model of Netflix and how the user interface allows them to collect data and research upon the viewing habits of their users. It moreover also focuses on the integration of Artificial intelligence in the Recommendation engine at Netflix.

Keywords: *Recommendation Engine, Netflix, Over-the-top (OTT) Media Service, Artificial Intelligence*

INTRODUCTION

In October 2006 Netflix initiated a platform seeking for a reliable, efficient and more accurate movie recommendation system which would successfully and entirely change the previously existing system known as "cinematch". To change with the technology and adapt to the dynamics is very essential nowadays for various web platforms as well as to other web providers on the web platform in order to gauge the interest of the users. Taking Amazon's example, their main strategy is basically to present users with what they want so that their likeness towards the product and service is more. Similarly, talking about the Netflix platform, their key driver, objective and motivation is therefore to help clients choose movies of their taste and preferences through their recommendation model in order to maintain a long-term association and subscription.

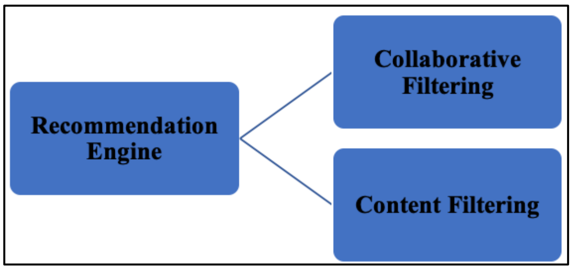
The new era of technology advancement has given a major boom to Netflix therefore leading to the broadening of the models based on algorithms blended with the concept of machine learning. Further, the paper discusses the methods and concepts that is being used currently at Netflix and how their recommendation system architecture is structured. The introduction is further divided into building an understanding of the concept and model of recommendation engines, Netflix, Artificial Intelligence and Recommendation Agent in Netflix.

Concept of Recommendation Engines

The rise of the digital markets has drastically augmented the selections of obtainable goods and services to consumers. The consumers are served with a very customized product combination adjusted to cater to their needs which helps smoothen the sales funnel. Recommendation lists, generated by the recommendation engines, are one of the most productive efficacies of machine learning in day-to-day involvement. (Kotu & Deshpande, 2019). Recommendation Systems are an agent of sorts for the user of its technology. This paper solely focuses upon the utilization of the same by Netflix. The recommendation engines use behavioral and preferential information. Internet search engines are a good example of such content-based systems, as they extract results based on keywords. The frequency of usage of a searched term is used to evaluate a document's rank, and the comparative incidences of the target terms are used to assess document similarity (Salton & Buckley, 1988). Recommendation systems are a framework applied to user-level forecasting that can be developed even with a few alternatives. For illustration, any individual in a situation where they are supposed to select a movie for their consumption, might have a large number or a small number of options to select from. The recommendation systems will assess the user's preferences and the past behavior of other consumers, based on which the engine will produce a recommendation to the individual with a target to suggest them with the most relevant movie, series, or shows as per their preferences. The recommendation system functions on data which are majorly accumulated from a variety of sources. They are: an individual's articulated likings or selections amid alternative movies, preferences for product attributes (e.g. director, producer, actor etc.), other users with similar decisions before, expert judgements (critics, journalists etc.) and individual features that assist forecast preferences as stated by Ansari *et al.,* (2000).

Using the above features the recommendation list is generated. An effective recommendation engine is built using any or all of these five types of data, to be able to make accurate recommendations as more information is collected. The functioning of a recommendation engine can be defined into two classes based upon their source of information. First-class is based upon collaborative filtering, which creates a linear combination of the general universe's (total population of data available within the information systems) preferences to predict an individual's preferences. The second class, identified to be content filtering, generates recommendations based upon the consumer's preferences for the product's attributes, i.e. the movie genre, actors, directors, etc.

Figure 1: Classification of Recommendation Engine Based on Source of Information



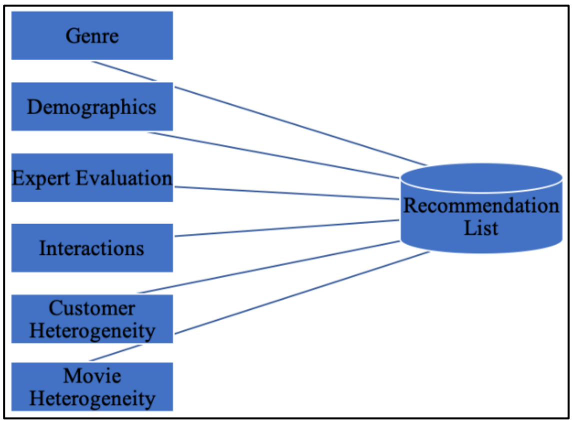
Source: Authors' Compilation (based on literature)

The recommendations made by a recommendation engine are based upon various heterogeneous patterns that are identified by a recommendation system that is enabled by Artificial Intelligence. These patterns are recognized by the recommendation engine by analyzing the datasets that are available using various predefined algorithms. As given by Ansari *et al.,* (2000), there are three major patterns studied by a recommendation engine. The first one is customer heterogeneity. This database is made using the rating provided by the consumers for various non-similar movies. Second one is product (movie) heterogeneity: a database that consists of various information based upon the movies and their attributes over several data points. The third one is customer and product heterogeneity: a conjoint analysis of both the datasets compiled together for a more meaningful output.

Model Specification and Variable Definition

The conceptual model for the functioning of a recommendation system consists of both forms of heterogeneity while considering all the sources of data that is being available for the information system of a recommendation system, as well as both the classes of recommendation engine i.e., defined by the logical reasoning behind the output recommendation list (refer to Figure 1). The model regarding the recommendation system is visualized in Figure 2.

Figure 2: Conceptual Model for Recommendation System



Source: Authors' Compilation (based on literature)

The recommendation list has six broad components which form the recommendation engine model. Firstly, the content is categorized on the bases of its genre, it broadly divides the movies and classifies them based on the theme and the storyline of the video content. Then it identifies the demographics of the viewers. Demographic value takes into consideration the age, gender, and similar features of the target audience into account for the recommendation system. The list also factors in the expert's evaluation which consists of the reviews by critics and industry experts which acts as a key variable that can be taken into consideration by the recommendation systems. This is followed by the interaction patterns which are built upon the reviews collected by the users that function as an important variable which helps in improving the recommendation engine systems. Customer heterogeneity contributes to the recommendation engine which suggests movies that are more frequently watched by a similar group of people, i.e., if a large group of people watch movie A and then watch movie B. So, the engine is more likely to recommend movie B, if you are searching for or watching movie A. This is combined with movie heterogeneity, using which the recommendation engine takes into consideration how different features of the movie are perceived by different customer groups. For example, if you recently watched a movie that cast a particular actor. Then the engine is more likely to further suggest your movies in which the same actor was featured.

Netflix

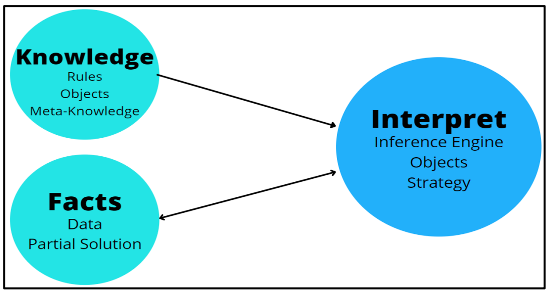
Netflix Ltd. is an American production and OTT (Over the top) media-services provider company. The company's primary business model is its subscription-based streaming service that offers internet-based streaming of a collection of films and television programs. Netflix operates on a subscription-based model. The user pays for a monthly subscription and is given access to all the shows, movies, documentaries, and other content available on Netflix. This subscription-based framework gives the user an advertisement-free experience to their collection of shows and movies with unrestricted access. The foremost attraction of the platform initiates from its capacity to produce a pleasant sense of choice and immediacy as explained by Havens (2018). Netflix uses machine learning and deep learning algorithms to build their artificial intelligence, to help their algorithms "learn" without human assistance. Machine learning gives the platform the ability to automate millions of decisions as per user activity as stated by Codecademy (2020). The necessity for recommendation engines and a personal touch is an outcome of a phenomenon known as the "era of abundance".

The need of this study is to gain a deeper understanding of how Netflix uses AI to enable effective usage of the recommendation engines in their business model. This paper also aims to gain a deeper insight into the functioning of recommendation systems for Netflix's Over-the-top media services by focusing upon the usage of AI technology which acts as a key enabler for Recommendation Engines for efficient and effective utilization to gain a competitive advantage for Netflix. Netflix functions at the juncture of technology and storytelling. Netflix has devised its segment of Internet television. Their core product revenue source is a subscription-based service that gives their customers access to watch movies and TV shows whenever they want on a large range of devices with internet connectivity. The Internet television segment is young and full of competition; thus, innovation is vital. An important innovation to the product is the recommender system embedded in the user interface which helps the viewers find videos to watch in every session. Their recommender system is built on a deep neural network, combined with unsupervised and supervised machine learning algorithms which contribute towards creating the complete Netflix experience.

Artificial Intelligence

"Artificial Intelligence (AI) is a field of science and engineering concerned with the computational understanding of what is commonly called intelligent behavior, and with the creation of artifacts that exhibit such behavior." as explained by Shapiro (1992). AI is a combination of computer structures implemented and designed that are proficient in resolving problems that frequently require the ability of human intellect. Such problems resemble natural everyday tasks and/ or a highly complicated situation (e.g. visual or natural language understanding). For solving them, AI algorithms mainly deal with a huge variety of data types as pointed out by Haton (2006). There are three major domains of AI that are identified by Shapiro (1992). They are computational psychology, computational philosophy and advanced computer science. Artificial Intelligence has reached some boundless heights and spectacular successes in the field of knowledge-based systems. These methods rely on knowledge that is inspired by human expertise. The following figure 3, demonstrates the principle of knowledge-based systems.

Figure 3: Principle of Knowledge-Based AI Systems



Source: Haton (2006)

AI technology is growing due to their essential characteristics that give striking properties to Artificial Intelligence systems. Features like the possibility of dealing with inadequate, conflictual, and missing data which makes the system scalable along with the organization, its size and requirements. Ease of investigating and scalability which makes it possible to incrementally process large requirements. Providing explanation and the reasoning behind the model's workings, which is important both for debugging flaws and for orienting new human experts. A normally developed knowledge representation model is built upon small, independent portions of knowledge which simple decision-making functions. Most of such models are built upon the following conditioning function:

IF Condition THEN Conclusion

The function is utilized by experts in various domains to evaluate the decision-making and logical process of experiential information. In a data-driven method, it is utilized from left to right (forward chaining). Whereas in a goal-driven method, it's used in a right to left (backward chaining) approach.

Recommendation Engine at Netflix

Here are a few samples of how the recommendation system of Netflix works based on the classification in the model developed in Figure 2. The figures attached below are taken:

1. *Based on Genre*

Netflix helps its members by suggesting their movies or shows of different genres. It broadly divides the movies and classifies them based on the theme and the storyline of the video content. For example, if a movie is themed around Thrill or Comedy, then all such movies are bundled together.

2. *Based on Demographic*

Netflix uses demographic data to aids its members by suggesting movies or shows based on similar clusters. Demographics means a classification of the population based on age, race, and sex. In this way, they can target different segments of the audience. For example, if a majority of the users in a particular age group prefer to watch a particular cluster of content, then the recommendation engine uses the same to improve the accuracy of the engine.

3. *Based on Expert Evaluation*

Expert evaluation is a key component in the recommendation list manufactured by Netflix. The reviews by critics and industry experts may act as an important variable that can be taken into consideration by the recommendation systems. With this, they can suggest to their viewers some of their best productions from their content catalogue.

4. *Based on Interactions*

Netflix collects the interaction data from the users. And based on the historical data like the viewing history or the watchlist of the user, they are able to analyze user behavior and customer preference. This is used to further improve the recommendation engine results.

5. *Based on Customer Heterogeneity*

Netflix suggests its viewers' movies or shows based on customer heterogeneity. Customer heterogeneity acknowledges the fact that just like human beings, every customer is different and so are their preferences and interests. It embraces the concept that every single customer is different and that customization and personalization are important, hence they came up with the concept of a recommendation engine.

6. *Based on Movie Heterogeneity*

Netflix has induced heterogeneity into the modelling for their recommendation systems. The recommendation engines take into consideration how different features of the movie or the show affects the different customer segments. For example, if you recently watched a movie that cast a particular actor. Then the engine is more likely to further suggest your movies in which the same actor was featured.

LITERATURE REVIEW

The paper makes use of the triangulation theory. Triangulation theory suggests multiple approaches to research a question. The goal is to increase the accuracy and the confidence in the findings through validation of a proposition using multiple independent variables (Heale & Forbes, 2013). Triangulation is the utilization of different speculations, strategies, information sources, or specialists inside the examination

identifying with a solitary wonder (Zhang & Wildemuth, 2009). Both qualitative and quantitative examination is utilized in the investigation of this paper. The qualitative content analysis helps gain an understanding of the literature that's available. Therefore, contents were investigated for the topical idea and drawn the induction by looking at those subjects (Zhang & Wildemuth, 2009), which is very crucial for the enhanced understanding towards fulfilling the objectives of the research paper. Artificial intelligence has made enormous progress in the field of information-based frameworks. These frameworks use information that depends on fundamental human aptitude. These key attributes give alluring properties to master frameworks. Different frameworks have perceived the matchless quality of a limited quantity of data to empower scholarly dynamics, either in an independent technique or in contiguous collaboration with an individual. From that point forward, significant advancement has been made in comprehension and applied thinking. AI is worried about the cost and execution of PC frameworks that are equipped for tackling issues that typically require the capacity of individuals. These unpredictable issues can't be addressed by exemplary calculation techniques. To tackle these issues, AI controls representative data just as mathematical information in software engineering (Haton, 2016). AI is commonly known as intelligent behavior along with the creation of artifacts that display this behavior. AI has also been defined as a field of engineering and science along with software and computational knowledge. There are goals that the researchers must examine before researching this field. The first one is computational psychology. The objective of computational psychology is to gain a deeper understanding of human intelligent behavior and its basic functioning by generating computer programs that perform in the same way people do. The second one is Computational Philosophy whose objective is to develop a computational consideration of the human category of intelligent behavior, and not be restricted to the data structure and procedures that a human mind might use and the third one is Advanced computer science. The objective of Advanced Computer Science is to push the boundaries of the frontier of what is known about how to program on computers, specifically in the direction of tasks that are perceived to be unprogrammable, but people can perform them (S.C Shapiro, 1992). This research paper examines the application of Artificial intelligence (AI) in various programming problems. Even though AI was criticized as being an impossible and unachievable goal, it is now recognized by the general public as an application that can be used to solve everyday problems.

It is also important to note how the proposed framework has achieved an immense change on the OTT channels. It is likewise called the current customization framework which further groups itself into two classes that utilize various wellsprings of data to make suggestions. The first being communitarian sifting, which impersonates informal proposals. Through this, it can foresee the inclinations of the individual as a direct weighted mix when contrasted with the public's inclination. The subsequent one being content separation which makes suggestions based on inclinations of the buyer for item credits. Notwithstanding the communitarian and content sifting, this paper expresses that five data sources can be utilized to make suggestions as follows : (i) an individual's inclinations among elective items, (ii) inclinations as far as item credits, (iii) Other individuals' decisions or inclinations, (iv) master decisions, and (v) singular attributes that might anticipate inclinations. A decent suggestion framework ought to have the option to join this load of five kinds of data to have the option to make a decent proposal framework. Suggestion frameworks are models for the singular level forecast

ISSN 2348-2869 Print
ISSN 2348-5434 Online

that can be useful regardless of whether there are a couple of substitutes. For instance, an individual choosing new arrivals of music, plays, or films might have a couple of decisions (Essegaier *et al.,* 2000). Netflix applies exceptional information concentrated strategies to underwrite the survey exercises of the individuals. The organization's key extraordinary potential relies upon its energetic procedures for client commitment. Constitutive of its user-interface and content-based structure, Netflix's observational direction toward gorge or "information base watching" has a significant capacity in outlining the brand as a moderator of its supporters. Netflix experience is about how we relate it to ourselves as well as other people through the utilization of streaming media (Pilipets, 2019). It is important to investigate the different calculations that are expected to make the proposed system that is utilized by Netflix effectively grow its business scope. It further clarifies the job of related calculations and web search tools which proliferates into a proposal framework. It is also important to investigate the inspiration that is utilized behind assessing the way to deal with further developing the suggestion framework. In this existence where the contest is prepared, development in the business field is urgent, and to get by in such a rivalry Netflix needed to think of a proposal motor that furnishes the clients with customization. The proposal framework that Netflix utilizes isn't simply founded on one calculation however an assortment of calculations serving diverse use cases that are consolidated together to give the watchers a total Netflix experience. Netflix likewise tests its calculations on two gatherings of individuals, one being the current individuals from Netflix and the other being the new individual from the Netflix family. Netflix likes to test their new individuals since they have not encountered an alternate form of the item previously. Their reactions will in general be reminiscent of the viability of the elective form of the calculation and as opposed to the change from old to new that is ordinarily capable by the current individuals (Carlos & Neil, 2015).

METHODOLOGY

Data Source Selection

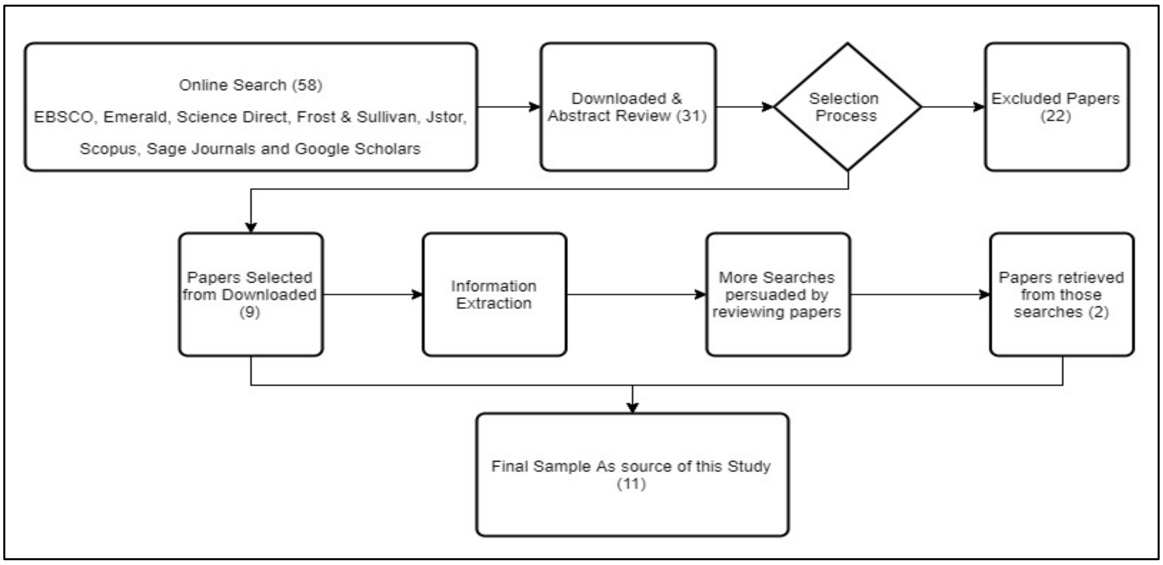
For potential writing thinking about the examination system, online data sets were investigated. As the examination depends on the Recommendation Engine at Netflix, hence the inquiry begins from "Predecessors of Recommendation Framework empowered through Artificial insight on Netflix Platform". Further, certain catchphrases were appended to limit the list items to a more ideal outcome. Watchwords, like the following, “Recommendation Engine", “Artificial Intelligence", "Netflix" and “OTT", were used to further look for online sources.

Selection by Inclusion and Exclusion

Although there were a large number of search results for the searches conducted, primarily around 58 articles were nominated for the sample selection, from which 31 articles were downloaded and collected for review. The investigation theme was checked recently to an evaluation of downloaded articles to decide if the substance of articles is coordinated through to the examination foundation. These articles post review was distributed into three categories, i.e. group one consisting of the articles that contain literature related directly to the Recommendation Engines and their basic functioning, group two consisting of the articles that contain literature related directly to Netflix, their day to day functioning, and the integration of Recommendation Engines and lastly, for group three consisted of papers based upon how AI function and acts as an enabler for Recommendation Engines. The selected articles were stored separately and an in-depth review was conducted on the same.

Further, the downloaded articles after a review were filtered based on their relevance to this paper’s core theme and quality of content based on the authors’ utility for the same, which ultimately led to a selection of nine papers out of the 31 articles as downloaded. The selected articles presented a few newer concepts and models to the researcher, which further led to the detection of another set of two relevant articles. Thus, finally providing us with a set of 9+2 = 11 articles, which act as a source of sample for this study. These chosen test articles were considered for building a hypothetical establishment toward Recommendation Engines. Furthermore, more papers were created from the catalog of value papers. Google search led to understanding the significance of some mind-boggling terms and further various online articles and sources were accessed for a deeper understanding of the functioning of the recommendation engines and the Netflix business model. The methodology of paper choice has been shown underneath in Figure 10.

Figure 10: Literature Selection Process



Source: Authors' Compilation (based on literature)

Statement of Problem

Netflix gained a very huge competitive advantage via the concept of Personalization on the consumer level (Victor, 2007). Netflix doesn't have any particular slogan or a tagline, or any focus upon any set genre. Netflix as a brand works on "personalization" (Liftigniter, 2016). It utilizes a large amount of data collected to improve its recommendation system, which also further assists them with acquiring experiences on client content inclinations and review propensities. Netflix tends to keep the data collected and generated by them as private and doesn't release much out to the public. In addition, there is a very minor number of researches conducted on such a similar theme. With merely a very minor number of researchers looking into Netflix's personalization feature. By get-together and surveying essential information from genuine Netflix streaming clients, this investigation intends to convey a more profound understanding of the relationship between the watcher/client and Netflix's suggestion framework, and a synopsis of client fulfillment in regards to the proposed framework.

Research Design

This paper aims to investigate the different factors which affect the choice of content selection for a Netflix user based on the recommendation list that is available to the consumer. This study is descriptive in nature based upon a structured questionnaire which was used to collect primary data from a sample of 120 respondents in the National Capital Region of India. The primary data is collected from the respondents using a convenience sampling method. The dataset is collected in the time frame of January 2020 to May 2020. Further, Exploratory Factor Analysis (EFA) has been conducted to evaluate the factors influencing the user behaviour in content selection and to evaluate the factors which contribute to the selection of content for the recommendation engine. Additionally, a linear regression analysis has been performed to investigate the variables influencing the user behaviour based on recommendation engine using the EFA scores for factors influencing the movie choice from the recommendation engine as the dependent variable and the EFA scores for the factors which influence the content suggested by the recommendation engine. R Programming is used to conduct the analysis and the hypothesis testing on the sample dataset.

Research Question, Objective & Hypothesis

The objective of this study is to evaluate the various factors that affect user behaviour, factors that contribute to the formation of a recommendation list based on which the recommendations are made. And, to investigate if there is any relation between the recommendation list structure and the user behaviour on content selection.

RQ: Does the recommendation list structure have a direct impact on the user behaviour?

H0: The recommendation list structure does not have a direct impact on the user behaviour.

H1: The recommendation list structure does have a direct impact on the user behaviour.

The data was collected through the Google Forms comprising twenty-three questions in total (8 Ranking Questions, 7 Rating Scale Questions, 1 Multiple Choice Question, 2 Yes or No Questions, and 5 Questions based on demographics.) The above twenty-three questions have been divided into three categories, i.e. Demographics, Perception and Satisfaction, and finally, recommendation engine. The respondents were contacted through various social media platforms Facebook, Twitter, LinkedIn, etc. A sample survey of 120 respondents was considered. Their responses were interpreted and analyzed with the help of graphical representations. To conduct the study of the primary data as collected, the researchers utilized various descriptive statistics tools to gain a better insight into the collected data. Further, linear regression was applied to the collected data as an inferential statistics tool to conduct hypothesis testing.

DATA ANALYSIS AND INTERPRETATION

Demographic Analysis

Table 1: Demographic Profile for the questionnaire data collected

Demographic profile	Scale	Number of respondents (Frequency)	Response Rate (%)
Age	15 or below	3	02.50
	Between 16 – 26	97	80.93
	Between 27- 37	13	10.83
	Between 38-48	5	04.17
	48 and above	2	01.67
Gender	Male	63	52.50
	Female	57	47.50
Education	Diploma	43	35.83
	Undergraduate	54	39.17
	Postgraduate	23	25.00

Source: Authors' Calculations

As indicated in Table 1, the sample data majorly consists of 80.83 percent of respondents from the age group of 16-26 years old and almost evenly spread between males and females. 45 percent of the respondents have completed or are pursuing their undergraduate degree while the others are either pursuing higher education or have completed their postgraduate degrees.

Value of Recommendation Systems according to Users

Regarding Netflix's recommendation system, respondents were asked to evaluate and rate two given statements based on their perception of the value of their Netflix access: "Recommendation system is somehow accurate in predicting my interests" and "I value this feature and want to continue to use it". On a scale, where 1 is "Strongly disagree", 3 is "neutral", 5 is "strongly agree", participants rated to what extent they agree with the statement.

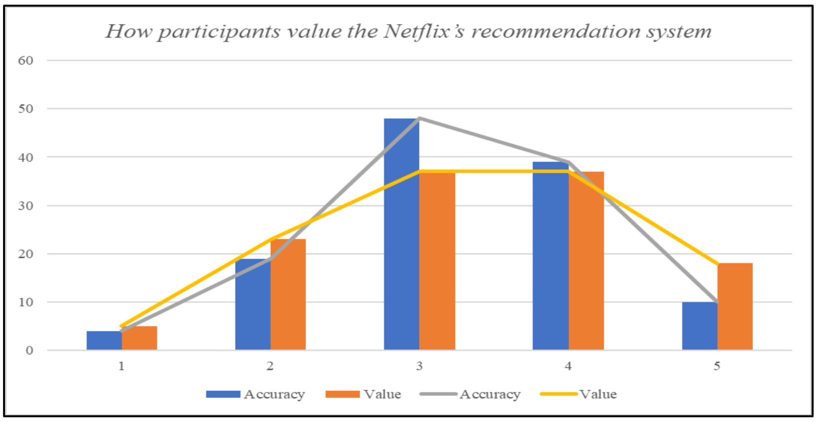
Table 2: How participants value Netflix's recommendation system

Descriptive tool	Mean	Median	Mode	Standard Deviation	Sample Variance	Kurtosis	Skewness
Accuracy of Recommendations	3.266	3	3	0.941	0.886	-0.194	-0.191
Value of Recommendation Engine	3.333	3	4	1.079	1.165	-0.685	-0.171

Source: Authors' Calculations

For the first statement, 120 respondents gave a 3.26 mean rating which was slightly lower than the average of 3.33 for the second statement. Also, the recommendations predicting accuracy had a relatively low standard deviation of 0.941 compared to 1.079.

Figure 11: Distribution of how Netflix's users value its recommendation system



Source: Authors' Compilation

Figure 11 shows the graphical distribution for two statements. The users' rating for the accuracy parameter shows that the left side of the tail was longer than the right side thus it's a left-skewed distribution. The concentration of mass of the distribution for the data was more in the middle to the right side of the diagram. A similar trend has also been shown for the data distribution for the value perception by the users. The medians for both the measures were smaller than the means (3 versus 3.26; 3 versus 3.33 respectively).

Hypothesis Testing

In the research question & hypothesis section, the research paper proposed a research Question along with one hypothesis. To inspect the projected hypothesis, an online survey was conducted using Google Forms. In the collected data, the following items were classified as shown in table 3.

Table 3: Variable Definition for the questionnaire data collected

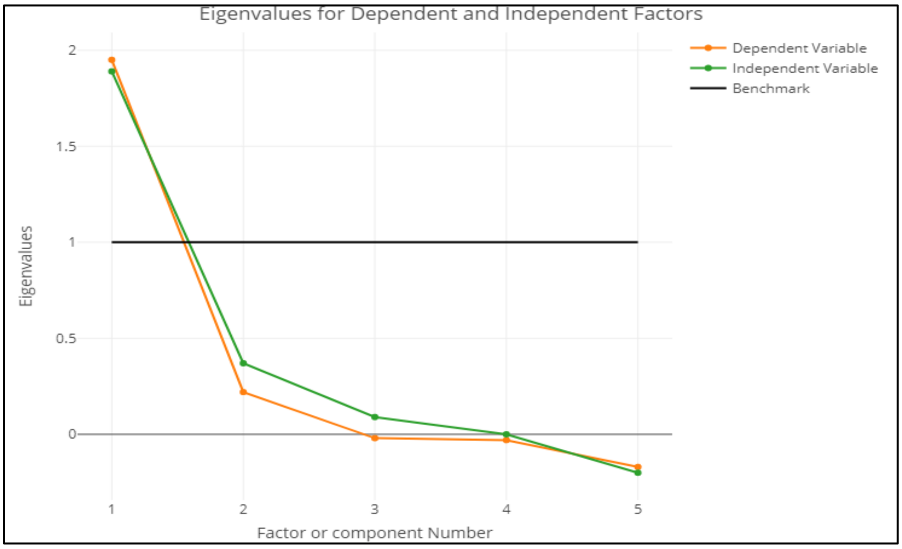
Variable and Type	Data Nomenclature	Literature Review
Dependent Variables		
How you usually choose a new show to watch after finishing current viewing content:		
Based on Netflix's recommendations	Choice_Rec_Rank	Yan, 2017
Word of mouth (include social media)	Choice_WoM_Rank	
Traditional media promotion	Choice_MediaPromo_Rank	
Online media promotion	Choice_OnlinePromo_Rank	
Film review aggregator (IMDB, Rotten Tomatoes, Metacritic, etc)	Choice_Critic_Rank	
User Behaviours (EFA)	User Behaviour	
Independent Variables		
The genre setting has a preference in the recommendation by Netflix	Genre_Setting	Ansari, Essegaier, & Kohli, 2000 and Carlos A & Neil, 2015
Demographic Values of viewers influence the recommendation list given by Netflix	Demographic_Value	
Video Content Critics influence the recommendations by Netflix	Expert_Critics	
The reviews of viewers influence the recommendation by Netflix	User_Review	
The Customer Heterogeneity indicators influence the recommendations by Netflix	Cust_Hetro	
The Model Heterogeneity indicators influence the recommendations by Netflix	Movie_Hetro	
Recommendation List Structure (EFA)	Recommendation List Structure	

Source: Authors' Compilation (based on literature)

Two EFA were performed by employing five items of the Likert scale for the dependent variable and six items of the Likert scale for the independent variable. The overall reliability for both the EFAs and their respective components were evaluated based on Cronbach's alpha benchmarks. Table 4 shows the results of factor loadings, variation explained and reliability analysis. The first factor explaining the factors affecting the choice made by the user for the content that they select to watch from the recommendation list has the reliability statistics of 0.715 and the second factor which explains the factors which influence the content suggested by the recommendation engine has a reliability score of 0.733. These values indicate that the variables have high internal consistency which is above the critical value of 0.70 (DeVellis, 2012). Additionally, the Kaiser-Meyer-Olkin (KMO)'s Measure of Sampling Adequacy (MSA) was 0.75 and 0.73 for Factor 1 and Factor 2 respectively, thus supporting the data adequate for factor analysis (Kaiser & Rice, 1974).

Bartlet's test was conducted to evaluate whether the correlation matrix is an identity matrix or not. The results for Bartlet's test showed the chi-square value to be 327.75 with a p value of 1.86e-40, which is below the threshold level of 0.05 therefore concluding that the data is significant for EFA (Bartlett, 1951; Gorsuch, 2014). The first factor explains the variation by 39 percent and the second factor explains the variation by 31.5 percent, therefore explaining a total variation of 70.5 percent. Cattell's scree test was used to investigate the significant number of factors needed for the analysis. As per the general rule of thumb, only those factors with >1 should be considered important for analysis purposes (Brown, 2015).

Figure 12: Scree Plot



Source: Authors' Compilation

Table 4: Factor Loadings, Variation Explained and Reliability statistics for the Factors created

Factor	Variables	Factor Loadings	Variation Explained	Cronbach Alpha
Dependent Variables	Choice_Rec_Rank	0.646	0.390	0.715
	Choice_MediaPromo_Rank	0.799		
	Choice_OnlinePromo_Rank	0.773		
User Behaviour	Choice_Critic_Rank	0.541		
Independent Variable	Genre_Setting_Rank	0.567	0.315	0.733
	Demographic_Value_Rank	0.600		
	Expert_Critics_Rank	0.599		
Recommendation List Structure	User_Review_Rank	0.560		
	Cust_Hetro_Rank	0.516		
	Movie_Hetro_Rank	0.564		

***Note:** Choice_WoM_Rank was removed from EFA as the factor loading was below the benchmark of 0.5 (Hair *et al.,* 2010).

Source*:* Authors' Calculations

Based on the scree plots, the eigenvalues suggest taking 1 factor for both the fctors. Table 4 represents the factor loading of each variable to the extracted factors. The factor loadings are partial correlation coefficients of factors to their respective items. And since all the factors which are more than the benchmark of 0.5 they can be considered as practically significant (Hair *et al.,* 2010)

Table 5: Regression Analysis summary for Linear Model between User Behaviour and Recommendation List Structure

Dependent Variable	**Independent Variable**	**Standardized Coefficients**	**t**	**Sig.**	**R square**	**F value**
		Beta				
User Behaviour	Constant	5.808e-17	0.000	1.000	0.09047	12.84
	Recommendation List Structure	0.3296	3.583	0.000495		

Source: Authors' Calculations

Regression Analysis

Table 5 indicates the regression analysis results with an overall p-value for the model being 0.000495 which is lower than the level of significance at 95 percent, confidence interval (i.e. 0.05) making the overall model significant and valid. Therefore, we will reject the null hypothesis and accept the alternate hypothesis which concludes that the choice of content by the users is significantly impacted by the recommendation engine and the structure. The model explains 9.04 percent of the variation with the F value of 12.84. The model indicates that the recommendation list and the structure based on which it is formed, like the genre settings, demographic values, expert critics, user reviews, customer, and heterogeneity have a significant impact on the way a user selects the content that they prefer to watch and select from the recommendation list. Further, the user behaviour with regards to the choice of content is influenced by the rank of any content on the recommendation list, media promotion, online promotion, critic ranks, and word of mouth.

CONCLUSION AND IMPLICATIONS

The research study had an equal distribution for the sample which consists of users who have a high frequency of usage of the Netflix platform consisting of 84.11 percent of respondents who were recurrent streaming consumers who used the platform once a week regularly. Youthful respondents displayed a higher response rate than the older participants. In addition, word-of-mouth and online promotional activities are the main influencing factors for users in their decision-making process on the content selection and discovery of new content. Word of mouth was identified to be the most preferred method of new content discovery above all the other modes. In the rapidly developing importance and significance of the impact of Social Media and Online promotions, they play an important role in how streaming users are introduced and influenced to watch new content. When the respondents were requested to rank the most preferred method to discover new content and influencing factors for their viewing decisions, word-of-mouth recommendations made by their peers they are acquainted with and have an aptitude of trust with, was regarded as the best way to select new content. Based on investigation, it was revealed that the recommendation system was not the most favored cause for respondents in new content discovery. And, was to be classified as the second most used way for the consumers to make their next watching decision. Although a large portion of the total respondents mentioned that they have watched content from their personalized recommendation list and 84.17 percent of the respondents browsed through the list regularly. The research concludes that the recommendation system of Netflix affects the consumer's decisions based upon the recommendations as made by the Netflix platform. The recommendation engine is affected by the genre settings, the demographic data, expert evaluations, user interactions, customer heterogeneity as well as movie heterogeneity. The recommendations made by Netflix's system are utilized by a majority of the consumers and this feature is valued by these users of the platform too. This paper can be insightful to all the readers and managers with a keen interest in understanding customer preference and how they can leverage the use of machine learning and clustering technology to extract deeper insights from the data points that are collected. This paper can give a different perspective and can help them understand the process of the influencing factors for their customer's decision-making process and how they can benefit from the insight into manufacturing a better product that targets their audiences' needs and wants with an attempt to boost their product or service's value for the customer. Using Machine Learning algorithms and clustering coupled with Natural Language Processing (NLP) can help any analytical research project gain a deep enormous understanding of the target audience. The study will provide an individual with the general view of the watching habits of the Netflix users. Armed with the results gathered, we are able to determine how our prototype might be linked with their current behavior.

SCOPE FOR FUTURE STUDY

Viewing the results of this paper there are several proposals for future exploration. Due to the impediment of small sample size, it would be worthwhile for impending examinations to administer the review to a bigger sample size and more muddled socioeconomics based on nationality. Accomplishing a worldwide viewpoint on how worldwide clients contemplate the proposed framework is likewise fundamental as Netflix is in the midst of fast worldwide development.

REFERENCES

1. Ansari, A., Essegaier, S., & Kohli, R. (2000). Internet Recommendation Systems. Journal of *Marketing Research*, 37(3), 363–375. doi:https://doi.org/10.1509%2Fjmkr.37.3.363.18779.

2. Bartlett, M. S. (1951). The effect of standardization on a χ 2 approximation in factor analysis. *Biometrika*, 38 (3/4), 337–344.Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

3. Codecademy. (2020). Netflix Recommendation Engine. Retrieved from https://www.codecademy.com/articles/how-netflix-recommendation-works-data-science

4. DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed). California: Sage publications

5. Gomez-Uribe, C. A., & Hunt, N. (2015). The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4), 1-19.

6. Gorsuch, R. L. (2014). Exploratory factor analysis. New York: Routledge.

7. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate data analysis. New Jersey: Prentice Hall.

8. Haton, J. P. (2006). A brief introduction to artificial intelligence. *IFAC Proceedings* Volumes, 39(4), 8-16.

9. Havens, T. (2018). Netflix: Streaming channel brands as global meaning systems. In *from Networks to Netflix* (pp. 321-331). Routledge.

10. Heale, R., & Forbes, D. (2013). Understanding triangulation in research. *Evidence-based nursing*, *16*(4), 98-98.

11. Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, 34 (1), 111–117.

12. Kotu, V., & Deshpande, B. (2019). Recommendation engines. *Data Science* (Second Edition), 343-394

13. Liftigniter. (2016, March 15). The Power of Netflix's Personalized Video Recommendations. Retrieved from https://www.liftigniter.com/netflix-personalizedvideo

14. Maxxor. (2017). Netflix's Cultural Influence on Global Culture. (https://www.maxxor.com/) Retrieved from https://www.maxxor.com/blog/netflix-cultural-influence-global-culture/

15. Pilipets, E. (2019). From Netflix Streaming to Netflix and Chill: The (Dis) Connected Body of Serial Binge-Viewer. *Social Media + Society*, 5(4), 2056305119883426.

16. Rose, J. A. (1988). Expert Systems by A. Bonnet, JP Haton, JM Truong-Ngoc Prentice Hall, New York, 1988 (£ 11.95). *Robotica*, *6*(4), 343-344.

17. Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5), 513-523

18. Shapiro, S. C. (1992). *Encyclopedia of artificial intelligence second edition*. John.

19. Thurmond, V. A. (2001). The point of triangulation. *Journal of nursing scholarship*, 33(3), 253-258.

20. Victor. (2007, February 14). Netflix's Competitive Advantage. (Victor, Ed.) *Internet, Organizations*, 1(1). Retrieved from http://noisebetweenstations.com/personal/weblogs/?p=2011

21. Vom Brocke, J. and Sinnl, T. (2011), "Culture in business process management: a literature review", *Business Process Management Journal, 17 (2)357-378.* https://doi.org/10.1108/14637151111122383

22. Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly,* xiii-xxiii.

23. Wildemuth, B. M. (Ed.). (2016). *Applications of social research methods to questions in information and library science*. ABC-CLIO.

24. Yan, J. (2017). *Identifying Online Streaming User Value in the Netflix Recommendation System.* Drexel University.